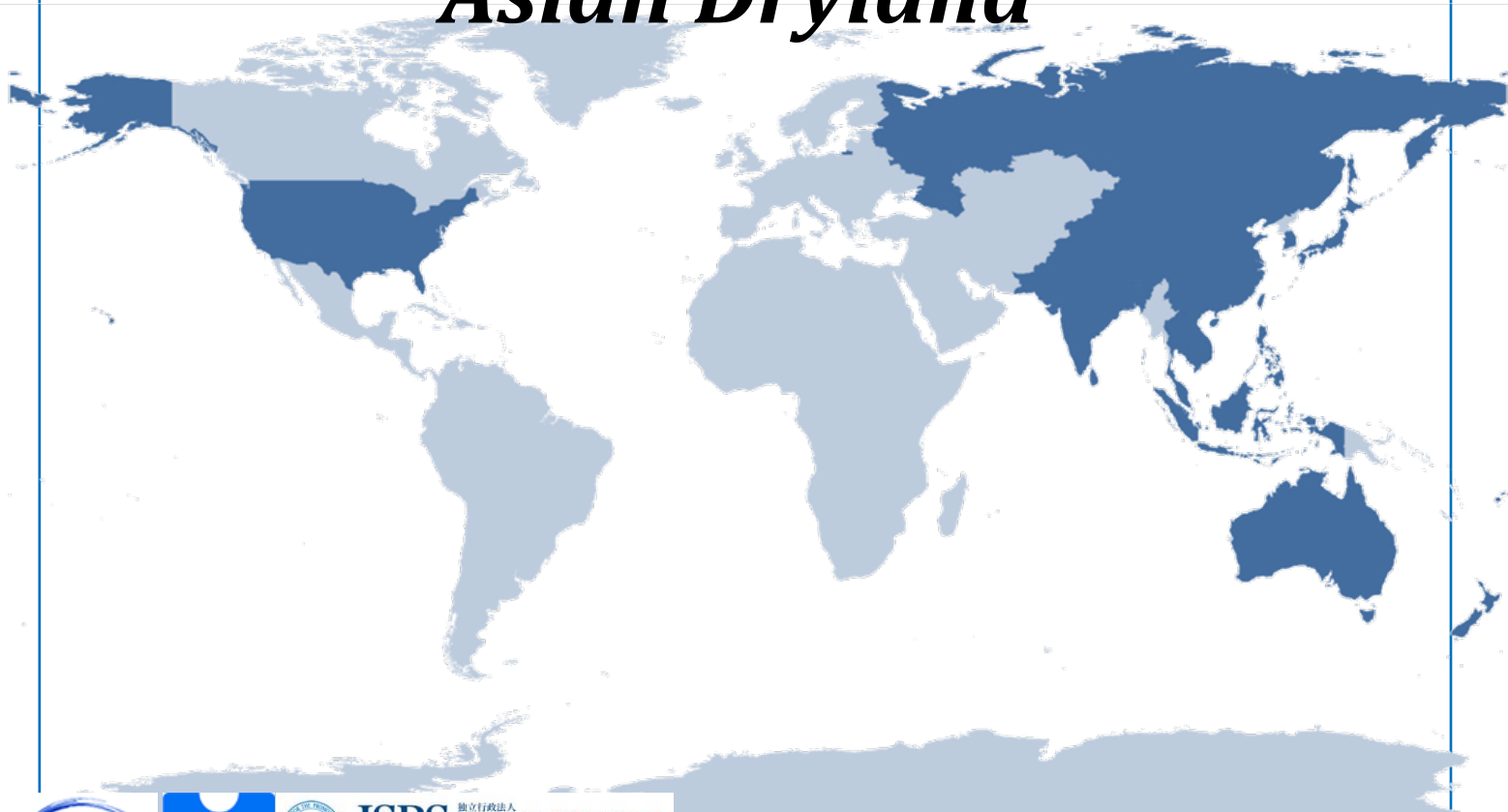


“Intercomparison of Landsurface Process Modelling at Asian Dryland”



The following collaborators worked on this project:

Jun Asanuma, TERC, University of Tsukuba, Japan, asanuma@suiri.tsukuba.ac.jp

Dennis Ojima, Colorado State University, USA, Dennis.Ojima@colostate.edu

Ailikun, MAIRS-IPO, Chinese Academy of Science, China, ailikun@gmail.com

Kazuaki Yorozu, Hydro. & Water Resour. Res. Lab, Kyoto Univ., Japan, yorozu@hywr.kuciv.kyoto-u.ac.jp

“Intercomparison of Landsurface Process Modelling at Asian Dryland”

Project Reference Number: ARCP2011-03CMY-Asanuma
Final Report submitted to APN



OVERVIEW OF PROJECT WORK AND OUTCOMES

Non-technical summary

For better prediction of future climate at Asian dryland by using climate models, a community of Asian scientists, including young scientists, was build to inter-compare numerous landsurface process models. These models are subprograms of climate models to reproduce water, energy and vegetation processes at landsurface, and in this project they are used independently, i.e. “off-line”, of climate models. Asian Dryland Model Intercomparison Project, ADMIP, is an international project driven by this community. The data needed to drive the models at the target sites were collected and archived: One of the finest data set in Asian dryland was constructed. Then, these data were used to drive a number of state-of-the-art numerical models of landsurface processes. The outputs generated by these models were also archived, and they were subject to mutual comparison. This way of comparison is called model intercomparison, and shed a light into differences in reproducibility of landsurface processes caused by different coding between the models. This gave an insight into current status of modeling skills of landsurface processes at Asian dryland.

Objectives

The main objectives of the project were:

1. To increase modeling capacity in Asian countries of landsurface phenomena at dryland
2. To build a “land surface modelling community” in Asia for future cooperation

Amount received and number years supported

The Grant awarded to this project was:

US\$ 57,700 for Year 1:

US\$ 47,100 for Year 2:

Activity undertaken

Three major international workshops and 1 interim meeting. Numerous conversations over the project mailing lists.

2010

- | | |
|------------|--|
| May 23 | A small meeting was held at National Olympics Memorial Youth Center |
| Jun. 22 | A presentation was given by J. A. to introduce ADMIP at HESSS2, held at Tokyo Univeristy , Komaba, Tokyo. |
| Jul. 11-12 | The APN-MAIRS joint workshop on ADMIP (ADMIP kick-off meeting) was held at Beijing Foreign Expert Hotel, Beijing, China |
| Jul. 13-16 | The 2nd Summer School on Land Surface Observing, Modelling and Data Assimilation at Beijing Normal University, Beijing, China |
| Jul. | The project home page was opened at http://hywr.kuciv.kyoto-u.ac.jp/admip/index.html hosted by Kyoto University |
| Jul. | The mailing list among the project members (admip@suiri.tsukuba.ac.jp) were made, under the support of Terrestrial Environment Research Center, University of Tsukuba |

2011

- | | |
|---------|--|
| Feb. 10 | A task force meeting was held at Katsura Campus of Kyoto University, Japan |
|---------|--|

- | | |
|------------|---|
| Jul. 13-14 | The 2nd workshop on ADMIP was held at Ning Wo Zhuang Hotel, Lanzhou, China |
| Dec. 1-2 | ADMIP interim meeting was held at Hokkaido University, Sapporo, Japan |
| 2012 | |
| May 16-18 | The 3rd ADMIP workshop was held at Sylvan Dale Guest Ranch, Loveland, Colorado, USA |

Results

In order to improve climate predictions of landsurface processes at Asian dryland, Asian scientists, including young scientists, were gathered to start a project called ADMIP, Asian Dryland Model Intercomparison Project, which aims to inter-compare numerous landsurface process models, that are subprograms of climate models to reproduce water, energy and vegetation processes at landsurface. The project protocol, which states the details of model runs and the methods of intercomparison of the model outputs, was composed through a long discussion among the project members. By following the protocol, then, the data needed to drive the models were collected, archived and used to drive land surface models (LSMs), which computes energy and water balance at the earth surfaces, and terrestrial ecosystem models (TEMs), which takes care of carbon uptake and intake by the ecosystem at the land surface. The outputs of these models were also archived and subject to mutual comparison to shed light into differences between the models caused by different schemes and different modelling approaches. While all of the models selected in this project are among the most elaborated, the preliminary results revealed by this project suggested that skills of current climate models to reproduce water and ecosystem processes are still under its developing stage at Asian dryland. More data and more elaboration on the testing of the models are needed. The current effort is following the similar efforts in Europe, Australia and America, but one of the first in Asia. The driving data and the model outputs archived by this project will surely serve as a basis for future works, as will do with the researchers community build in this project. The current project also worked as a capacity building of the young scientists, which will help promote further modelling works.

Relevance to the APN Goals, Science Agenda and to Policy Processes

As already stated, this project aims to increase research capacities in Asian countries, and to build research communities for Asian dryland research. These aims relates to three of the four specific areas cited in APN scientific agenda, 2010-2015; namely, climate change and climate variability, ecosystems, biodiversity and land use; changes in the terrestrial domains. Moreover, it contributes specifically to APN Goal 1, supporting regional cooperation, and Goal 3, improving scientific capabilities of Asian nation.

Self evaluation

This project aims to increase modeling capacity of drylands in Asian countries, and will seek to build a “land surface modeling community” in Asia for future cooperation. About ten researchers, who constantly attended the meetings and the conferences, and who, therefore, are listed as the participants of the smallest Sapporo meeting in Dec. 2011, were the core members of the project. A community of Asian researchers to conduct a model intercomparison was successfully constructed. At the beginning of the project, few among them are experienced with the model intercomparison works. Now that they have become well-experienced, their capacity has been expanded. For these reasons, the project fulfills the initial aims very well. I am sure that this project will be a basis of the further modeling works in Asia.



Potential for further work

During this project, the long-term research plan, called here as “Project protocol”, was constructed for the period of 5 years or longer. We are in the middle of the tasks stated in this plan: About half of the tasks in the protocol were already finished, while the others ahead are left to be finished in the next two to three years.

Publications (please write the complete citation)

- Asanuma, J. and Yorozu, K., 2010: Introduction to ADMIP, Asian Drylands Landsurface Modeling Intercomparison Project, a paper presented at Hydrology delivering Earth System Science to Society (HESSS2), Jun. 22, 2010, Tokyo, Japan.
- Asanuma, J., 2011: Processing Data from KBU, a paper presented at APN-MAIRS joint workshop on Intercomparison of Land Surface Process Modelling at Asian Dryland, Jul. 13-14, 2011, Lanzhou, China.
- Miyazaki, S., Yorozu, K., Asanuma, J., and Kondo, M., 2012: Reproducibility of hydrological processes with a land surface model at semi-arid grassland, Proc. 2012 Annual Conf. Japan Soc. Hydrol. & Water Resour., pp.154-155.
- Yorozu, K. and others, 2012: Comparison of ADMIP model results on stage 0.5 and stage 0.5 re-run, a paper presented at The 3rd International Workshop on Asian Dryland Model Intercomparison Project (ADMIP), 16-18, May, 2012, Loveland, Colorado, USA.

References

- Henderson-Sellers, A., Pitman, A. J., Love, P. K., Irannejad, P. & Chen, T. H., 1995: The Project for Intercomparison of Land Surface Parameterization Schemes (PILPS): Phases 2 and 3, *Bulletin of the American Meteorological Society*, 76, 489-503
- Henderson-Sellers, A., Yang, Z.-L. & Dickinson, R. E., 1993: The Project for Intercomparison of Land-surface Parameterization Schemes, *Bulletin of the American Meteorological Society*, 74, 1335-1349
- Lee, X., Massman, W., and Law, B., 2004: Handbook of Micrometeorology, Kluwer Academic Pub., pp.250.
- Manabe S. 1969: Climate and the ocean circulation: 1, the atmospheric circulation and the hydrology of the Earth's surface. *Monthly Weather Review*, 97, 739–805.
- Oki, T. & Kanae, S. 2006: Global Hydrological Cycles and World Water Resources, *Science*, 313, 1068-1072
- Pitman, A. The evolution of, and revolution in, land surface schemes designed for climate models, *International Jour. Climatology*, 2003, 23, 479-510.
- Qian, T., Dai, A., Trenberth, K. E. & Oleson, K. W., 2006: Simulation of global land surface conditions from 1948 to 2004. Part I: Forcing data and evaluations, *J. Hydrometeorology*, 2006, 953-975.

Acknowledgments

Asia-Pacific Network for Global Change Research, APN-GCR, and Monsoon Asia Integrated Regional Study, MAIRS, were two major funding sources of ADMIP, without whose financial support our activities were not be possible. Japan Society for the Promotion of Science, JSPS, is another through its Grant-in-Aid for Scientific Research (B). In addition, logistic and/or in-kind supports were also provided by Cold and Arid Regions Environmental and Engineering Research Institute, CAREERI, CAS, and Global COE program “Establishment of Center for Integrated Field Environment Science” at Hokkaido University, Colorado State University,

when a meeting or a workshop was held. The ADMIP home page and its mailing lists were hosted by Hydrology and Water Resources Research Laboratory at Kyoto University, and Terrestrial Environment Research Center, TERC, at University of Tsukuba, Japan, respectively. Hydrology and Water Resources Research Laboratory at Kyoto University also provided support in preparing the data for model runs. TERC and MAIRS-IPO mainly provided various in-kind supports to the project. The concept of ADMIP was first proposed by Prof. Congbin Fu at Institute of Atmospheric Physics, CAS, China, and Prof. Michael Manton at Monash University, Australia, continuously various supports to the project.



Preface

Researchers from Japan, China, USA, Korea, Mongolia and so on, gathered to conduct an international project called, ADMIP (Asian Dryland Model Intercomparison Project), to inter-compare land surface process models. This reports its activities at the end of the financial term of APN-GCR.

Table of Contents

ACTIVITY UNDERTAKEN	4
RESULTS.....	5
POTENTIAL FOR FURTHER WORK.....	6
PUBLICATIONS (PLEASE WRITE THE COMPLETE CITATION).....	6
REFERENCES	6
ACKNOWLEDGMENTS	6
PREFACE	8
TABLE OF CONTENTS	8
1.0 INTRODUCTION.....	9
1.1 WHY DRYLAND ?.....	9
1.2 WHAT IS LANDSURFACE PROCESS MODELLING?	10
1.3 WHAT IS MODEL INTERCOMPARISON?	10
1.4 TARGET PHENOMENA AND SCIENTIFIC QUESTIONS	11
1) TARGET PHONEMENA	11
2) SCIENTIFIC QUESTIONS	11
1.5 PURPOSE OF THIS PROJECT?.....	11
2.0 METHODOLOGY.....	11
2.1 TARGET SITES.....	11
A. TONGYU (CHINA).....	12
B. KHERLEN BAYAN ULAAN (KBU, MONGOLIA).....	12
C. PINGLIANG (CHINA)	12
2.2 INTERCOMPARISON STAGES (BLIND POLICY)	12
1) STAGE 0: ANALYSES OF EXISTING DATA SET.....	13
2) STAGE 0.5: W/"DEFAULT" PARAMETER SET	13
3) STAGE 1: W/"OBSERVED" PARAMETER SET:	13
4) STAGE 2: W/"CALIBRATED" PARAMETER SET:.....	13
2.3 REGISTERED MODELS	13
2.4 MODEL INPUT/OUTPUT VARIABLES AND INPUT PARAMETERS.....	15
1) INPUT VARIABLES (AS TIME SERIES)	15
2) INPUT PARAMETERS (AS FIXED VALUE IN TIME)	15
3) OUTPUT VARIABLES.....	16

2.5 DATA ISSUES	16
1) DATA SHARING POLICY.....	16
2) DATA FORMAT	16
3) DATA PREPARATION.....	16
4) DATA SUBMISSION.....	18
2.6 TIMELINE OF THE PROJECT	20
1. TIMELINE OF ACTIVITIES.....	20
2. MEETINGS AND CONFERENCES.....	20
3.0 RESULTS & DISCUSSION	25
3.1 DATA SET CONSTRUCTED.....	25
3.2 MODEL INTERCOMPARISON.....	27
4.0 CONCLUSIONS	29
5.0 FUTURE DIRECTIONS	29
REFERENCES	29
APPENDIX.....	31
LIST OF YOUNG SCIENTISTS.....	31
GLOSSARY OF TERMS.....	32
ORGANIZATIONS	32
OTHERS	32
ACRONYMS OF PARTICIPATING MODELS	32
ACRONYMS FOR MODEL VARIABLES (ADOPTED FROM ALMA CONVENTION)	33

1.0 Introduction

1.1 Why dryland ?

Dryland account for 40% of the earth's land surface and also a similar fraction of Asian land surface (FAO, UNCCD, Figure 1). Characterized by dry climate, low vegetation cover and low nutrient, its ecosystem and the society that depends thereon, have inherently large vulnerability to the external perturbations, such as climate change and land use change. IPCC AR4 predicts drier climate in the extratropical arid regions during the latter half of this century, and overgrazing and/or land use change are triggering desertification and land degradation in Asian drylands, especially in the transitional zones (Fu, 2009). Because it concerns the poorest population groups, essentially living off threatened natural resources, the desertification and land degradation are challenges for achieving the Millennium Development Goals (UNCCD).

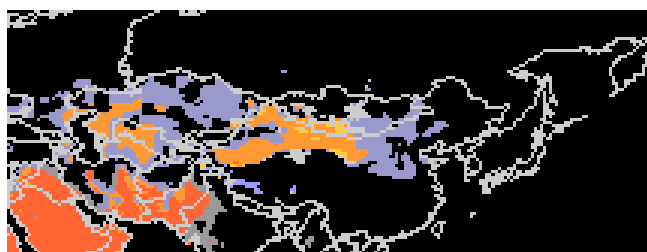


Figure 1.1: Dryland in Asia(FAO)

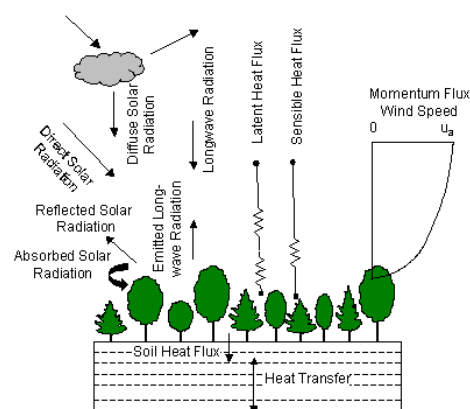


Figure 1.2: Schematic figure of processes in LSMs



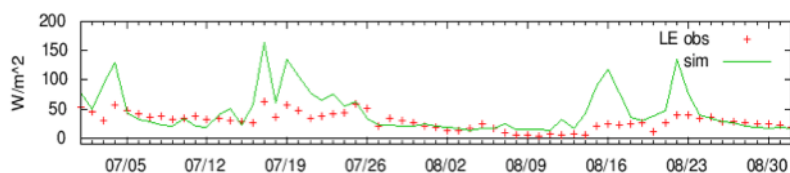


Figure 1.3: An example of LSMs simulation results of evaporation at Mongolian grassland (line) compared with the observations (cross).

1.2 What is landsurface process modeling?

In order to facilitate sustainable land management in Asian drylands, increasing our predictive capability of land surface processes is indispensable. This can be achieved through improvement of land surface process modelling, that refers, in this proposal, to land surface models (LSMs, Figure 1.2) and terrestrial ecosystem model (TEMs), each of which are briefly described below.

Land surface models (LSMs) were initially developed as a sub-module of GCMs (general circulation model) to incorporate energy and water exchange at the land surface into global atmospheric simulation (Manabe 1969). Since mid 1980's, LSMs have been developed separately from those of their parent GCMs. Now, most sophisticated LSMs in the world compute processes such as heat and water budget for different land cover/use, transpiration and carbon assimilation of plants, river discharge, groundwater, urban environment, and even usage and control of river water by human beings (Pitman, 2003; Oki and Kanae, 2006).

Terrestrial ecosystem models (TEMs) has been also developed in conjunction with GCM simulation, particularly to incorporate response of ecosystem to the elevated CO₂ concentration in the projected future. TEMs specialize in carbon dynamics, such as plant growth, ecological succession and soil organic processes at longer time scale larger than, e.g., months, while LSMs, that major in water and energy at shorter time scales such as hours and days, take vegetation as a static component. Though these models are initially developed for a sub-module of GCMs as stated, they can be also used independently of GCMs (offline), and therefore can serve as a modelling tool of land surface environment at the regional and point scale.

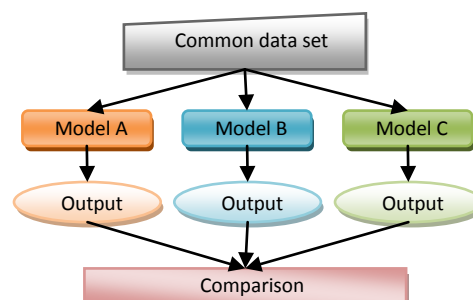


Figure 1.4: Schematic diagram of model intercomparisons

1.3 What is model intercomparison?

The current most-sophisticated LSMs and TEMs utilize latest scientific knowledge on hydrosphere, atmosphere and biosphere, and use state-of-the-art computer simulation technique. Nevertheless, uncertainties in the model results are, in general, large, and the origins of these uncertainties are still a matter of a research target. In particular, it is already known that ability of these models to reproduce land processes at Asian dryland surface processes are still limited (Figure 1.3), partly because they are not sufficiently tested at short grass vegetation in Asian dryland, where plant activities are strongly regulated by water availability. In other words, Asian dryland is one of largest gap region of land surface modelling works in the world. Note that most of LSMs and TEMs are initially developed for tropical and boreal forests as a major target. In addition, until recently, the observed data



that is needed to drive and validated these models were not easily available from Asian drylands.

The model intercomparison (Henderson-Sellers et. al, 1993, 1995; Figure 1.4) is a way toward improvement of LSMs and TEMs, where multiple models are run with a common driving meteorological data set, and the results are compared each other to identify relative performance of each model as well as ensemble characteristics common to all of the models tested. This facilitates intensive tests of LSMs at the targeted site as well as best-estimate of model uncertainties from the inter-model variations.

1.4 Target Phenomena and Scientific Questions

1) Target Phenomena

Target phenomena that will be tested through the intercomparison of LSMs and TEMs are, energy, water and carbon exchange at the land surface under water-limited environment, and/or under the control of vegetation growth, at the temporal scales of 3-hourly, daily, and monthly.

2) Scientific Questions

Scientific questions that will be answered through the intercomparison of TEMs and LSMs are as follows

- 1) What is the ability of models to reproduce energy, water, and carbon exchange at dryland surface (reproducibility)?
- 2) Do current ecosystem models, developed mainly for forest, reproduce energy, water and carbon exchange?
- 3) Does the current complexity of models effectively simulate land surface processes? (complexity issues, also related to benchmarking)
- 4) Does the multi-model ensemble of LSMs outperform single model output? (ensemble)

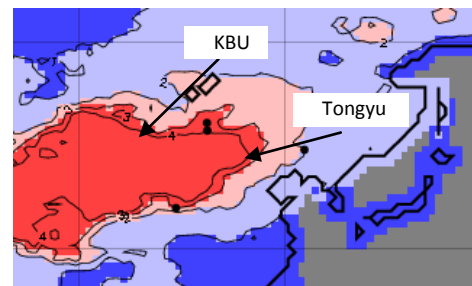


Figure 1.5: Target sites. Contours and colors indicate radiative dryness index of Budyko.

1.5 Purpose of this project?

In the proposed project, several LSMs and TEMs will be intercompared by using the common data recently obtained at Asian dryland. LSMs and TEMs will be selected from the candidates and their revisions widely used in both research and operational institutes in the world. Two selected site will provide meteorological data to drive these models. These sites started their observations within this century, and have continued the observations to acquire high-quality and continuous data that are needed to drive LSMs and TEMs in the intercomparison. All of these circumstances made the proposed works possible, only recently.

Through these intercomparison works, **this project aims to increase modelling capacity of drylands in Asian countries, and will seek to build a “land surface modelling community” in Asia for future cooperation.**

2.0 Methodology

2.1 Target Sites

At the beginning, three target sites were chosen for the intercomparison, mainly with the data availability. The observed data at these sites will be archived and used in the



intercomparison.

a. Tongyu (China)

contact person: Dr. Ailikun (MAIRS-IPO, ailikun@gmail.com)

target period: 2003-2009

note: GEWEX-CEOP registered

http://www.eol.ucar.edu/projects/ceop/dm/insitu/sites/ceop_ap/Tongyu/Cropland

b. Kherlen Bayan Ulaan (KBU, Mongolia)

contact person: Dr. Jun Asanuma (U. Tsukuba, asanuma@suiri.tsukuba.ac.jp)

target period: 2003-2007

note: GEWEX-CEOP and IGBP-AsiaFlux registered

http://www.eol.ucar.edu/projects/ceop/dm/insitu/sites/ceop_ap/Northern_Mongolia/Kherlen

http://asiaflux.yonsei.ac.kr/network/021KBU_1.html

c. Pingliang (China)

contact person: Dr. Jun Wen (CAREERI, jwen@lzb.ac.cn)

future candidate

2.2 Intercomparison stages (Blind policy)

In general, a model will perform better when it is provided with more information about the target location it is applied. However, different models will perform differently with different degree of information they are given about the target. Some model may perform well even when it does not have any geographical information. Another model may improve its performance magnificently better than others, when it is given some information.

In order to evaluate model differences with respect to how they depends on the amount of information they are given about the geographical, three stages with different degree of information given to the models were defined, as an experimental design of the intercomparison.

In the stage 0.5, no information will be given to the models, except time series of meteorological variables and the location of the sites, i.e. longitude and latitude. This is called a “blind” stage. The performance of the model when nothing is known will be tested at this stage. In the following stage 1, the models will be given information about the target site, such as soil and vegetation properties. These are the information the models are given when they are run within the climate models. stage 2 will provide the models with data for calibration, that are, the answers.

In summary, stage 0.5 is designated to differentiate models with their performance when they are “blind”, while the models are evaluated with the same degree of information as in the climate models under stage 1.0. Stage 2 will compare the model performance when they are fully “tuned”.



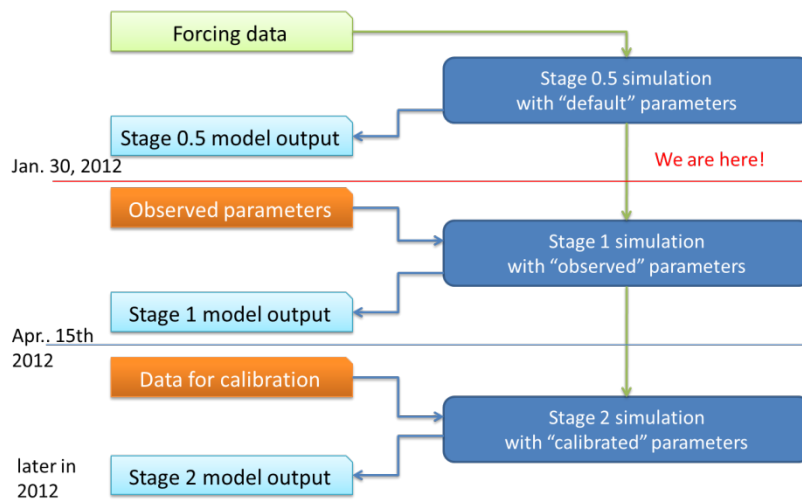


Fig. 2.1: Schematic diagram for the blind policy in ADMIP

1) Stage 0: analyses of existing data set

Basic investigation of the results from previous intercomparison or analyses. e.g., GSWP 1 and 2, GLDAS, and CEOP-MOLTS, will be made to evaluate current overall performance.

2) Stage 0.5: w/"default" parameter set

"Totally-blind" stage. No observed information is available at this stage. The parameter values without any measured knowledge will be used. The decision is left to the model operators. Data being provided will be

for LSMs, climatological values of LAI derived from MODIS
for TEMs, none

3) Stage 1: w/"observed" parameter set:

"Half-blind" stage. Parameters in the models will be determined from observations if reliable information is available. Otherwise, the decision is left to the model operators. All of the model operators will be requested to report their parameter selection. Data being provided will be:

for LSMs, LAI (from MODIS); albedo; vegetation type & properties including root profiles; soil type and properties
for TEMs, soil type and properties

4) Stage 2: w/"calibrated" parameter set:

"Calibration" stage. Calibration will be performed by tuning the model results with the observations. All of the model operators will be requested to report their parameter selection, or how they calibrated the parameters. The objective function for this calibration will be a matter of future discussion. Data being provided for calibration will be, in addition to those listed in stage 1,

for LSMs: energy, water and carbon fluxes
for TEMs: energy, water and carbon fluxes; above- and below-ground biomass; LAI

2.3 Registered models

Below are the list of the models registered to the ADMIP project, and the researchers in charge of the models.



BAIM V2.0

Biosphere-Atmosphere Interaction Model ver. 2
Kazuo MABUCHI, Meteorological Research Institute, Japan

JULES 2.0

Joint UK Land Environment Simulator, ver 2
Hong Jinkyu, Korea

SiBCrop

Simple Biosphere Model-Crop, ver 1
Erandi Lokupitiya, USA

CABLE_UNSW

Community Atmosphere Biosphere Land Exchange, ver 1.0
Jason P. Evans, Climate Change Research Centre, University of New South Wales, Australia

CLM 3.5

Community Land Model, ver 3.5
Jun Asanuma, Tsukuba University, Japan

CoLM

Common Land Model, ver 3
Guo Weidong, Institute of Atmospheric Physics, Chinese Academy of Sciences, China

MATSHIRO

Minimal Advanced Treatments of Surface Interaction and Runoff, ver 5.6
Shin Miyazaki, Hokkaido University, Japan
Standard version, Groundwater version and Isotope version

Noah 2.7

Noah land surface model, revised ver 2.7
Chen Yingying, Chinese Academy of Sciences, China

SiB 2

Simple Biosphere Model, revised ver 2
Chen Yingying, Chinese Academy of Sciences, China

SiBUC

Simple Biosphere including Urban Canopy, ver 1
Kazuaki Yorozu, Kyoto University, Japan

SM/TEA

Soil model
Zhang Xia, Institute of Atmospheric Physics, Chinese Academy of Sciences, China

Biome-BGC

Biome-BGC
Kazuhito Ichii, Fukushima University, Japan

SEIB-DGVM

Spatially Explicit Individual-Based Dynamic Global Vegetation Model, ver 2.53
Kaoru Tachiiri, Japan Agency for Marine-Earth Science and Technology, Japan



VISIT

Vegetation Integrative Simulator for Trace gases, ver 1
Akihiko Ito, National Institute for Environmental Studies, Japan

SSiB

The Simplified Simple Biosphere model
Qian Li, Institute of Atmospheric Physics, Chinese Academy of Sciences, China

HAL

HAL
Masahiro Hosaka, Meteorological Research Institute, Japan

DayCent

DayCent
Dennis Ojima, Colorado State University, USA

NoahMP

Noah with multiple physics
Guo-Yue Niu, The University of Arizona, USA

2.4 Model input/output variables and input parameters

(note: ALMA variable names in **Bold**)

Model input variables are selected through the investigation of each participating models.

1) Input variables (as time series)

The variables listed below will be prepared as time series from observed data set.

- for LSMs (at every 30min or so)

- **SWdown** and **LWdown** (incoming shortwave & longwave radiation, maybe diffuse/direct)
- **Tair** (air temperature), **Qair** (specific humidity), **CO2air** (CO2 concentration), and **Wind** (wind speed) at specified height
- **Snowf** and **Rainf** (snowfall and rainfall, or their sum)
- **Psurf** (Surface pressure)
- LAI (as time series if available)

- for TEMs (daily, or with longer time interval)

- max. and min. temperature
- precipitation, VPD, and **SWdown**
- current and historical land use (grazing history)

2) Input parameters (as fixed value in time)

The variables listed will be either prepared on the basis of the ground observation or derived from satellite measurements. Any post-observation processing of observed records should be documented.

- for LSMs

- Vegetation type/properties (that can be measured with confidence)
- Soil type/properties (that can be measured with confidence)
- Elevation, slope, aspect



- for TEMs (in addition to the above)

- Soil texture
- Additionally, soil C and N (for validation?)
- Peak live aboveground biomass, root biomass (for validation?) and snow cover
- Soil respiration (for validation?)

3) Output variables

The variables listed below should be submitted to the project for further analyses. Each variable listed below were ALMA compliant. The detail description can be derived from [output variables table](http://hywr.kuciv.kyoto-u.ac.jp/admip/output_table.html) (http://hywr.kuciv.kyoto-u.ac.jp/admip/output_table.html).

- from LSMs (Hourly)

Mandatory

[SWup](#), [LWup](#), [Qh](#), [Qle](#), [Qg](#), [SoilTemp](#) (layer-averaged soil temperature)
[Evap](#), [Qs](#), [Qsb](#), [SWE](#) and/or [SnowDepth](#), [SoilMoist](#), [WatertableD](#)

Mandatory if applicable

[GPP](#), [NPP](#), [NEE](#), [AutoResp](#), [HeteroResp](#)

- from TEMS (Daily)

Mandatory

[Qle](#), [SoilTemp](#)
[SoilMoist](#)
[GPP](#), [NPP](#), [NEE](#), [AutoResp](#), [HeteroResp](#), [ToLivBiom](#)

Mandatory if applicable

[SWup](#), [LWup](#), [Qh](#), [Qg](#)
[Evap](#), [Qs](#), [Qsb](#), [SWE](#) and/or [SnowDepth](#), [WatertableD](#)

2.5 Data issues

1) Data sharing policy

First, all forcing data and model outputs will be shared by the participants of ADMIP. A key paper will be published with all participants as well as the data provider as the authors. At 1 year after the publication of the key paper, forcing and input data that were used to run the models and the model outputs will be available to the public, under the condition that it complies the data policy of the data sources.

2) Data format

ALMA-compliant netCDF format will be used for submitting, archiving and exchanging the forcing and model output data. Specific notes to ADMIP are given below.

sign convention	flux values should be upward positive
leap year issues	365-day calendar should be used, and, therefore, data at Feb 29th will be dropped after the data submission.
the start of the day	00UTC (0 o'clock in UTC)
soil layer depth	information should be submitted in the netCDF file or in a separate text.
soil temperature	SoilTemp in ALMA convention is layer-averaged soil temperature.

3) Data preparation

The data working group will be organized to prepare necessary data. All of submission of



the data will be directed to ADMIP web page.

a. LAI and albedo from MODIS

CAREERI group (Dr. J. Wen) and Mr. Kondo (Fukushima U.) is in charge. LAI (at 500 m resolution) and albedo (at 5km resolution) was and will be derived from MODIS products. At KBU, LAI will be calibrated with in situ observations. 2 data sets is prepared as follows.

1) Climatological monthly LAI from 10-years data for the blind stage 0.5 prepared by Fukushima U.

Table 2.1: Climatological monthly LAI derived from MODIS

Site	KBU	TGY
Jan	0.0	0.0
Feb	0.0	0.0
Mar	0.0	0.0
Apr	0.0	0.0
May	0.23	0.25
Jun	0.28	0.33
Jul	0.52	0.69
Aug	0.56	1.22
Sep	0.37	0.72
Oct	0.2	0.25
Nov	0.0	0.0
Dec	0.0	0.0

It should be noted that an expected value of LAI=0 during winter may not lead to reasonable results, and that it may need to be replaced with an infinitesimally small value, such as LAI=0.1, for some models (Thanks to Prof. Jinkyu Hong).

2) Time series of LAI (Fukushima U.) and albedo (CAREERI) at every 8 days, will be constructed.

b Soil properties

Dr. Yang K.(Tongyu) and Dr. Asanuma (KBU) will derive parameters needed.

c. Grazing history:

Dr. Aili (Tongyu) and Dr. Asanuma (KBU) are in charge. Dr. Byanbahuu also gave details at KBU.

- ✧ Tongyu: heavily overgrazed since 1960's, with increasing grazing pressure upto now. No grazing since 2003 when it was fenced
- ✧ KBU: heavily overgrazed, especially in winter, in summer moderately overgrazed (Dr. Byanbahuu will investigate further)



d Root profile (depth) from the literature (at KBU)

- ✧ Tongyu (10-35cm, no root below 35cm) and KBU(upto 50 cm)

e. Long term forcing data set for spin-up

Long term forcing data set was prepared for spinning up LSMs and TEMs.

- ✧ Tongyu (prepared by Dr. Yang K.)
date set constructed by Institute of Tibetan Plateau Research
period 1981-2008, 3 hourly
- ✧ KBU (prepared by J. Asanuma)
data set constructed by NCAR (Qian et al., 2007, J.Hydro.Met) for Community Land Model.
period 1948-2004, 3 hourly

4) Data submission

1) Data format

ALMA-compliant netCDF format is used for submitting model output data. The detailed description of requested variables can be found in [output variables table](http://hywr.kuciv.kyoto-u.ac.jp/admip/output_table.html). (http://hywr.kuciv.kyoto-u.ac.jp/admip/output_table.html) at ADMIP home page.

Some notes specific to ADMIP are also given in “section 6, 2) Data format”.

2) File naming convention

Each participants will submit their model output to the project for further analysis and basic comparison. To facilitate these data analysis, a convention for file names should be followed:

[model ID]_[stage ID]_[station ID].nc

Where MODEL: model acronym and STATION: station name identifier (such as Tongyu or KBU(Kherlenbayan-Ulaan)

Each modeler use specified model inputs integration of their models and submit that information with their output by filling in the [requested sheet](#).

a Model ID

An acronym will be given to each of participating models as an unique ID. This ID will be used in the file name for the model outputs. It is in the form,

<model acronym><version>[<modified organization>][<modified version>]

ex1) “CLM40”

Community Land Model ver. 4.0, original version as developed at NCAR

ex2) “CLM40UTKB13”

Community Land Model modified at Univ. Tsukuba after the original version of ver. 4.0. Its modified version of 1.3.

Table 2.2: Model acronyms of the participating models

Model name	Model acronym
BAIM (Biosphere-Atmosphere Interaction Model)	BAIM
JULES (Joint UK Land Environment Simulator)	JULES



SiB-Crop (Simple Biosphere Model-Crop)	SiBC
CABLE (Community Atmosphere)	CABLE
CLM (Community Land Model)	CLM
CoLM (Common Land Model)	CoLM
MATSIRO (Minimal Advanced Treatments of Surface Interaction and Runoff)	MATSIRO
MATSIRO-GW (Minimal Advanced Treatments of Surface Interaction and Runoff-Groundwater)	MATSIRPG
Iso-MATSIRO (Minimal Advanced Treatments of Surface Interaction and Runoff-Groundwater)	IMATSIRO
Noah-CAS (Noah land surface model)	NCASA
SiB2-CAS (Simple Biosphere Model)	SCASA
SiBUC (Simple Biosphere including Urban Canopy)	SiBUC
SM (Soil model)	SM
Biome-BGC	BBGC
SEIB (Spatially Explicit Individual-Based Dynamic Global Vegetation Model)	SEIB
VISIT (Vegetation Integrative Simulator for Trace gases)	VISIT
HAL(Hydrology, Atmosphere and Land (HAL) model)	HAL
DayCent	DAYCENT
Noah MP	NOAHMP

b Stage ID

Table 2.3: Stage IDs

Stage no.	Stage ID
stage 0.5	stage05
stage 1.0	stage10
stage 2.0	stage20

c. Station ID

Table 2.4: Station IDs

Sation	Station ID
Tongyu	Tongyu
Kherlen Bayan Ulaan(KBU)	KBU

3) Data upload

Each participants will submit their model output and their default parameters after archiving them into one file at [the data submission page](#).

a. Stage 0.5 submission

Each modeler use default parameters for stage 0.5 integration of their models and submit



that information with their output by filling in the [requested sheet \(http://hywr.kuciv.kyoto-u.ac.jp/admip/doc/default_parameter.xls\)](http://hywr.kuciv.kyoto-u.ac.jp/admip/doc/default_parameter.xls). If you can't upload or upload system doesn't work well, please send your output as attachment.

2.6 Timeline of the project

In order to discuss the methodologies described above, and to prepare the data needed for the intercomparison as well as results of the intercomparison, 4 meetings have been held in conjunction with this project. Full summaries, participants lists, major presentation files, and pictures of these meetings are printed in the appendix or contained in DVD-ROM attached to this report, and therefore only brief descriptions are given herein.

1. Timeline of Activities

2009

July, 22 Possibility of an international project on intercomparison of landsurface process models by using data obtained at Asian dryland was discussed at the 2nd MAIRS International workshop on Asian Dryland Study.

2010

Apr. A proposal submitted to APN-ARCP by J. Asanuma (J.A.) and D. Ojima was approved.

Apr. A proposal submitted to JSPS by J. Asanuma and others was approved.

May 23 A small meeting was held at National Olympics Memorial Youth Center, Yoyogi, Japan, with 5 participants.

Jun. 22 A presentation (Asanuma and Yorozu, 2010) was given by J. A. to introduce ADMIP at HESS2, held at Tokyo University, Komaba, Tokyo.

Jul. 11-12 The APN-MAIRS joint workshop on ADMIP (ADMIP kick-off meeting) was held at Beijing Foreign Expert Hotel, Beijing, China.

Jul. 13-16 The 2nd Summer School on Land Surface Observing, Modelling and Data Assimilation at Beijing Normal University, Beijing, China.

Jul. The project home page was opened at <http://hywr.kuciv.kyoto-u.ac.jp/admip/index.html> hosted by Kyoto University.

Jul. The mailing list among the project members (admip@suiri.tsukuba.ac.jp) was opened under the support of Terrestrial Environment Research Center, University of Tsukuba.

2011

Feb. 10 A task force meeting was held at Katsura Campus of Kyoto University, Kyoto, Japan.

Jul. 13-14 The 2nd workshop on ADMIP was held at Ning Wo Zhuang Hotel, Lanzhou, China.

Dec. 1-2 ADMIP interim meeting was held at Hokkaido University, Sapporo, Japan.

2012

May, 16-18 The 3rd ADMIP workshop was held at Sylvan Dale Guest Ranch, Loveland, Colorado, USA.

2. Meetings and Conferences

The methodologies given in this chapter were discussed over long discussions between the participating researchers. These discussions are made during the series of meetings and conferences, and through the mailing list during the two and half year of this project. Below, these meetings and conferences will be listed and described briefly. Full information will be given in Appendix.



1. A small meeting at National Olympics Memorial Youth Center

Date: May 23, 2010
Place: National Olympics Memorial Youth Center, Yoyogi, Tokyo
Participants: Drs Mabuchi, Itoh, Ichii, Miyazaki, Yorozu, Tachiiri, Asanuma
Summary: Expectations and prospects of ADMIP were discussed. Past experiences of model intercomparison projects were introduced by some of the participants.

2. The APN-MAIRS joint workshop on Asian Dryland Model Intercomparison Project (ADMIP) –“Kickoff Meeting”

Date: July 11-12, 2010
Place: Beijing Foreign Expert Hotel, Beijing, China
Hosted and supported by: MAIRS, CAS, APN and JSPS
Participants: see the list in Appendix
Summary: Current status of the landsurface models and terrestrial ecosystem models were introduced by the participants. Various experiences and applications of these models at dry climate regions were discussed. Possible outcomes of intercomparison of these models at Asian dry land were further discussed. A detailed plan of the intercomparison of these models with use of observations gained at Asian dryland was presented. Possible candidates of the target sites of the project were listed, and 3 sites, Tongyu, KBU, and Pinliang, were chosen from them. The data policy, the time line of the project, and the next meeting place were decided.

Training course: Right after the workshop, a training workshop jointly sponsored by CAS and UNESCO was held at Beijing Normal University, and the young participants of the project also attended the training. This workshop worked as a capacity building of this project. The course materials and the participant list are presented in the Appendix.





Fig. 2.2: 1st APN-MAIRS joint workshop on ADMIP at Beijing Foreign Expert Hotel, Beijing on July 11-12, 2010.



Fig. 2.3: The 2nd Summer School on Land Surface Observing, Modeling and Data Assimilation 13-16 July 2010, Beijing Normal University, Beijing, China

3. The 2nd APN-MAIRS joint workshop on Intercomparison of Land Surface Process Modeling at Asian Dryland

Date: July 13-14, 2011

Place: Ning Wo Zhuang Hotel, Lanzhou, China

Also hosted and supported by :

CAREERI, APN, MAIRS, CAS and JSPS

Participants: see the list in Appendix

Summary The progress of the project was reviewed. It was confirmed that two sites, Tongyu and KBU, are the final selections of the target site of the model application, while Pingliang is a future candidate. A future time line of the project was proposed and agreed, where the model intercomparison works will be separated into 3 stages with different objectives and data set.



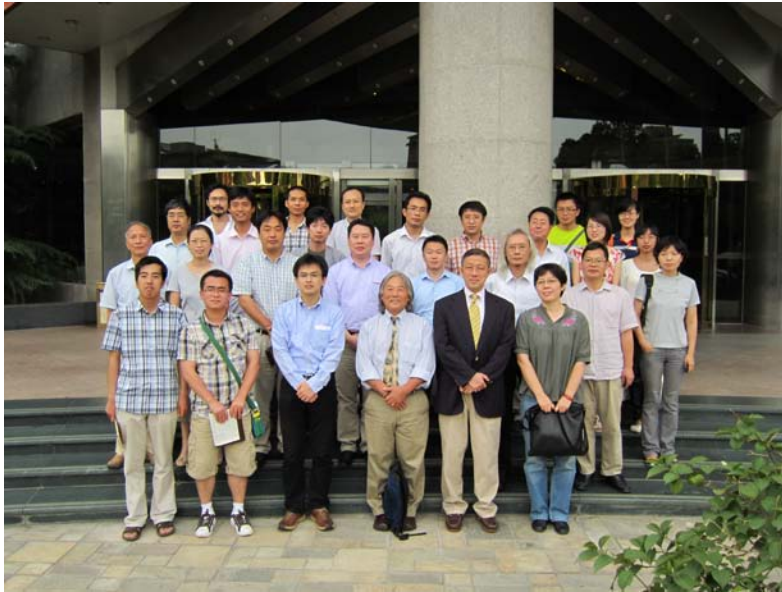


Fig. 2.4: The 2nd APN-MAIRS joint workshop on Intercomparison of Land Surface Process Modeling at Asian Dryland, Lanzhou, China

4. Asian Dryland Model Intercomparison Project (ADMIP) Interim Meeting

Date: Dec 1-2, 2011

Place: Hokkaido University, Sapporo, Japan

Hosted and supported by :

Global COE program “Establishment of Center for Integrated Field Environment Science” at Hokkaido University, APN, MAIRS, CAS and JSPS

Participants: see the list in Appendix

Summary: The progress of the project, especially that of the data preparation needed to drive the models in the intercomparison, were discussed. It was found that the year 2003, the beginning of the target period of the simulation, was found to be an anomalous year, and there are difficulties in generating the initial conditions of the state variables in LSMs and TEMs, such as soil moisture and biomass both above- and below-ground. In order to solve this issue, it was decided that the project prepare historical records of the meteorological data, which can be used in the initializing process of the models (spin-up process). The updated schedule of the project, as well as the related data preparation, was also discussed.





Fig. 2.5: ADMIP interim meeting at Sapporo, Japan

5. The 3rd International Workshop on Asian Dryland Model Intercomparison Project (ADMIP)

Date: May 16-18, 2012

Place: Sylvan Dale Guest Ranch, Loveland, Colorado, USA

Hosted and supported by :

Colorado State University, APN, MAIRS, CAS and JSPS

Participants: see the list in Appendix

Summary: The current status of the project was reviewed. The first results of the intercomparison between the models at the target sites were presented. A plan to write a research paper for a publication in a peer-reviewed journal was proposed, and agreed. The future plans of the project were also discussed.





Fig.2.6: 3rd APN-MAIRS joint workshop on ADMIP at Colorado, USA

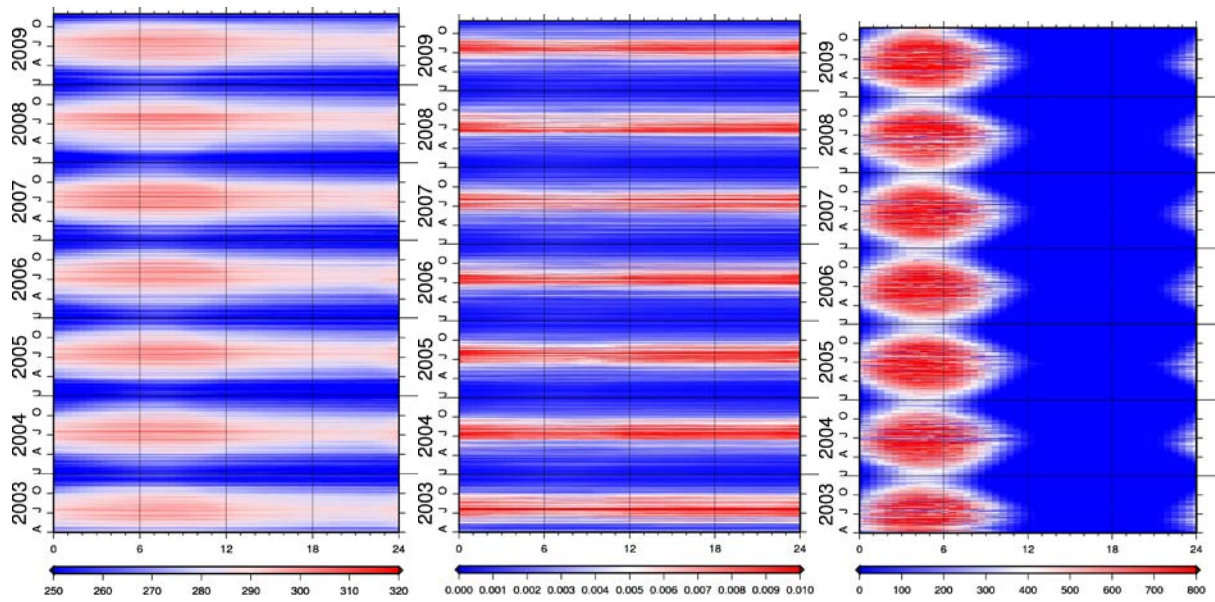
3.0 Results & Discussion

3.1 Data set constructed

A forcing data set, that can drive LSMs and TEMs, were derived from the observational data at Tongyu and KBU. This process includes quality-controlling (QC), which evaluate the quality of each observational value to remove less-qualified data points, and gap-filling (GF) which fill in the gaps generated in the process of QC by mainly using statistical techniques (Lee, Massman, and Law, 2004). Through these processes, continuous, i.e. without any gaps, and well-qualified data series with high temporal resolution, i.e. every 30 minutes, were generated for KBU and Tongyu. An example of these data at KBU are plotted in Fig. 3.1.

In addition, in order for constructing an initial state of the model variables (spin-up_, such as vegetation amount and soil moisture, a historical record of meteorological variables were derived for each of the target sites.

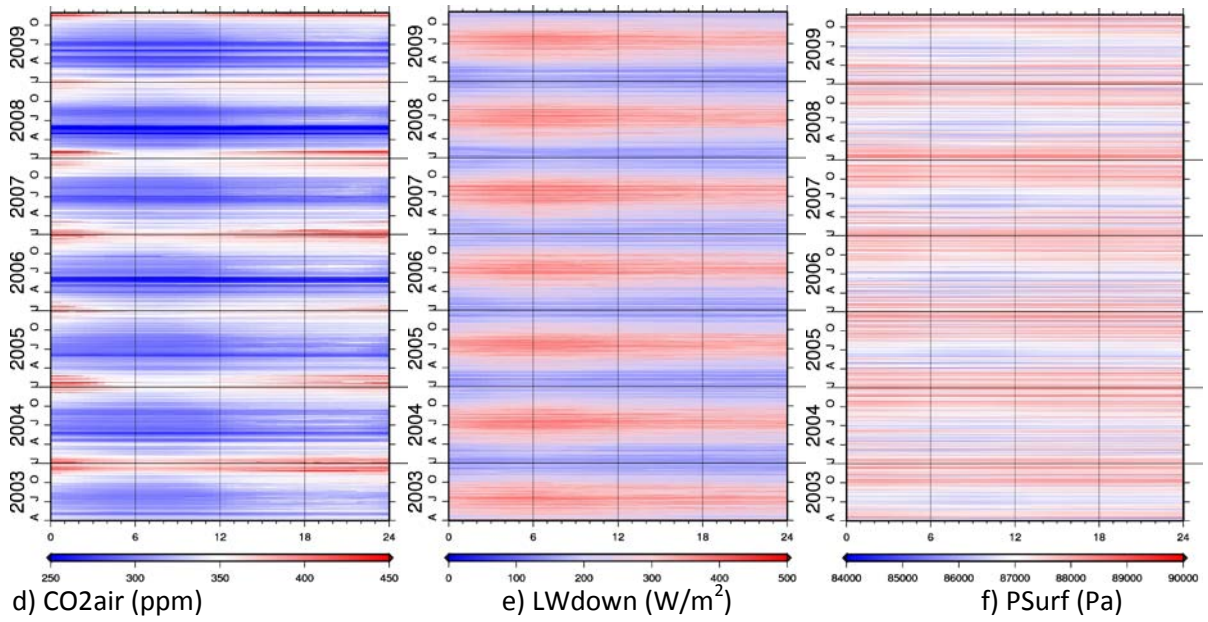




a) Tair (K)

b) Qair (fraction)

c) SWdown (W/m²)



d) CO2air (ppm)

e) LWdown (W/m²)

f) PSurf (Pa)



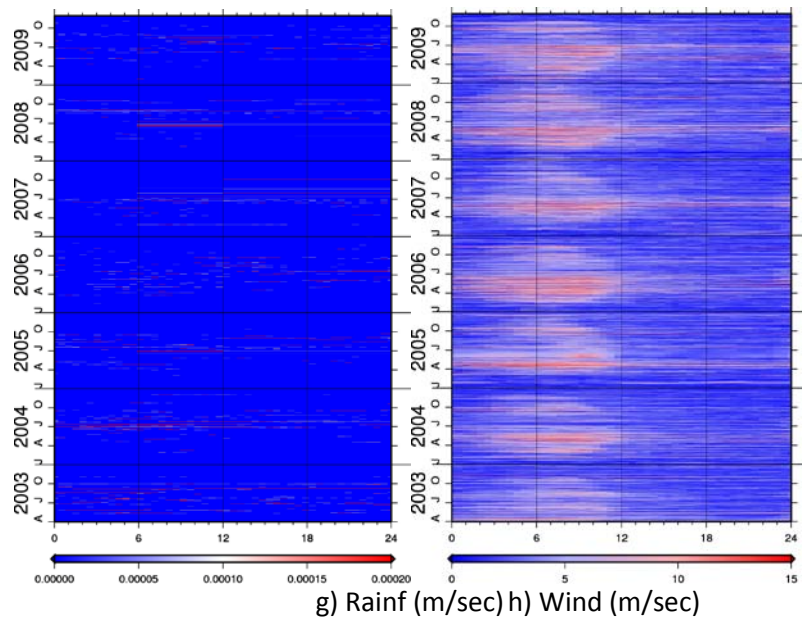


Fig 3.1: Example of the forcing data generated for KBU. The x-axes indicate time of the day, while the y-axes are days. Subtitles are in the ALMA convention, and the units are presented with the subtitles. (Asanuma, 2011)

3.2 Model Intercomparison

Table 3.1 summarizes the status of model output submission at the time of the 3rd workshop in Colorado in May, 2012. “Sign” indicates the sign of the values may be opposite. Though we strictly ask participants to follow ALMA convention, an international standard among the model intercomparison community, some of the model outputs are not compliant with it.

Fig 3.2 and 3.3 show the comparison between monthly model outputs of Q_{le} , latent heat flux, and Q_h , sensible heat flux, from the participating models. It is found that differences between models are larger with Q_h than with Q_{le} . This is in contrast to general knowledge about the physical process in the landsurface modeling. While sensible heat flux from the surface to the atmosphere, that is the heating of air by the land, is mainly concerned only with the temperature of the land and the atmosphere, latent heat flux from the ground, that is the evaporation from the land into the air also involves moisture of the land. Therefore, modeling of the former is regarded as easier than that of the latter. However, the results shown in Figures 3.2 and 3.3 are opposite: The model works better with evaporation than with atmospheric heating. This is probably due to characteristics of dry climate: Evaporation from the landsurface, i.e., latent heat flux, is fully restricted by small amount of rainfall at the monthly time scale, therefore it does not have larger degree of freedom, therefore it is easier to be predicted. On the other hand, sensible heat is controlled by the balance of heat between the air and the land, and has larger degree of freedom, in contrast.



Table 3.1. Output submission status of the participating models, summarized at the 3rd workshop in May, 2012 (Yorozu et. al., 2012).

	SWup	LWup	Qle	Qh	Qg
BAIM	ok	ok	ok	ok	ok
BBGC			ok		
DAYCENT			stage 0.5		
SEIB		ok	ok		
VISIT			ok		
HAL	stage 0.5	stage 0.5	stage 0.5	stage 0.5	stage 0.5
NCASA	ok	ok	ok	ok	ok
MATSIRO	ok	ok	ok	ok	sign
SCASA	ok	ok	ok	ok	sign
SM	ok	ok	ok	ok	sign
SSiB2	ok	ok	ok	ok	ok
CLM	Still data handling				
CoLM	Still data handling				

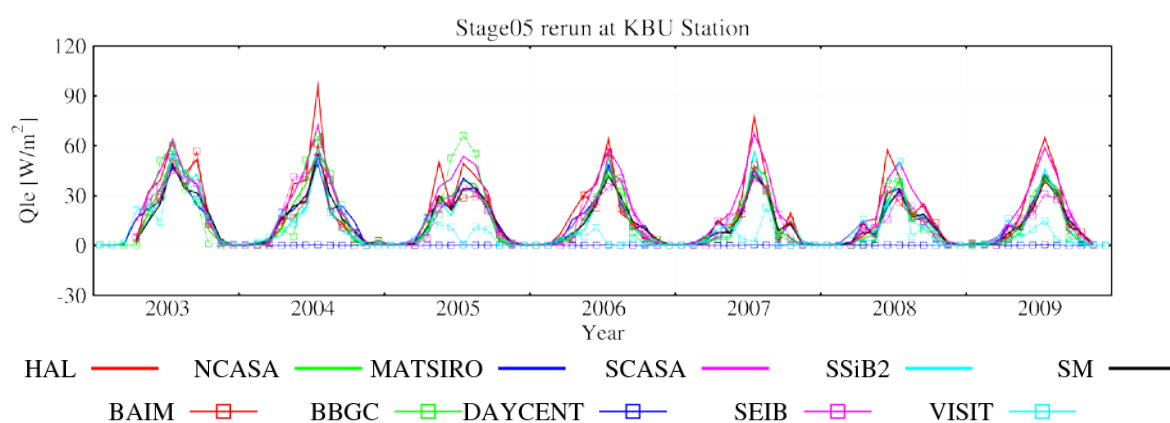


Fig 3.2: Comparison between Qle, latent heat flux, computed by different models (Yorozu et. al., 2012).

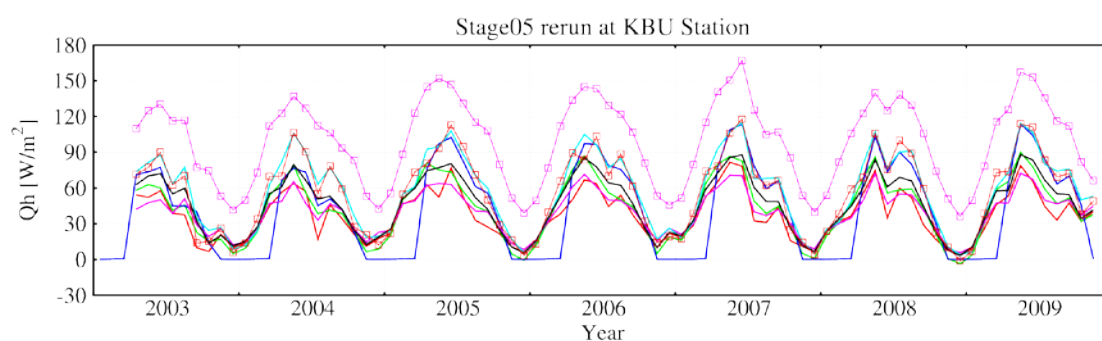


Fig 3.3: Same as Fig.3.2 but with comparison between Qh (Yorozu et. al., 2012).



4.0 Conclusions

The purpose of this APN funded project is to increase modelling capacity at dryland in Asian countries, and will seek to build a “land surface modelling community” in Asia for future cooperation.

Asian scientists, including young scientists, were gathered to build an international project called ADMIP, Asian Dryland Model Intercomparison Project, which aimed to inter-compare numerous landsurface process models. The project protocol of ADMIP, which states the details of model runs and the methods of intercomparison of the model outputs, was composed from scratch through a long discussion among the project members. A part of the project protocol was given in the section 2.0 “Methodologies”.

At the time of the beginning of the project, not many were experienced with the model intercomparison works. Now, the project protocol, that describes the model intercomparison methodologies, was composed from scratch by this community, and that the first model intercomparison was done. Therefore, it is able to conclude that our capacity was expanded, and that “the community” was constructed, which assures the successful end of the project.

5.0 Future Directions

While the project fully satisfied its first goal of the project, it is still at the middle of the tasks stated in the project protocol; about half of the tasks was finished, while the others ahead are left untouched. At the time of writing, the project finished the stage 0.5 comparison, which is outlined in section “3.0 Results & discussion”, and it is about to start with stage 1.0 works. Therefore, stage 1.0 and 2.0, with further intercomparison, are surely next to do.

References

- Asanuma, J. and Yorozu, K., 2010: Introduction to ADMIP, Asian Drylands Landsurface Modeling Intercomparison Project, *a paper presented at Hydrology delivering Earth System Science to Society (HESSS2)*, Jun. 22, 2010, Tokyo, Japan.
- Asanuma, J., 2011: Processing Data from KBU, *a paper presented at APN-MAIRS joint workshop on Intercomparison of Land Surface Process Modeling at Asian Dryland*, Jul. 13-14, 2011, Lanzhou, China.
- Henderson-Sellers, A., Pitman, A. J., Love, P. K., Irannejad, P. & Chen, T. H., 1995: The Project for Intercomparison of Land Surface Parameterization Schemes (PILPS): Phases 2 and 3, *Bulletin of the American Meteorological Society*, 76, 489-503
- Henderson-Sellers, A., Yang, Z.-L. & Dickinson, R. E., 1993: The Project for Intercomparison of Land-surface Parameterization Schemes, *Bulletin of the American Meteorological Society*, 74, 1335-1349
- Lee, X., Massman, W., and Law, B., 2004: Handbook of Micrometeorology, Kluwer Academic Pub., pp.250.
- Manabe S. 1969: Climate and the ocean circulation: 1, the atmospheric circulation and the hydrology of the Earth's surface. *Monthly Weather Review*, 97, 739–805.
- Miyazaki, S., Yorozu, K., Asanuma, J., and Kondo, M., 2012: Reproducibility of hydrological processes with a land surface model at semi-arid grassland, *Proc. 2012 Annual Conf. Japan Soc. Hydrol. & Water Resour.*, pp.154-155.
- Oki, T. & Kanae, 2006: S. Global Hydrological Cycles and World Water Resources, *Science*, 313, 1068-1072.
- Pitman, A. The evolution of, and revolution in, land surface schemes designed for climate models, *International Jour. Climatology*, 2003, 23, 479-510.



- Qian, T., Dai, A., Trenberth, K. E. & Oleson, K. W., 2006: Simulation of global land surface conditions from 1948 to 2004. Part I: Forcing data and evaluations, *J. Hydrometeorology*, 2006, 953-975.
- Yorozu, K. and others, 2012: Comparison of ADMIP model results on stage 0.5 and stage 0.5 re-run, *a paper presented at The 3rd International Workshop on Asian Dryland Model Intercomparison Project (ADMIP)*, 16-18, May, 2012, Loveland, Colorado, USA.



Appendix

Project home page:

ADMIP home page hosted by Kyoto University at
<http://hywr.kuciv.kyoto-u.ac.jp/admip/index.html>

Funding sources outside the APN

MAIRS: Cofunding. Almost half of the expenses of the project are covered by MAIRS, in addition to in-kind support by its staffs at International Project Office (IPO).

JSPS: Cofunding. Financial support through its Grant-in-Aid for Scientific Research (B), titled “Intercomparison of landsurface models and terrestrial ecosystem models at dry and semi-dryland” (PI: J. Asanuma)

2010: 3,500,000JPY

2011: 2,200,000JPY

2012: 2,300,000JPY

TERC, University of Tsukuba, Japan:

In-kind supports through 1) hosting the mailing list of the project, 2) providing logistic supports for the conferences and meetings including one personnel, and 3) providing whole financial accountings for whole APN funding.

Hydrology and Water Resources Research Laboratory at Kyoto University:

In-kind support through 1) hosting the project home page, and 2) assisting preparing the data needed for model runs.

CAREERI, CAS, China:

in-kind supports as the local host of the 2nd workshop as well as covering some local expenses.

Global COE program “Establishment of Center for Integrated Field Environment Science” at Hokkaido University:

in-kind support as the local host of the interim meeting.

Colorado State University:

in-kind supports as the local host of the 3rd workshop, as well as covering some meeting expenses.

List of Young Scientists

Prof. Kazuaki Yorozu

Affiliation and Contacts:

Hydrology and Water Resources Research Laboratory
Department of Civil and Earth Resources Engineering
Kyoto University, C1, Nishikyo-ku, Kyoto 615-8540, Japan
Tel: +81-75-383-3363 Fax: +81-75-383-3360
E-mail: yorozu@hywr.kuciv.kyoto-u.ac.jp

Involvement in the project activities

Input/output data managing, land surface process simulation and model comparison analysis

Comments:



First, my science community was expanded because of participating the project. Second, the skill for communication has been enhanced through discussing many topics with participants. I'm sure participating the project would be useful experience.

Masayuki Kondo

Affiliation:

Faculty of Symbiotic Systems Science, Fukushima University, Japan

Contact:

Tel/Fax: +81-24-548-5256

E-mail: redmk92@gmail.com

Comments and involvements:

Involvement in the ADMIP project not only helped me to learn how to succeed in international collaboration, but also gave an opportunity to get acquaintance with both domestic and overseas senior researchers. Discussing science with those people had founded clear confidence on personal level. In the project, my role was to provide outputs and simple analysis results of one of participated ecosystem models. Thankfully to senior scientists, I was also entrusted with preparing climatological LAI for land surface models. From the beginning, this project had been steadily steered towards a high scientific goal, and it has also fulfilled a sound stance of promoting young scientists.

Glossary of Terms

Organizations and projects

CAS	Chinese Academy of Science
MAIRS	Monsoon Asia Integrated Regional Study
CAREERI	Cold and Arid Regions Environmental and Engineering Research Institute
JSPS	Japan Society for the Promotion of Science

Others

HESSS	Hydrology delivering Earth System Science to Society
LSM	Landsurface Model
TEM	Terrestrial Ecosystem Model
ALMA	Assistance for Land-surface Modelling activities

Acronyms of Participating models

BAIM	Biosphere-Atmosphere Interaction Model
JULES	Joint UK Land Environment Simulator
SiB-Crop	Simple Biosphere Model-Crop
CABLE	Community Atmosphere
CLM	Community Land Model
CoLM	Common Land Model
MATSIRO	Minimal Advanced Treatments of Surface Interaction and Runoff)
MATSIRO-GW	Minimal Advanced Treatments of Surface Interaction and Runoff-Groundwater
Iso-MATSIRO	Minimal Advanced Treatments of Surface Interaction and Runoff-Groundwater
Noah-CAS	Noah land surface model
SiB2-CAS	Simple Biosphere Model
SiBUC	Simple Biosphere including Urban Canopy



SM	Soil model
SEIB	Spatially Explicit Individual-Based Dynamic Global Vegetation Model
VISIT	Vegetation Integrative Simulator for Trace gases
HAL	Hydrology, Atmosphere and Land (HAL) model

Acronyms for model variables (adopted from ALMA convention)

Wind_N	Near surface northward wind component
Wind_E	Near surface eastward wind component
Rainf	Rainfall rate
Snowf	Snowfall rate
Tair	Near surface air temperature
Qair	Near surface specific humidity
PSurf	Surface pressure
SWdown	Surface incident shortwave radiation
LWdown	Surface incident longwave radiation
CO2air	Near surface CO2concentration
SWnet	Net shortwave radiation
LWnet	Net longwave radiation
Qle	Latent heat flux
Qh	Sensible heat flux
Qg	Ground heat flux
Qf	Energy of fusion
Qv	Energy of sublimation
Qtau	Momentum flux
Qa	Advective energy
DelSurfHeat	Change in surface heat storage
DelColdCont	Change in snow cold content
Snowf	Snowfall rate
Rainf	Rainfall rate
Evap	Total Evapotranspiration
Qs	Surface runoff
Qrec	Recharge
Qsb	Subsurface runoff
Qsm	Snowmelt
Qfz	Re-freezing of water in the snow
Qst	Snow throughfall
DelSoilMoist	Change in soil moisture
DelSWE	Change in snow water equivalent
DelSurfStor	Change in Surface Water Storage
DelIntercept	Change in interception storage
SnowT	Snow Surface Temperature
VegT	Vegetation Canopy Temperature
BaresoilT	Temperature of bare soil
AvgSurfT	Average surface temperature
RadT	Surface Radiative Temperature
Albedo	Surface Albedo
SWE	Snow Water Equivalent
SWEVeg	SWE intercepted by the vegetation
SurfStor	Surface Water Storage
SoilMoist	Average layer soil moisture
SoilTemp	Average layer soil temperature
SMLiqFrac	Average layer fraction of liquid moisture



SMFrozFrac	Average layer fraction of frozen moisture
SoilWet	Total Soil Wetness
PotEvap	Potential Evapotranspiration
ECanop	Interception evaporation
TVeg	Vegetation transpiration
ESoil	Bare soil evaporation
EWater	Open water evaporation
RootMoist	Root zone soil moisture
CanopInt	Total canopy water storage
EvapSnow	Snow Evaporation
SubSnow	Snow sublimation
SubSurf	Sublimation of the snow free area
ACond	Aerodynamic conductance
PotEvap	Potential Evapotranspiration
ECanop	Interception evaporation
TVeg	Vegetation transpiration
ESoil	Bare soil evaporation
EWater	Open water evaporation
RootMoist	Root zone soil moisture
CanopInt	Total canopy water storage
EvapSnow	Snow Evaporation
SubSnow	Snow sublimation
SubSurf	Sublimation of the snow free area
ACond	Aerodynamic conductance
LWup	Upward long-wave broadband radiation
GPP	Gross Primary Production
NPP	Net Primary Production
NEE	Net Ecosystem Exchange
AutoResp	Autotrophic Respiration
HeteroResp	Heterotrophic Respiration
TotSoilCarb	Total Soil Carbon
TotLivBiom	Total Living Biomass



1.0 The APN-MAIRS joint workshop on Asian Dryland Landsurface Process Model Intercomparison Project (ADMIP) at Beijing in July, 2010

1.1 Agenda

The APN-MAIRS joint workshop on Asian Dryland Land Surface Process Model Intercomparison Project (ADMIP)

11-12 July 2010

Beijing Foreign Expert Hotel, China

draft agenda ver1.1

11 July, 2010 (Sunday)

DAY 1- morning

- 8:30 - 9:00 **Registration**
Opening (Chair: Ailikun)
- 9:00 - 9:05 **Welcome address by M. Manton, vice chair of MAIRS SSC**
- 9:05 - 9:20 **Jun Asanuma: Scope of the meeting**
- 9:20 - 10:00 **Dennis OJIMA: Regional land-atmosphere considerations of arid and semi-arid land systems of Monsoon Asia**
- 10:00 - 10:40 **Jason EVANS: Land surface model evaluation & suggestions for ADMIP**
- 10:40 - 10:55 **Tea break (group photo)**
- Lessons learned from past model intercomparisons (Chair: D. Ojima)**
30 minutes each talk
- 10:55 - 11:25 **William PARTON: Results from the PILPS and other model comparison efforts**
- 11:25 - 11:55 **Zongliang YANG: TBD**
- 11:55 - 12:25 **Akihiko ITOH: Intercomparison of terrestrial ecosystem models for clarifying uncertainties in carbon cycle**
- 12:25 - 12:55 **YANG Kun:**
- 12:55 - 13:15 **Masayuki KONDO: Lessons learned from Japan-MIP**
- 13:15 - 14:40 **Lunch**

DAY 1- afternoon

- Current Issues in Asian Dryland and its Modeling Perspective (Chair: J. Evans)**
20 minutes each talk
- 14:40 - 15:00 **Moshin IQBAL**
- 15:00 - 15:20 **Purevjav GOMBOLUDEV**



Participating models & Past Intercomparison 1 (Chair: J. Evans)

20 minutes each talk

15:20 - 15:40 Guo-Yue NIU

15:40 - 15:55 Kazuaki YOROZU

15:55 - 16:15 Jun WEN

16:15 - 16:30 Tea break

16:30 - 17:30 Discussion 1: Key processes and questions to be answered
(Chair: J. Asanuma)

18:00 - 18:30 Reception Banquet by MAIRS

12 July, 2010 (Monday)

DAY 2- morning

Participating models & Past Intercomparison 2 (Chair:
Zongliang YANG)

20 minutes each talk

9:00 - 9:20 Kazuo MABUCHI

9:20 - 9:40 Xia ZHANG

9:40 - 10:00 Kaoru TACHIIRI

10:00 - 10:20 Hisashi SATO

10:20 - 10:40 Jia YANG: Dynamic Land Ecosystem Model (DLEM) --
Data-model intercomparison at LBA and NACP Sites

10:40 - 11:00 Tea break

Introduction to target site candidates (Chair: K. Mabuchi)

30 minutes each talk

11:00 - 11:30 Renjian ZHANG: Tongyu Station

11:30 - 12:00 Jun ASANUMA: Kherlen Bayan Ulaan Station

12:00 - 13:30 Lunch

DAY 2- afternoon

13:30 - 15:30 Discussion 2: Target sites & data, and project protocol (Chair: J.
Asanuma)

1) target sites, 2) time lines and mile stones, 3) data policies & formats,
4) data center, etc



1.2 List of Participants

**APN-MAIRS joint workshop on Asian Dryland LandSurface Process Model
Intercomparison Project (ADMIP)
11-12 July 2010, Beijing, China**

List of Participants

Dr. Ailikun Director, MAIRS-IPO Institute of Atmospheric Physics, Chinese Academy of Sciences (IAP), 40# Huayanli, Qi Jia Huo Zi, Chaoyang District Beijing 100029, China Tel. +86-10-82995264 Fax +86-10-82995161 E-mail: aili@mairs-essp.org	Institute of Meteorology and Hydrology Juulchiny gudamj 5, Ulaanbaatar 46, Mongolia Tel: 976-11-326606 Fax: 976-11-326614 E-mail :p_gombo@hotmail.com
Dr. Jun ASANUMA Assoc. Professor Terrestrial Environment Research Center Tsukuba University Voice&Fax: 029-853-6704 E-mail: asanuma@sui.tsukuba.ac.jp	Dr. Xiaodan GUAN College of Atmospheric Science Lanzhou University 222#Tianshui Southroad, Lanzhou, 730000, P.R.China Tel: +86-931-8913139 E-mail: guanxd05@lzu.cn
Mr. Gankhuyag BATJARGAL Forecast Research Laboratory, Institute of Meteorology and Hydrology Juulchiny gudamj 5, Ulaanbaatar 46, Mongolia E-mail: batjargal2006@yahoo.com	Prof. Jinky HONG Global Environmental Lab, Yonsei University Seoul, Korea 120-749 Tel: +82-2-2123-7576 Fax: +82-11-9987-3897 E-mail: jkhong@yonsei.kr
Dr. Yingying CHEN Academy of Sciences Shuangqing Rd. 18 Haidian District Beijing 100085, China E-mail: chenyy@itpcas.ac.cn	Prof. Kazuhito ICHII Faculty of Symbiotic Systems Science, Fukushima University 1 Kanayagawa, Fukushima, 960-1296, Japan Tel: +81-24-548-5256 Fax: +81-24-548-5256 E-mail: ichii@sss.fukushima-u.ac.jp
Prof. Jason EVANS University of New South Wales Sydney, NSW, 2052, Australia Tel: +61-2-9385 7066 Fax: +61-2-9385 7123 E-mail: jason.evans@unsw.edu.au	Mr. Muhammad IJAZ Global Change Impact Studies Centre (GCISC), National Centre for Physics (NCP) Complex, Quaid-i-Azam University Campus, Shahdara Road, Islamabad - 44000, Pakistan Tel: +92-51-2077300/Ext 457, 455 Mobile: +92-300-5315021 E-Mail: m.ijaz@gcisc.org.pk ; muhammad.ijazmalik@gmail.com
Dr. Purevjav GOMBOLUDEV Climate Sec.	



Dr. Akihiko ITO
Center for Global Environmental Research,
National Institute for Environmental Studies
16-2 Onogawa, Tsukuba, Ibaraki 305-8506,
Japan
Tel: +81-29-850-2981
Fax: +81-29-850-2960
E-mail: itoh@nies.go.jp

Dr. M. Mohsin IQBAL
Head, Agriculture Section
Global Change Impact Studies Centre (GCISC)
National Centre for Physics (NCP) Complex;
Quaid-i-Azam University Campus
P.O. QAU-45320, Islamabad -44000, Pakistan
Cell: +92-321-5171386
E-mail: mohsin.iqbal@gcisc.org.pk;
gcisc@comsats.net.pk;
drmohsiniqbal@gmail.com

Dr. Minseok KANG
Department of Atmospheric Sciences, Yonsei
University
Sinchon-dong 134, Seoul, Korea, 120-749
E-mail: ms-kang@yonsei.ac.kr

Dr. Hyojung KWON
Department of Atmospheric Sciences, Yonsei
University
Sinchon-dong 134, Seoul, Korea, 120-749
E-mail: hyojungkwon@yonsei.ac.kr

Dr. Mingxing LI
RCE-TEA
Institute of Atmospheric Physics
Chinese Academy of Sciences
P.O. Box 9804 Beijing 100029
P. R. CHINA
Tel: +86-10-82995138
Fax: +86-10-82995135
E-mail: limx@tea.ac.cn

Prof. Xin LI
Cold and Arid Regions Environmental and
Engineering Research Institute,

Chinese Academy of Sciences, Lanzhou,
730000
Tel: +86-931-4967249
Email: lixin@lzb.ac.cn

Ms. Xiaolu LING
College of Atmospheric Science
Lanzhou University
222#Tianshui Southroad,
Lanzhou, 730000, P.R.China
E-mail: lingxl08@lzu.cn

Prof. Zhuguo MA
RCE-TEA
Institute of Atmospheric Physics
Chinese Academy of Sciences
P.O. Box 9804 Beijing 100029
P. R. CHINA
Tel: +86-10-82995047
Fax: +86-10-82995135
E-mail: mazg@tea.ac.cn

Dr. Kazuo MABUCHI
Geochemical Research Department,
Meteorological Research Institute
1-1 Nagamine, Tsukuba-city, Ibaraki 305-0052
Japan
Tel: +81-29-853-8722
Fax: +81-29-853-8728
E-mail: kmabuchi@mri-jma.go.jp

Prof. Michael MANTON
Vice Chair of MAIRS SSC
School of Mathematical Sciences,
Monash University, Australia
Tel: (03) 9669 4444;
E-mail: michael.manton@sci.monash.edu.au

Dr. Guo-Yue NIU
Department of Geological Sciences
The University of Texas at Austin
Department of Geological Sciences
Austin, TX 78712. Yongkang Xue
Work: (512) 471-5355



1.3 Presentation slides



ADMIP
Asian Dryland Landsurface Modeling
Intercomparison Project
KICKOFF MEETING

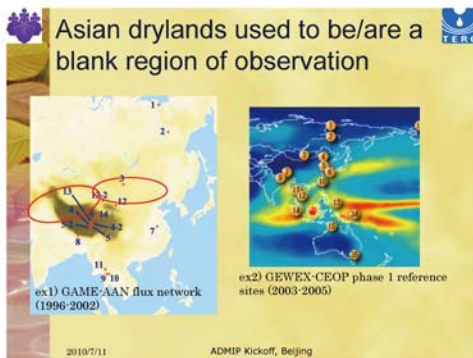
Jun Asanuma
Terrestrial Environment Research
Center
University of Tsukuba

2010/7/11 ADMIP Kickoff, Beijing JSPS 日本学術振興会



MOTIVATIONS

2010/7/11 ADMIP Kickoff, Beijing




Asian drylands used to be/are a blank region of observation

ex1) GAME/AAN flux network (1996-2002)

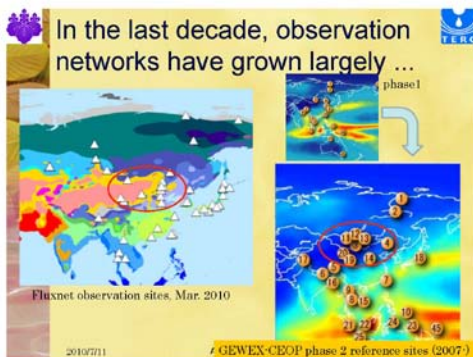
ex2) GEWEX-CEOP phase 1 reference sites (2003-2005)

2010/7/11 ADMIP Kickoff, Beijing



What are the consequences?

2010/7/11 ADMIP Kickoff, Beijing

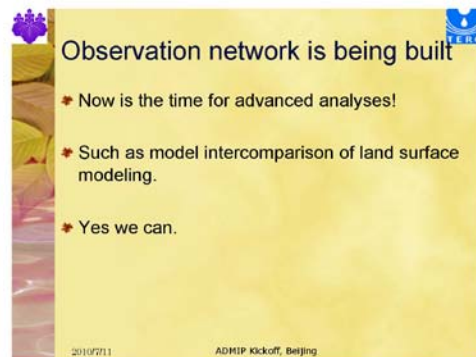


In the last decade, observation networks have grown largely ...

Fluxnet observation sites, Mar. 2010

GEWEX-CEOP phase 2 reference sites (2007-)

2010/7/11



Observation network is being built

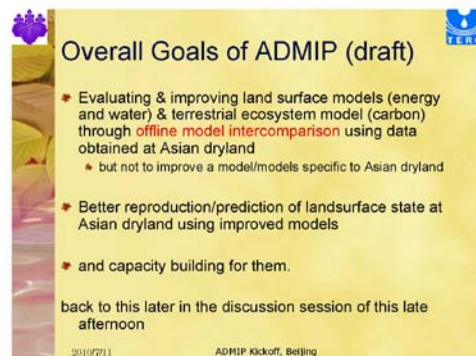
- Now is the time for advanced analyses!
- Such as model intercomparison of land surface modeling.
- Yes we can.

2010/7/11 ADMIP Kickoff, Beijing



ABOUT ADMIP

2010/7/11 ADMIP Kickoff, Beijing



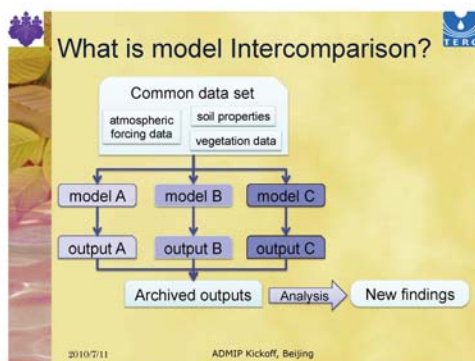
Overall Goals of ADMIP (draft)

- Evaluating & improving land surface models (energy and water) & terrestrial ecosystem model (carbon) through **offline model intercomparison** using data obtained at Asian dryland
 - but not to improve a model/models specific to Asian dryland
- Better reproduction/prediction of landsurface state at Asian dryland using improved models
- and capacity building for them.

back to this later in the discussion session of this late afternoon

2010/7/11 ADMIP Kickoff, Beijing





ADMIP – current framework

- Research framework
 - under MAHASRI, MAIRS-dryland & CEOP-dryland
- Supporting funds (currently)
 - APN (Asia Pacific Network for Global Change Research)
 - ESSP-MAIRS
 - MEXT-JSPS, Japan

2010/7/11 ADMIP Kickoff, Beijing

ADMIP – possible long term plans

- phase 1 (being started)
 - PILPS2-type intercomparison
 - off line & observation-based drivers and inputs
 - on a point/points basis
 - 2-3 years
- phase 2 (in the future)
 - GSWP-type ensemble & regional analysis
 - synthesized drivers and inputs
 - regional analysis
 - 2-3 years

2010/7/11 ADMIP Kickoff, Beijing

SCOPE OF THE MEETING

2010/7/11 ADMIP Kickoff, Beijing

This meeting is

- to identify
 - project objectives to achieve
 - tasks to be done to achieve the project objectives
 - data (observation sites) to be used in the intercomparison
 - the time line and mile stones to be followed
 - data policies and format to be followed.
- and hopefully
 - draft of project protocols will be presented at the

2010/7/11 ADMIP Kickoff, Beijing

Agenda – day 1

11 July, 2010 (Sunday)

DAY 1: morning

9:00- 9:15 Welcome address by M. Matsuo, vice chair of MAIRS SSC

9:15- 9:30 Sun. Asanuma: Scope of the meeting

9:30- 10:00 Evans OJIMA: Regional land-atmosphere considerations of acid and semi-acid land systems of Monsoon Asia

10:00- 10:45 Evans EVANS: Land surface model evaluation & suggestions for ADMIP
Lessons learned from past model intercomparisons (Chair: D. Ojima)

10:45- 11:15 Williams FARTON: Results from the FLUP and other model comparison efforts

11:15- 11:45 Zengliang YANG: Ensemble-based Methods for Intercomparing Land Surface Models

11:45- 12:15 Nakaike ETOH: Intercomparison of terrestrial ecosystem models for clarifying uncertainties in carbon cycle

12:15- 12:30 YANG Run:

12:30- 12:45 Matsuyuki KUNDO: Lessons learned from Japan-MIP

DAY 1: afternoon

Current Issues in Asian Dryland and its Modeling Perspective (Chair: J. Evans)

14:45- 15:00 Maekawa KOBAL

15:00- 15:15 Purnanjan GOMBOLOTTSEV

Participating models & Past Intercomparison 1 (Chair: J. Evans)

15:20- 15:40 Guo-Yue SHU Zengliang YANG: The Noah Land Surface Model with Multi-Physics Options

15:40- 15:55 Kanaka YOSHIZU

15:55- 16:10 Sun WEN

16:10- 17:30 Discussion 1: Key processes and questions to be answered (Chair: J. Asanuma)

2010/7/11 ADMIP Kickoff, Beijing

Agenda – day2

12 July, 2010 (Monday)

DAY 2: morning

Participating models & Past Intercomparison 2 (Chair: Zengliang YANG)

9:00- 9:20 Kanun MAIRE CHU

9:20- 9:40 Xia ZHANG

9:40- 10:10 Kamei TSUCHIDA: SEB DGVN (Spatially Explicit Individual Based Dynamic Global Vegetation Model) A model description

10:00- 10:20 Hiroyuki SATO

10:20- 10:45 Hu YANG: Dynamic Land Ecosystem Model (DLEM) – Data model intercomparison at LBA and NAOP Sites

Introduction to target site candidates (Chair: X. Maekawa)

11:00- 11:15 Ruiqiang ZHANG: Tongren Station

11:15- 11:45 Jon ASANUMA: Thar Desert Japan/Tsao Station

DAY 2: afternoon

13:00- 13:15 Discussion 2: Target sites & data, and project protocol (Chair: J. Asanuma)

13:15- 13:30 Target sites, 23 future data and data stations, 21 data policies & formats, 42 data needs, etc.

2010/7/11 ADMIP Kickoff, Beijing

Structure of this meeting

- 1st day
 - 2 key note presentations by Profs Ojima and Evans
 - 5 presentations on past model intercomparisons
 - 2 presentations from dryland countries
 - 3 presentations on introduction to model or/and model intercomparison
 - Discussion on the core of ADMIP
- 2nd day
 - 5 presentation continues to the previous day.
 - Introduction to the 2 observation sites which could provide data for this project
 - Discussion

2010/7/11 ADMIP Kickoff, Beijing

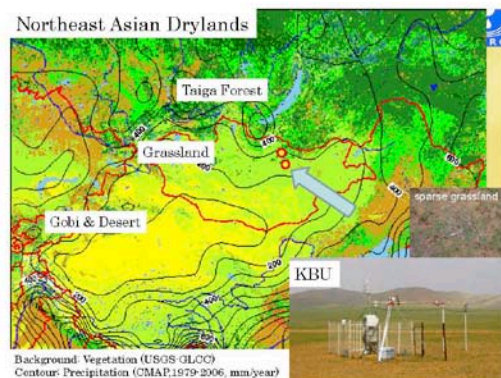




Introduction to Kherlenbayan Ulaan (KBU) flux site

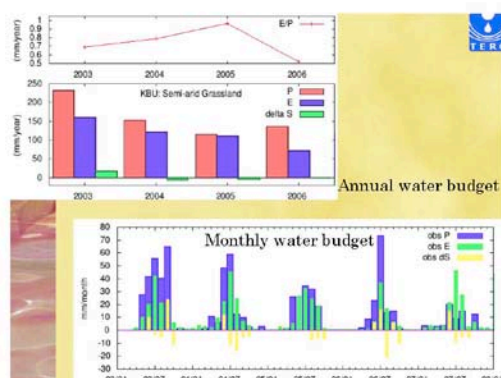
J. Asanuma
Terrestrial Environment Research Center
U. Tsukuba

G. Davaa
Institute of Meteorology and Hydrology
NAMHEM, Mongolia



Climate, land cover, vegetation

- Topography:
 - plane & flat, 1235m asl
- Climate: semi-arid
 - temp: 1.2 degC, precip. : 200 mm/year
- Landuse /cover
 - uniform & sparse grassland (grazed overly)
- Vegetation
 - dominant species: *Stipa krylovii*, *Carex sp.*, *caragana stenophylla*, *Cleistogenes squarrosa*, *Artemisia frigida*
 - LAI < 1, canopy height=0.1-0.2m
- Soil: chestnut soil, seasonally frozen



Observation

- started on Mar., 2003 (soil obs. from fall, 2002)
- registered as
 - CEOP 2 reference site
 - Asiaflux-registered site
- next to operational met. station (data available)
- initially designed to evaluate grazing effect on surface heat, water and carbon exchange
 - fenced and unfenced
 - later, fence was broken

Kherlenbayan Ulaan(KBU): Instruments

Observations

- Eddy Covariance (3.5m above ground)
- Kaijo SAT-540
- Licor LI7500
- 4 radiation components
- Kipp&Zonen CNR1
- radiational surf. temp.
- temp. & humidity
- Vaisala HMP45D
- precipitation (tipping bucket)
- Soil
 - temp. : Climated C-PT
 - moisture : Campbell CS616
 - heat : REBS PHF-1.1

sparse grassland 10 AMV Data WS

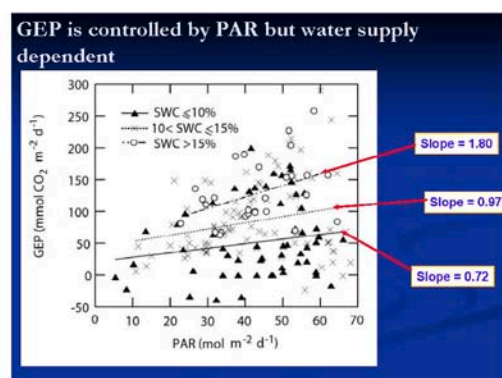
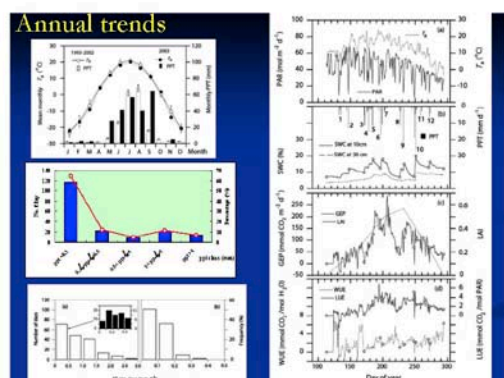
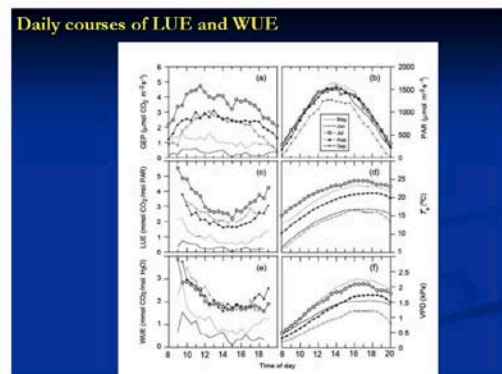
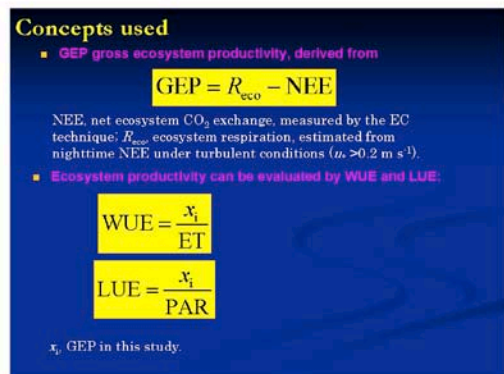
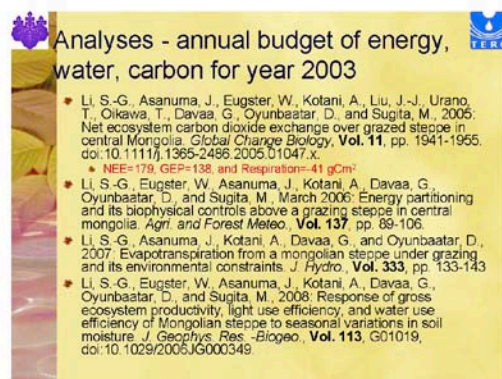
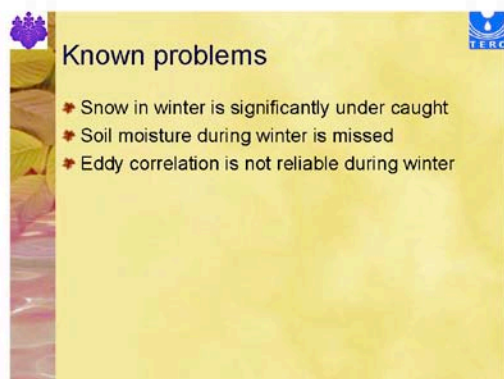
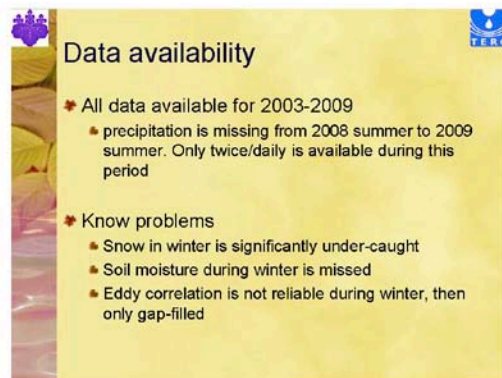
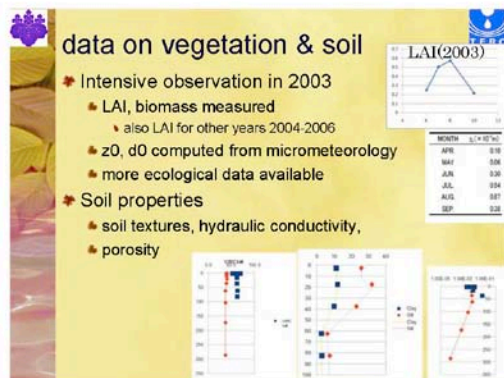
Routinely observed variables

- radiation
 - 4 components of radiation, surface radiative temp.
- general meteorology
 - wind speed, wind direction, temperature, relative humidity
- eddy covariance flux
 - sensible & latent heat flux, NEE
- soil
 - soil temperature
 - soil moisture
 - soil heat flux

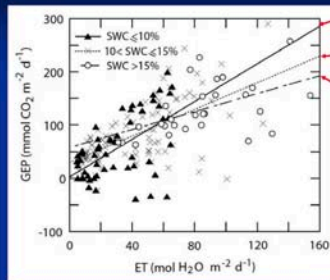
Post-observation data processing

- Data intervals
 - eddy fluxes averaged over 30min.
 - other variables are also at 30 min interval
 - daily, monthly and annual average computed
- Quality control
 - mainly for eddy covariance measurements
 - follows Asiaflux protocol: variance-based, empirical
- Gap filling
 - Using Reichstein's statistical model.

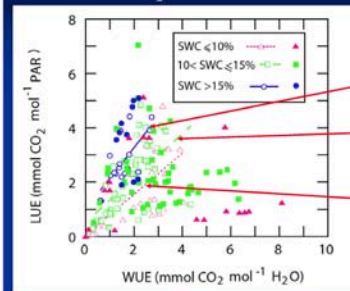




GEP and ET are closely related depending on water availability



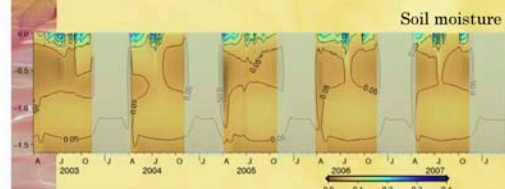
LUE and WUE are coupled but maximal LUE and WUE are decoupled



Some analyses

- Sato, T., M. Tsujimura, T. Yamanaka, H. Iwasaki, A. Sugimoto, M. Sugita, F. Kimura, G. Davaa, and D. Oyumbaatar (2007). Water sources in semiarid northeast Asia as revealed by field observations and isotope transport model. *J. Geophys. Res.*, 112, D17112, doi:10.1029/2006JD008321.
 - Partitioning of ET into E & T using δ¹⁸O for 4 times in 2003: T/ET=35-59% on average
- Asanuma, J., K. Takata, H. Fukui, 2010: Key Controlling Factors of Surface Water Budget at the Semi-Arid Grassland in Northeastern Eurasia, *J. Hydro. Eng. Japan*, vol 54
 - 4-year annual budget and CLM sensitivity analyses

some figures



Discussion core of protocol (draft)

J. Asanuma

This morning's comment

- "Modellers need to be involved in setting up model comparison" by Prof. Parton
- Let's discuss!
- in the following, some idea come from discussion with Japanese researchers, today's talks, and protocols of the previous intercomparison.

Goals of ADMIP (draft)

- Evaluating & improving land surface models (LSM, energy and water) & terrestrial ecosystem model (TEM, carbon) through **offline model intercomparison** using data obtained at Asian dryland
 - but not to improve a model/models specific to Asian dryland
- Better reproduction/prediction of landsurface state at Asian dryland using improved models
- and capacity building for them.

2010/7/11

ADMIP Kickoff, Beijing

Target phenomena & variables to be looked at

- Energy, water and carbon exchange
 - Water-limited process
 - Vegetation (growth) controlling process
 - Daily time scale, 3hourly, monthly ...
- Surface(soil) moisture & temperature
- Radiation components



Science questions to be answered

- ✦ What is the ability of models to reproduce energy, water, and carbon exchange in dry climate? (reproducibility)
- ✦ Do current ecosystem models, developed mainly for forest, reproduce energy, water and carbon exchange? (from Dr. Itoh)
- ✦ Does the current complication of models effectively simulate land surface processes? (complexity issues, also related to benchmarking)
- ✦ Does interexchange of information between a LSM and a TEM improve the model results? (offline coupling)
- ✦ Does the multi-model ensemble of LSMs outperform single model output? (ensemble)
- ✦ > proposal basis

Experimental setup

- ✦ Stage 0: basic analyses
 - of previous model intercomp.
 - Wha
- ✦ Stage 1: simple & straightforward comparison
 - comparison within LSMs and TEMs (reproducibility issue)
 - comparison between ET from LSMs, TEMs and empirical eqns. (complexity issue, benchmarking?)
- ✦ Stage 2: offline, one-way coupling experiment
 - rerun TEMs with soil moisture from LSMs
 - rerun LSMs with vegetation parameters from TEMs (lot of LSMs has done already)
 - compare with Stage1 results.
 - Sensitivity analysis

- ✦ look at GSWP1 or 2(?) at dryland
 - ET, and its relation with SM
 - => simultaneously with intercomparison

- ✦ Parameters (default, or tuned)
 - leave parameter selection to modelers, except unambiguously determined from obs. (Evans)
 - Need good data for parameters (K. Yang)
 - Compare ensemble with frequently used parameters and with optimized parameters (Yang)
 - need detailed document for parameter selection (Parton)
 - sensitivity analysis?


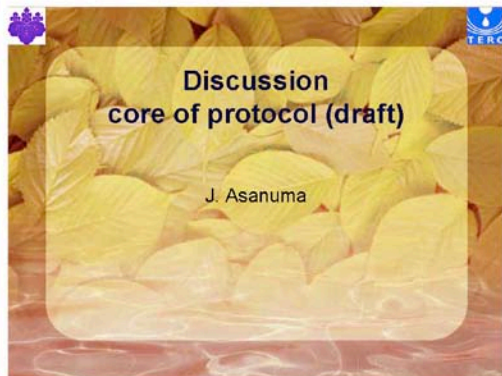

Discussion tomorrow (agenda change)

- ✦ More individual presentations, 5 presentations
- ✦ Introduction to candidates of the target sites
 - Tongyu (inner Mongolia) by Zhang.
 - KBU (Mongolia) by Asanuma
 - ?? (Shinxian) by Yang Kun
- ✦ Discussion
 - selection of obs. sites; 2(1 main, 1 backup) out of 3 sites

Experimental setup – 1 data preparation


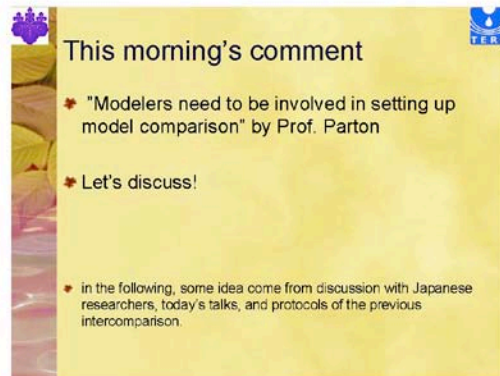

- ✦ Preparing driving data
 - atmospheric forcing
 - vegetation, soil properties




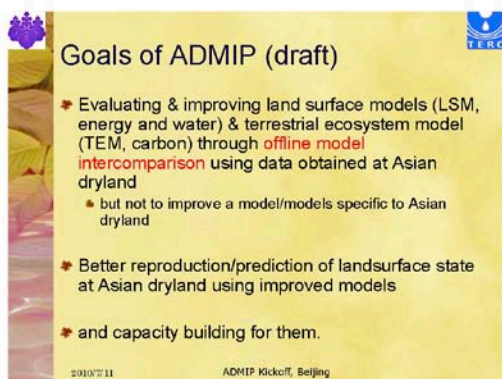

Discussion core of protocol (draft)

J. Asanuma

This morning's comment


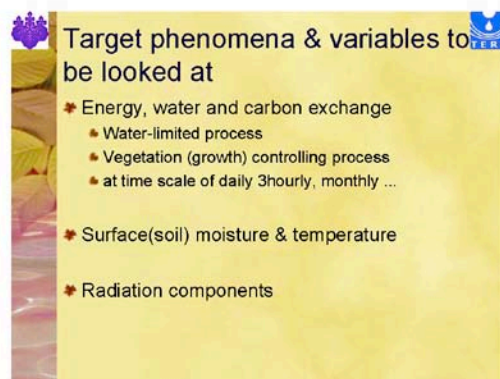

- * "Modelers need to be involved in setting up model comparison" by Prof. Parton
- * Let's discuss!
- * in the following, some idea come from discussion with Japanese researchers, today's talks, and protocols of the previous intercomparison

Goals of ADMIP (draft)


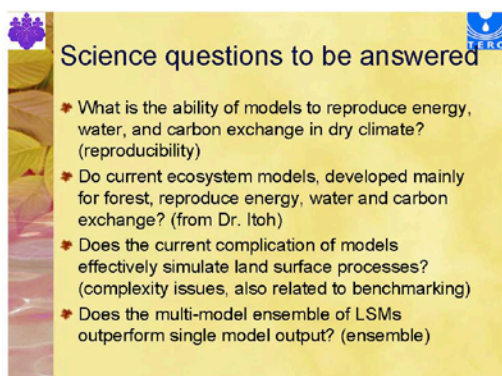

- * Evaluating & improving land surface models (LSM, energy and water) & terrestrial ecosystem model (TEM, carbon) through **offline model intercomparison** using data obtained at Asian dryland
 - * but not to improve a model/models specific to Asian dryland
- * Better reproduction/prediction of landsurface state at Asian dryland using improved models
- * and capacity building for them.

2010/7/11 ADMIP Kickoff, Beijing

Target phenomena & variables to be looked at


- * Energy, water and carbon exchange
 - * Water-limited process
 - * Vegetation (growth) controlling process
 - * at time scale of daily 3hourly, monthly ...
- * Surface(soil) moisture & temperature
- * Radiation components

Science questions to be answered


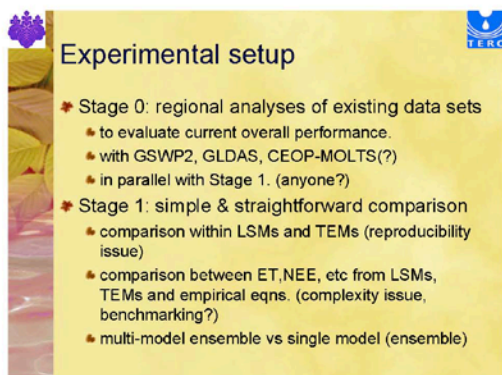

- * What is the ability of models to reproduce energy, water, and carbon exchange in dry climate? (reproducibility)
- * Do current ecosystem models, developed mainly for forest, reproduce energy, water and carbon exchange? (from Dr. Itoh)
- * Does the current complication of models effectively simulate land surface processes? (complexity issues, also related to benchmarking)
- * Does the multi-model ensemble of LSMs outperform single model output? (ensemble)






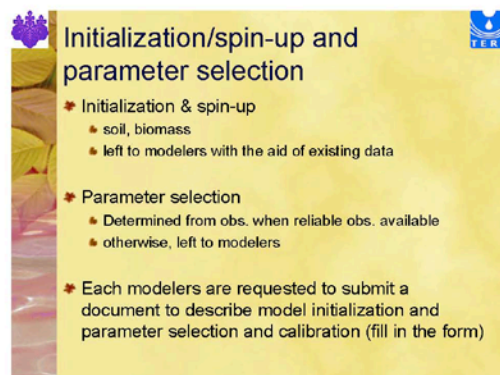

Target sites, periods

- * Tongyu (contact: Ailikun)
 - * period: 2003-2009
- * KBU (contact: J. Asanuma)
 - * period: 2003-2007
- * Pingliang(contact: Jun WEN)
 - * period: 2004-2008

Experimental setup

- * Stage 0: regional analyses of existing data sets
 - * to evaluate current overall performance.
 - * with GSWP2, GLDAS, CEOP-MOLTS(?)
 - * in parallel with Stage 1. (anyone?)
- * Stage 1: simple & straightforward comparison
 - * comparison within LSMs and TEMs (reproducibility issue)
 - * comparison between ET, NEE, etc from LSMs, TEMs and empirical eqns. (complexity issue, benchmarking?)
 - * multi-model ensemble vs single model (ensemble)

Initialization/spin-up and parameter selection

- * Initialization & spin-up
 - * soil, biomass
 - * left to modelers with the aid of existing data
- * Parameter selection
 - * Determined from obs. when reliable obs. available
 - * otherwise, left to modelers
- * Each modelers are requested to submit a document to describe model initialization and parameter selection and calibration (fill in the form)



Forcing and ancillary data

- ✦ forcing
 - ✦ for LSMs
 - ✦ Wind, Tair, Qair, CO2air, Rain, Pout, SWdown, LWdown, LAI
 - ✦ time interval: 20 to 60 min
 - ✦ for TEMs
 - ✦ Tair, Rain, Qair, SWdown
 - ✦ time interval: ?
 - ✦ need documentation for post data processing
- ✦ ancillary data
 - ✦ soil properties
 - ✦ report from observers with all available information needed
 - ✦ e.g., texture, conductivity, S-parameter of Homburger, etc
 - ✦ soil respiration, biomass
 - ✦ vegetation
 - ✦ LAI: estimated from multiple sources of satellite LAI with validation with ground-observed LAI (Panton & Yorozu)
 - ✦ biomass: estimated from LAI with reference to obs
 - ✦ fraction of C3/C4
- ✦ for validation
 - ✦ flux
 - ✦ soil: soil respiration, biomass

Model output data

- ✦ Output variables from LSMs
 - ✦ Qh, Qle, and more
 - ✦ interval: 20min-1hours, daily, monthly
 - ✦ albedo, conductance
- ✦ Output variables from TEMs (multiple years)
 - ✦ GPP, NEE, ET: daily
 - ✦ live plant biomass (above ground, convertible with LAI)
 - ✦ soil water & temp. () =>

Data Issues

- ✦ Data format & Archiving
 - ✦ File format: ALMA-compliant netCDF
 - ✦ archived at PALS at U. NSW.
- ✦ Data policy
 - ✦ All forcing data & model outputs will be shared by the participants of ADMIP
 - ✦ Key paper will be published with all participants as authors, as well the data provider.
 - ✦ At 1 year after the submission (or publication), **forcing and input data (if possible, observed data)** & model output will be available to public

Time line

- ✦ atmospheric forcing data & other obs data to PALS: mid september
- ✦ model output submission: early march
- ✦ ws in mid apr

Purpose of next workshop

- ✦ to review of the results
- ✦ to decide core results
- ✦ to outline publication

Time line & mile stones

	2010												2011											
	01	02	03	04	05	06	07	08	09	10	11	12	01	02	03	04	05	06	07	08	09	10	11	12
kickoff meeting																								
agenda																								
writing protocol																								
distribution for review																								
reviewing protocol																								
model rep. session																								
deadline for submission																								
forcing data & ancillary data																								
preparation																								
deadline for submission																								
available to the community																								
stage 1: basic model runs																								
model runs																								
deadline for model output																								
analysis of stage 1 model results																								
discussion																								
2010 meeting at Tsukuba																								
vis. at Tsukuba, Japan																								
workshop at U. Kyoto																								

other logistics

- ✦ Project home page will be at U. Kyoto managed by Dr. Yorozu
 - ✦ <http://kugi.kyoto-u.ac.jp/~admip/>
 - ✦ project documents, meeting schedule, link to PALS and else.
- ✦ Mailing list for project participants [admip]
 - ✦ Anyone can join from
 - ✦ <http://lists.suiri.tsukuba.ac.jp/listinfo/admip>
 - ✦ or send a mail to asanuma@suiri.tsukuba.ac.jp

Those who are supported by U. Tsukuba fund

- ✦ Please leave boarding pass stub, otherwise ...





* Blind policy

- * w/ default parameters
 - "default"
 - what is default parameters?
 - depends on degree of empiricity
- * w/ calibrated parameters (what is obj-func?)
 - best run! what is best?
 - should be documented
 - "most frequent" results, ensemble analyses (volunteers needed)



* look at GSWP1or2(?) at dryland

- ET, and its relation with SM
- => simultaneously with intercomparison

* Parameters (default, or tuned)

- leave parameter selection to modelers, except unambiguously determined from obs. (Evans)
- Need good data for parameters (K. Yang)
- Compare ensemble with frequently used parameters and with optimized parameters (Yang)
- need detailed document for parameter selection (Parton)
- sensitivity analysis?



Discussion tomorrow (agenda change)

- * More individual presentations, 5 presentations
- * Introduction to candidates of the target sites
 - Tongyu (inner Mongolia) by Zhang.
 - KBU (Mongolia) by Asanuma
 - ?? (Shinxian) by Yang Kun
- * Discussion
 - selection of obs. sites; 2(1 main, 1 backup) out of 3 sites



1.4 Meeting Summaries

Asian Dryland Landsurface Modeling Intercomparison

Project (ADMIP)

Kickoff Meeting

July 11-12, 2010,

Beijing Expert Hotel, Beijing, China

Meeting Summary (final)

The following is the summary of agreements made during the kickoff meeting of ADMIP held at Beijing, China on Jul. 11th and 12th, 2010.

1. Goals of the Project

The goal of ADMIP is evaluating and improving land surface models (LSMs, energy and water) & terrestrial ecosystem model (TEMs, carbon) through offline model intercomparison using data obtained at Asian dryland, towards better reproduction and prediction of landsurface state using improved models, and necessary capacity building.

2. Target Phenomena and Scientific Questions

1) Target Phenomena

Energy, water and carbon exchange at the land surface under water-limited environment, and/or under the control of vegetation growth, at the temporal scales of 3-hourly, daily, and monthly.

2) Scientific Questions

- What is the ability of models to reproduce energy, water, and carbon exchange at dryland surface (reproducibility) ?
- Do current ecosystem models, developed mainly for forest, reproduce energy, water and carbon exchange?
- Does the current complexity of models effectively simulate land surface processes? (complexity issues, also related to benchmarking)
- Does the multi-model ensemble of LSMs outperform single model output?



(ensemble)

3. Target Sites

Three target sites were chosen, data of which will be archived and used for the intercomparison.

1) Tongyu (China)

contact person: Dr. Ailikun (MAIRS-IPO, ailikun@gmail.com)

target period: 2003-2009

note: GEWEX-CEOP registered

http://www.eol.ucar.edu/projects/ceop/dm/insitu/sites/ceop_ap/Tongyu/Cropland

2) Kherlen Bayan Ulaan (KBU, Mongolia)

contact person: Dr. Jun Asanuma (U. Tsukuba, asanuma@sui.ri.tsukuba.ac.jp)

target period: 2003-2007

note: GEWEX-CEOP and IGBP-AsiaFlux registered

http://www.eol.ucar.edu/projects/ceop/dm/insitu/sites/ceop_ap/Northern_Mongolia/

[Kherlen](#)

http://asiaflux.yonsei.ac.kr/network/021KBU_1.html

3) Pingliang (China)

contact person: Dr. Jun Wen (CAREERI, jwen@lzb.ac.cn)

target period: 2004-2008

note: site info will be given at ADMIP homepage.

4. Model Runs

1) Stage 0: analyses of existing data set

Basic investigation of the results from previous intercomparison or analyses. e.g., GSWP 1 and 2, GLDAS, and CEOP-MOLTS, will be made to evaluate current overall performance.

2) Initialization (common to stage 1 and 2)

Initialization, or spin-up, of soil water, soil heat storage and biomass in the model will be left to the model operators. The model operators will be requested to report their initialization.



3) Stage 1: w/"default" parameter set

Parameters in the models will be determined from observations if reliable information is available. Otherwise, the decision is left to the model operators. No calibration will be done in this stage. All of the model operators will be requested to report their parameter selection.

4) Stage 2: "calibrated" parameter set

Calibration will be performed by tuning the model results with the observations. All of the model operators will be requested to report their parameter selection.

5. Model input/output variables and input parameters

(note: ALMA variable names in **Bold**)

1) Input variables (as time series)

The variables listed below will be prepared as time series from observed data set.

- for LSMs (at every 30min or so)

- **SWdown** and **LWdown** (incoming shortwave & longwave radiation, maybe diffuse/direct)
- **Tair** (air temperature), **Qair** (specific humidity), **CO2air** (CO2 concentration), and **Wind** (wind speed) at specified height
- **Snowf** and **Rainf** (snowfall and rainfall, or their sum)
- **Psurf** (Surface pressure)
- LAI (as time series if available)

- for TEMs (daily, or with longer time interval)

- max. and min. temperature
- precipitation, VPD, and **SWdown**
- current and historical land use (grazing history)

2) Input parameters (as fixed value in time)

The variables listed will be either prepared on the basis of the ground observation or derived from satellite measurements. Any post-observation processing of observed records should be documented.

- for LSMs

- Vegetation type/properties (that can be measured with confidence)
- Soil type/properties (that can be measured with confidence)
- Elevation, slope, aspect



- for TEMs (in addition to the above)

- Soil texture
- Additionally, soil C and N (for validation?)
- Peak live aboveground biomass, root biomass (for validation?) and snow cover
- Soil respiration (for validation?)

3) Output variables

The variables list below will be computed with LSMs and TEMs and will be submitted to the project for further analyses.

- from LSMs

➤ **Energy**

SWup and **LWup** (outgoing shortwave and longwave radiation)

Qh(sensible heat), **Qle**(latent heat), **Qg**(ground heat)

Soil temperature

➤ **Water**

Evaporation (transpiration, canopy interception loss, ground))

Qs, **Qsb** (surface and subsurface runoff), Snow (SWE and/or depth)

Soil moisture, **WaterTableD**(water table depth)

➤ **Carbon**

GPP, **NEE**, **NPP**, respiration

- from TEMS (Daily)

Live plant biomass,

Soil water and temperature,

Qle(latent heat flux), **NEE**, and **GPP**

6. Data issues

1) Data format

ALMA-compliant netCDF format will be used for submitting, archiving and exchanging the forcing and model output data.

2) Data working group

The data working group will be organized to prepare necessary data. All of submission of the data will be directed to the data working group at its data focal point to check its format and quality before being archived at PALS (see below)



Jun Asanuma(convener & KBU site), Kazu Yorozu (data focal point, analyses),
Ailikun (Tongyu site), Jun Wen (Pingliang site), Jason Evans (liaison to PALS),
William Parton (adviser)

3) Data archiving at PALS

All the model input data, forcing data and output data will be archived finally at PALS (Protocol for the Analysis of Land Surface models) operated at University of New South Wales (Contact: J. Evans). PALS will be also used to analyze the results.

4) Data sharing policy

First, all forcing data and model outputs will be shared by the participants of ADMIP. A key paper will be published with all participants as well as the data provider as the authors. At 1 year after the publication of the key paper, forcing and input data that were used to run the models and the model outputs will be available to the public.

7. Time line

Year 2010

Aug.	drafting the intercomparison protocol and its distributing for review
31st, Aug.	finalizing the intercomparison protocol
15th, Sep.	deadline for forcing data and observation data submission
30th, Sep.	deadline for model registration
	forcing data and obs. data available to the community

Year 2011

10th, Mar.	deadline for model output submission.
mid Apr.	Workshop in Japan



	2010						2011											
	07	08	09	10	11	12	01	02	03	04	05	06	07	08	09	10	11	12
Kickoff meeting	▲																	
Agenda																		
drafting protocol	■	■																
distribution for review			■	■														
finalizing protocol			▲															
Model registration	■	■	■	■														
deadline for submission				▲														
Forcing data & Ancillary data																		
preparation	■	■	■															
deadline for submission				▲														
available to the community				▲														
Stage 1 basic model runs																		
model runs	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
deadline for model output									▲									
Analysis of stage 1 model outputs									■	■	■	■	■	■	■	■	■	■
Meetings																		
AMS meeting at Texas								▲										
WS at Tsukuba, Japan										▲								
WPGM or AOGS (?)														▲				

8. Miscellaneous

1) Project homepage

Project home page at <http://hywr.kuciv.kyoto-u.ac.jp/admip> maintained by Dr. Yorozu (U. Kyoto). It will archive project documents, will keep updated information of the project, such as meeting schedules and important links.

2) Project mailing list

Project mailing list, [ADMIP], can be joined at <http://lists.suiri.tsukuba.ac.jp/listinfo/admip>

3) Intercomparison Protocol

Intercomparison will be written on the basis of this meeting summary.



1.5 Pictures



2.0 The 2nd Summer School on Land Surface Observing, Modeling and Data Assimilation 13-16 July 2010, Beijing Normal University, Beijing, China

2.1 Program Schedule

The 2nd Summer School on Land Surface Observing, Modeling and Data Assimilation

13-16 July 2010, Beijing Normal University, Beijing, China

Program Schedule

Registration	
8:30 - 20:00, July 12 (Monday), the main hall of the Science and Technology Building	
7:00 - 8:00, July 13, RM502, Zeng Xianzi Building (Teaching Building 9)	Earphone Distribution
7:00-12:00, July 13, RM502, Zeng Xianzi Building (Teaching Building 9), Registration	
Opening Ceremony, Chaired by Prof. S. Liang	
8:30 – 9:30am, July 13 (Tuesday), RM502, Teaching Building 9	
8:30-8:35	Opening remarks, Prof. S. Liang
8:35-8:45	Speech by Prof. Guanhua Xu, Former Minister of Chinese Ministry of Science and Technology
8:45-8:55	Speech by Dr. Ramasamy Jayakumar, UNESCO Office in Beijing
8:55-9:05	Speech by Dr. Ailikun, Monsoon Asia Integrated Regional Study (MAIRS)
9:05-9:15	Welcome remarks by Vice President of Beijing Normal University
9:15-9:30	Logistics of the Summer School, Ms. Bo Jiang, Beijing Normal University
9:30 – 10:00 Group picture taking	
PART I: Data Sources, Chaired by Dr. J. Shi, July 13, 2010 (Tuesday):	
10:00 - 11:00	Overview of Satellite Remote Sensing Products for Land Data Assimilation, Prof. S. Liang
11:00 - 11:10	Break
11:10 - 12:10	Atmosphere Forcing Data in China Region Based on FY2 Satellite, Dr. Chunxiang Shi
12:10 - 14:00	Lunch

14:00 - 15:20	Second Generation Polar-Orbiting Meteorological Satellites of China: Fengyun 3 and its application on global monitoring, Prof. Peng Zhang
15:20 - 15:30	Break
15:30 - 16:50	NASA Satellite Remote Sensing and Assimilation Data and Services: A Tutorial of Data Accessing, Dr. Suhung Shen
16:50 - 17:00	Break
17:00-18:20	Microwave Modeling and Parameterization for Data Assimilation, Dr. Jiancheng Shi
PART II: Land Surface Modeling, Chaired by Dr. Ailikun, July 14, 2010 (Wednesday):	
8:30 - 9:50	Modeling Land and Atmosphere Interactions, Prof. Yongkang Xue
9:50 - 10:00	Break
10:00 - 11:20	Land Surface Modeling in Arid and Semi-arid Regions, Dr. Kun Yang
11:20 - 14:00	Lunch
14:00 - 15:20	Integration of MODIS Products and a Crop Simulation Model for Crop Yield Estimation, Prof. Hongliang Fang
15:20 - 15:30	Break
15:30 - 16:50	Uncertainty Quantification of Large Complex Dynamical Systems, Prof. Qingyun Duan
16:50 - 17:00	Break
17:00-18:20	Parameter Estimation and Error Analysis of Ecosystem Models: Component Analysis of Ecosystem Models, Dr. Dennis Ojima
PART III: Data Assimilation Methods, Chaired by Dr. Xin Li, July 15, 2010 (Thursday):	
8:30 - 9:50	Assimilating Remote Sensing Data into Land Surface Models: Theory, Method and a Software Tool I, Dr. Xin Li
9:50 - 10:00	Break
10:00 - 11:20	Cost-function Based Methods and Particle Filter Method: Theory and Application in Soil Moisture Estimation, Dr. Jun Qin
11:20 - 14:00	Lunch
14:00 - 15:20	Review of Parameterization and Parameter Estimation for Hydrologic models, Prof. Soroosh Sorooshian
15:20-15:30	Break
15:30 - 16:50	Forecast Error Covariance Inflation For Data Assimilation, Prof. Xiaogu Zheng



16:50 - 17:00	Break
17:00 - 18:20	Assimilating Remote Sensing Data into Land Surface Models: Theory, Method and a Software Tool II , Dr. Xin Li and Dr. Xujun Han
PART IV: Data assimilation applications, Chaired by Dr. Xin Li or Dr. Xianhong Xie, July 16, 2010 (Friday):	
8:30 - 9:40	Towards Operational Soil Moisture Monitoring, Prof. Jeffrey Walker
9:40 - 9:50	Break
9:50 - 11:00	The Application of Remotely Sensed Land Surface Properties in Regional Climate Model, Dr. Lixin Lu
11:00 - 11:10	Break
11:10-12:20	Land Surface Modeling and Its Applications, Prof. Xu Liang
12:20 - 14:00	Lunch
14:00 - 15:10	North American Land Data Assimilation System, Dr. Youlong Xia
15:20 - 15:30	Break
15:30 - 16:30	Distributed Hydrological Modeling and Data Assimilation, Dr. Xianhong Xie
16:30-16:40	Break
Special session on UNESCO-IHP, Chaired by Dr. Xin Li	
16:40 - 17:40	Role of UNESCO-IHP (International Hydrological Programme) in Global Hydrological System and Modelling and IHP Programme in East Asia, Dr. Ramasamy Jayakumar and Dr. Anil Mishra
17:40 - 18:20	Climate Change Adaptation: Two Case Studies of IHP from Indian Sub Continent, Dr. Bhanu Neupane
18:20 - 19:00	Returning earphones and University Temporary card Distribution of CD containing all lecture notes, slides, and pictures



2.2 List of Participants

List of participants of Land Surface Training course supported by APN and
MAIRS
13-16 July 2010, Beijing Normal University

Ms. Jigjsuren ODGARAV,
Hydrologist, Hydrology section,
Institute of Meteorology and Hydrology, Ulaanbaatar-46, Mongolia
E-mail is Oogii_0822@yahoo.com
Ph: 976-96600667, 976-11-312765
Fax: 976-11-326611

Dr Tserenchunt BATTUMUR
Research scientist at the Dryland Sustainability Institute
Research interest: Climate change, Use of Remote Sensing and GIS for Land Use
and Land Cover change studies.
Address: P.O. Box: 44/243
Partizan street 31, Sukhbaatar district 5
Ulaanbaatar 14250, Mongolia
Tell: (976) 99063747,
Fax: (976) 11-266173
E-mail: tserenchunt.b@mail.mn, bchunt_11@yahoo.com

Ms. Banzragch NANDINTSETSEG
Arid Land Research Center, Tottori University
〒680-0001, Tottori Prefecture, Tottori City, Hamasaka 1390, Tottori University,
Arid Land Research Center
Tel: (0857) 23 – 3411, 090-9464-5515
Fax: (0857) 29 – 6199
e-mail: nandia@alrc.tottori-u.ac.jp

Mr. Gankhuyag BATJARGAL
Forecast Research Laboratory, Institute of Meteorology and Hydrology
Juulchiny gudamj 5, Ulaanbaatar 46, Mongolia
E-mail: batjargal2006@yahoo.com

Mr. MUHAMMAD IJAZ
Present Position: Scientific Officer
Address: Global Change Impact Studies Centre (GCISC),
National Centre for Physics (NCP) Complex,
Quaid-i-Azam University Campus,
Shahdara Road, Islamabad - 44000, Pakistan
Tel: +92-51-2077300/Ext 457, 455
Mobile: +92-300-5315021
E-Mail: m.ijaz@gcisc.org.pk



muhammad.ijazmalik@gmail.com

Miss Nuzba Shaheen

Present Position: Scientific Officer

Address: Global Change Impact Studies Centre (GCISC),

National Centre for Physics (NCP) Complex,

Quaid-i-Azam University Campus,

Shahdra Valley Road, Islamabad, Pakistan.

Tel: +92-51-2077300/Ext 457, 455

Mobile No: 0332-5424554

Phone (office): +92-51-9230226

E-Mail: nuzba.shaheen@gcisc.org.pk

nuzba_gcisc@yahoo.com

Mr. Winai Chaowiwat

Room 203, 2nd Fl, Bldg. 2,

Water Resources Engineering,

Faculty of Engineering,

Chulalongkorn University,

Bangkok 10330

Thailand

E-mail: winaich@yahoo.com

Mr. Chalermrat Sangmanee

Southeast Asia START Regional Centre (SEA START RC),

Chulalongkorn University, Chulawich 1 Building, 5th Floor

Henri Dunant Road, Bangkok 10330, Thailand

Tel: (66 2) 218 9464 to 7

Mobile: (66 8) 6550 0024

Fax: (66 2) 251 9416

e-mail: chalermrat@start.or.th

Dr. Hisashi SATO

Graduate School of Environmental Studies, Nagoya University

D2-1(510) Furo-cho, Chikusa-ku, Nagoya, 464-8601, JAPAN

e-Mail: sato.hisashi@nagoya-u.jp

URL : http://seib-dgvm.com/hsato/English/index_E.html

TEL : +81-52-747-6546(Direct)

FAX : +81-52-747-6520

13-16 July Training course participants supported by APN and MAIRS (in China)

Dr. Zhihui ZHANG

Cold and Arid Regions Environmental and Engineering Research Institute
(CAREERI)



Lanzhou, China

Dr. WANG Xufeng
Cold and Arid Regions Environmental and Engineering Research Institute
(CAREERI)
Lanzhou, China

Dr. Zhang Tangtang
Cold and Arid Regions Environmental and Engineering Research Institute
(CAREERI)
Lanzhou, China

Dr. Wang Xin
Cold and Arid Regions Environmental and Engineering Research Institute
(CAREERI)
Lanzhou, China

Dr. Liu Lichao
Cold and Arid Regions Environmental and Engineering Research Institute
(CAREERI)
Lanzhou, China

Dr. Gao Yanhong
Cold and Arid Regions Environmental and Engineering Research Institute
(CAREERI)
Lanzhou, China

Dr. Su Na
Cold and Arid Regions Environmental and Engineering Research Institute
(CAREERI)
Lanzhou, China

Dr. Lian Jie
Cold and Arid Regions Environmental and Engineering Research Institute
(CAREERI)
Lanzhou, China

Dr. Ji Xibin
Cold and Arid Regions Environmental and Engineering Research Institute
(CAREERI)
Lanzhou, China

Dr. Zhihui ZHANG
Cold and Arid Regions Environmental and Engineering Research Institute



(CAREERI)
Lanzhou, China

Dr. Zhao Liwen
Cold and Arid Regions Environmental and Engineering Research Institute
(CAREERI)
Lanzhou, China

Dr. Li Suosuo
Cold and Arid Regions Environmental and Engineering Research Institute
(CAREERI)
Lanzhou, China

Dr. Guo Weidong
Cold and Arid Regions Environmental and Engineering Research Institute
(CAREERI)
Lanzhou, China

Dr. Wang Shaoying
Cold and Arid Regions Environmental and Engineering Research Institute
(CAREERI)
Lanzhou, China

Dr. Wen Xiaohang
Cold and Arid Regions Environmental and Engineering Research Institute
(CAREERI)
Lanzhou, China

Dr. Ma Di
Cold and Arid Regions Environmental and Engineering Research Institute
(CAREERI)
Lanzhou, China

Dr. Zhao Lin
Cold and Arid Regions Environmental and Engineering Research Institute
(CAREERI)
Lanzhou, China

Dr. Chen Siyu
College of Atmospheric Sciences, Nanjing University
22 Hankou Road, Nanjing 210093
P. R. China

Dr. Xie Jing



College of Atmospheric Sciences, Nanjing University
22 Hankou Road, Nanjing 210093
P. R. China

Dr. Ling Xiaolu
College of Atmospheric Sciences, Nanjing University
22 Hankou Road, Nanjing 210093
P. R. China

Dr. Zhao Jianfei
College of Atmospheric Sciences, Nanjing University
22 Hankou Road, Nanjing 210093
P. R. China

Dr. Niu Xiaorui
Institute of Atmospheric Physics
Chinese Academy of Sciences
P.O. Box 9804, Beijing 100029
P. R. China

Dr. Li Shan
Institute of Atmospheric Physics
Chinese Academy of Sciences
P.O. Box 9804, Beijing 100029
P. R. China

Dr. Lin liuying
Beijing Normal University

Dr. Xu Lu
Beijing Normal University

Prof. LIU Wenzhao
Institute of Water and Soil Conservation, CAS
26 Xi'nong Road,
Yangling Demonstration Area, Shaanxi
Province 712100, China
Tel: + 86 29 7012307
Fax: + 86 29 7012210
E-mail: wzliu@ms.iswc.ac.cn

Dr. Zhu Yuanjun
Institute of Water and Soil Conservation, CAS
26 Xi'nong Road,
Yangling Demonstration Area, Shaanxi



Province 712100, China

Dr. Wang Sheng
Gansu Bureau of CMA

Dr. Wang Heling
Gansu Bureau of CMA

Dr. Tu Gang
Lijin Bureau of CMA



2.3 Pictures

The 2nd Summer School on Land Surface Observing, Modeling and Data Assimilation July 13-16, 2010
第二届陆面观测、模拟与数据同化培训班 2010.7.13-16



3.0 The 2nd APN-MAIRS joint workshop on Intercomparison of Land surface Process Modelling at Asian Drylands, Lanzhou, China, 13-14 July 2011

3.1 Agenda

APN-MAIRS Joint Workshop on Intercomparison of Land surface Process
Modelling at Asian Drylands
July 13-14 July 2011, Lanzhou, China

13 July, 2011 (Wednesday)

DAY 1- morning

- 0:30 9:30 - 10:00 **Registration**
Opening (Chair: Ojima)
- 0:05 10:00 - 10:05 **Welcome address (Jun Asanuma)**
- 0:30 10:05 - 10:35 **Jun WEN: Introduction to Cold and Arid Regions Environment and Engineering Research Institute, Chinese Academy of Sciences**
- 0:40 10:35 - 11:15 **Shin MIYAZAKI: Interaction between climate and pasture in Mongolia**
- 0:40 11:15 - 11:55 **Hong JINKYU: Comparative analysis of two land surface models on the Tibetan Plateau**
- 1:30 11:55 - 13:25 **Lunch (group photo)**

DAY 1- afternoon

- First Results from ADMIP (Chair: Gomboluudev)**
30 minutes each talk
- 0:20 13:25 - 13:45 **Denis OJIMA: Ecosystem controls on land surface-atmosphere exchange of the Mongolian Plateau steppe**
- 0:20 13:45 - 14:05 **Yaomin SONG: Performances of NCAR_CLM4.0 and CoLM in Simulating Land-Atmosphere Interactions at Tongyu and KBU sites**
- 0:20 14:05 - 14:25 **Masayuki KONDO: Adoptability of Biome-BGC to the dryland in East Asia**
- 0:20 14:25 - 14:45 **YANG Kun: Land surface modeling for drylands of China**
- 0:20 14:45 - 15:05 **Xia Zhang: Simulation Study on Bare Soil Surface Water and Energy Balances at Tongyu Site using A Soil-Atmosphere-Transfer Model**
- 0:15 15:05 - 15:20 **Tea break**
- Discussion 1: Data Issues (Chair: Ailikun)**
10 to 15 minutes each talk
- 0:15 15:20 - 15:35 **Jun ASANUMA: Processing Data from Kherlen Bayan Ulaan Station**
- 0:15 15:35 - 15:50 **Kazuaki YOROZU: Compiling Tongyu data set from the CEOP database toward ADMIP experiments**
- 0:15 15:50 - 16:05 **YANG Kun: Soil Property DATA at Tongyu Station**
- 0:30 16:05 - 16:35 **Discussion**
- 2:00 18:00 - 20:00 **Dinner**

14 July, 2011 (Thursday)

DAY 2- morning

- Discussion 2: ADMIP Protocol and Plan for 2nd Year (Chair: J. Asanuma)**
- 2:00 10:00 - 12:00 **Discussion**



APN-MAIRS Joint Workshop on Intercomparison of Land surface Process
Modelling at Asian Drylands
13-14 July 2011, Lanzhou, China

Introduction to target site candidates (Chair: K. Mabuchi)
30 minutes each talk

1:30 12:00 - **Lunch**

DAY 2- evening

TBA **Dinner hosted by CARRERI**



3.2 List of Participants

**The 2nd APN-MAIRS joint workshop on Intercomparison of Land surface Process
Modelling at Asian Drylands
13-14 July 2011, Lanzhou, China**

List of Participants

Dr. Ailikun
Director
MAIRS-IPO
c/o Institute of Atmospheric Physics
Chinese Academy of Sciences
P.O. Box 9804, Beijing 100029
P. R. China
Tel: +86-10-82995264
Fax: +86-10-82995161
E-mail: ailli@mairs-essp.org; ailli@tea.ac.cn

Dr. Jun ASANUMA
Assoc. Professor
Terrestrial Environment Research Center
Tsukuba University
Voice&Fax: 029-853-6704
E-mail: asanuma@suiri.tsukuba.ac.jp

Dr. Purevjav GOMBOLUDEV
Climate Sec.
Institute of Meteorology and Hydrology
Juulchiny gudamj 5, Ulaanbaatar 46, Mongolia
Tel: 976-11-326606
Fax: 976-11-326614
E-mail: p_gombo@hotmail.com

Prof. Jinky HONG
Global Environmental Lab, Yonsei University
Seoul, Korea 120-749
Tel: +82-2-2123-7576
Fax: +82-11-9987-3897
E-mail: jkhong@yonsei.kr

Prof. Kazuhito ICHII
Faculty of Symbiotic Systems Science,
Fukushima University
1 Kanayagawa, Fukushima, 960-1296, Japan
Tel: +81-24-548-5256
Fax: +81-24-548-5256
E-mail: ichii@sss.fukushima-u.ac.jp

Mr. Masayuki KONDO
Faculty of Symbiotic Systems Science,
Fukushima University
1 Kanayagawa, Fukushima, 960-1296, Japan
E-mail: masa@sss.fukushima-u.ac.jp

Dr. Kazuo MABUCHI
Geochemical Research Department,
Meteorological Research Institute
1-1 Nagamine, Tsukuba-city, Ibaraki 305-0052
Japan
Tel: +81-29-853-8722
Fax: +81-29-853-8728
E-mail: kmabuchi@mri-jma.go.jp

Prof. S. MIYAZAKI
Graduate School of Environmental Science &
Division of Environmental Resources,
Graduate School of Agriculture, Hokkaido
N10W5, Sapporo, 060-0810, Japan
E-mail: miyashin@ees.hokudai.ac.jp

Prof. Dennis OJIMA
The Heinz Center,
900 17th St. NW, Washington, DC 20006,
United States
E-mail: ojima@heinzctr.org

Ms. Y. SAWAGUCHI
Terrestrial Environment Research Center,
University of Tsukuba
Email: y_sawa@suiri.tsukuba.ac.jp

Prof. Shufen SUN
LASG
Institute of Atmospheric Physics
Chinese Academy of Sciences
P.O. Box 9804, Beijing 100029
P. R. China



E-mail: ssf@lasg.iap.ac.cn

Dr. Yaoming SONG
Nanjing University of Information Science and
Technology (NUIST)
E-mail: sym35@163.com

Prof. Kaoru TACHIIRI
Frontier Research Center for Global Change,
Japan Agency for Marine-Earth Science and
Technology
3173-25 Showamachi, Kanazawa-ku,
Yokohama City
Kanagawa 236-0001, Japan
Tel: +81-45-778-5497
Fax: +81-45-778-5701
E-mail: tachiiri@jamstec.go.jp

Dr. Shuyu WANG
College of Atmospheric Sciences, Nanjing
University
22 Hankou Road, Nanjing 210093
P. R. China
E-mail: wsy@nju.edu.cn

Prof. Jun WEN
Cold and Arid Regions Environmental and
Engineering Research Institute (CAREERI)
Lanzhou, China
Tel. +86 931 4967105
Fax +86 931 4967105
E-mail: jwen@lzb.ac.cn

Prof. Kun YANG
Institute of Tibetan Plateau Research, Chinese
Academy of Sciences
Shuangqing Rd. 18 Haidian District
Beijing 100085, China
E-mail: yangk@itpcas.ac.cn

Ms. Ying YANG
MAIRS-IPO
c/o Institute of Atmospheric Physics
Chinese Academy of Sciences
P. O. Box 9804, Beijing 100029

P. R. China
Tel: +86-10-82995162
Fax: +86-10-82995161
E-mail: yangying@mairs-essp.org;
sec@tea.ac.cn

Prof. Kazuaki YOROZU
Hydrology and Water Resources Engineering
Laboratory, Department of Urban and
Environmental Engineering
Kyoto University
C1, Nishikyo-ku, Kyoto 615-8540, Japan
Tel: +81-75-383-3363
Fax: +81-75-383-3360
E-mail: yorozu@hywr.kuciv.kyoto-u.ac.jp

Prof. Lujun ZHANG
College of Atmospheric Sciences, Nanjing
University
22 Hankou Road, Nanjing 210093
P. R. China
Tel: 025-82594850
E-mail: ljjzhang@nju.edu.cn

Dr. Xia ZHANG
Institute of Atmospheric Physics
Chinese Academy of Sciences
P.O. Box 9804, Beijing 100029
P. R. China
Tel: +86-10-82995111
E-mail: zhangx@tea.ac.cn



3.3 Presentation slides

 <p style="text-align: center;"> Welcome Address and Current Status of ADMIP (Asian Dryland Model Intercomparison Project) </p> <p style="text-align: center;"> Terrestrial Environment Research Center, U. Tsukuba J. Asanuma </p>	<p style="text-align: center;">History of ADMIP</p> <ul style="list-style-type: none"> • Jul., 2009: Proposed at MAIRS-dryland meeting at Changchun • drafting proposals to APN • Apr. 2010: APN project approved • Jul., 2010: ADMIP kickoff meeting at Beijing – summary of the meeting follows.
<p style="text-align: center;">Goals of ADMIP</p> <ul style="list-style-type: none"> • Evaluating & improving land surface models (LSM, energy and water) & terrestrial ecosystem model (TEM, carbon) through offline model intercomparison using data obtained at Asian dryland – but not to improve a model/models specific to Asian dryland • Better reproduction/prediction of landsurface state at Asian dryland using improved models • and capacity building for them. <p style="font-size: small;">©2010 WCI ADMIP/Stacks/8 Staging</p>	<p style="text-align: center;">Target sites, periods</p> <ul style="list-style-type: none"> • Tongyu (contact: Ailikun) – period: 2003-2009 • KBU (contact: J. Asanuma) – period: 2003-2007 • Pingliang(contact: Jun WEN) – period: 2004-2008 • > candidate in the future stage
<p style="text-align: center;">Experimental setup</p> <ul style="list-style-type: none"> • Stage 0: regional analyses of existing data sets – to evaluate current overall performance. – with GSWP2, GLDAS, CEOP-MOLTS(?) – in parallel with Stage 1. (anyone?) • Stage 1: simple & straightforward comparison – comparison within LSMs and TEMs (reproducibility issue) – comparison between ET,NEE, etc from LSMs, TEMs and empirical eqns. (complexity issue, benchmarking?) – multi-model ensemble vs single model (ensemble) 	<p style="text-align: center;">Blind policy</p> <ul style="list-style-type: none"> • w/ default parameters – “default” – what is default parameters? – depends on degree of empiricity • w/ calibrated parameters (what is obj-func?) – best runs! what is best? – should be documented – “most frequent” results, ensemble analyses (volunteers needed)
<p style="text-align: center;">Initialization/spin-up and parameter selection</p> <ul style="list-style-type: none"> • Initialization & spin-up – soil, biomass – left to modelers with the aid of existing data • Parameter selection – Determined from obs. when reliable obs. available – otherwise, left to modelers • Each modelers are requested to submit a document to describe model initialization and parameter selection and calibration (fill in the form) 	<p style="text-align: center;">Data Issues</p> <ul style="list-style-type: none"> • Data format & Archiving – File format: ALMA-compliant netCDF – archived at PALS at U. NSW. • Data policy – All forcing data & model outputs will be shared by the participants of ADMIP – Key paper will be published with all participants as authors, as well the data provider. – At 1year after the submission (or publication), forcing and input data (if possible, observed data) & model output will be available to public



After the Kickoff ...

- Jul, 2010: Kick off workshop at Beijing
- Nov, 2010: Data sampled at Tongyu by Dr. Yang K.
- Feb, 2011: meeting at Kyoto between Dr. Yorozu and J. Asanuma about data preparation.
- Jun, 2011: Atmospheric forcing data became available at web.
 - PALS at U of NSW, Australia is ready (?)
- Jul, 2011: 2nd WS at Lanzhou

Purpose of this workshop

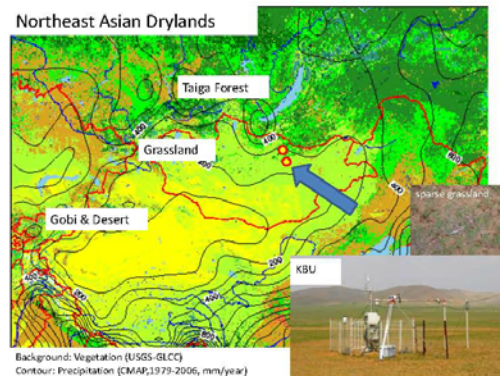
- to review **the first results** (1st day)
- to discuss **data** issues (1st day)
 - data policy
 - data preparation and quality
- to outline **future activities** (2nd day)
 - more detailed plan for the intercomparison
 - meetings during the 2nd year

Note:

- Powerpoint slides will be shared among the participants. If this is not preferable, remove those slides that should be confidential.
- Those who are supported by APN fund at U. Tsukuba, please contact with Ms. Sawaguchi for necessary documents.

Processing Data from KBU

J. Asanuma



Climate, land cover, vegetation

- Topography:
 - plane & flat, 1235m asl
- Climate: semi-arid
 - temp: 1.2 degC, precip. : 200 mm/year
- Landuse /cover
 - uniform & sparse grassland (grazed overly)
- Vegetation
 - dominant species: *Stipa krylovii*, *Carex sp.*, *caragana stenophylla*, *Cleistogenes squarrosa*, *Artemisia frigida*
 - LAI < 1, canopy height=0.1-0.2m
- Soil: chestnut soil, seasonally frozen

Observation

- started on Mar., 2003 (soil obs. from fall, 2002)
- registered as
 - CEOP 2 reference site
 - Asiaflux-registered site
- next to operational met. station (data available)
- initially designed to evaluate grazing effect on surface heat, water and carbon exchange
 - fenced and unfenced
 - later, fence was broken





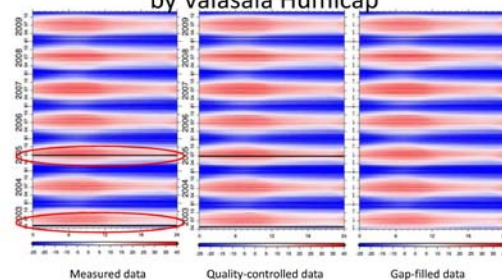
Processing the data

- AsiaFlux procedures
 - 1. Quality control to remove dubious data to create "gaps" in data series (QC)
 - 2. Filling "gaps" by using statistical (correlation) method (GF)
 - except for precipitation which were filled with operational measurements

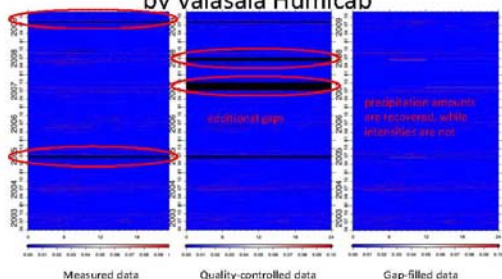
Know problems

- There are mainly 2 periods of missed observations
- Precipitation
 - Cannot distinguish between snow and rain
 - Significant undercatch of snow in winter
 - Longer period of missing precipitation missing than others
 - Filled with nearby operational (manual) measurements with 3 times in a day (0, 6, 12 UTC)
- There are misinformation about sensor heights at Asiaflux data site

Temperature measured by Vaisala Humicap



Temperature measured by Vaisala Humicap



Discussion on data issue

- 1. Target Period Change
- 2. Data Policy
- 3. Data Group
- 4. Data and information needed
- 5. Miscellaneous



<p style="text-align: center;">1. Target Period Change</p> <p>1) Tongyu (China) target period: 2003–2009 2003–2004</p> <p>2) Kherlen Bayan Ulaan (KBU, Mongolia) target period: 2003–2007 2003–2009 or possibly 2010</p>	<p style="text-align: center;">2. Data policy</p> <div style="background-color: #f08080; padding: 10px;"> <p>• 6. 4) Data sharing policy (as of Jul. 2010)</p> <p>– First, all forcing data and model outputs will be shared by the participants of ADMIP. A key paper will be published with all participants as well as the data provider as the authors. At 1 year after the publication of the key paper, forcing and input data that were used to run the models and the model outputs will be available to the public.</p> </div> <p style="text-align: right; font-size: small;">from Kickoff meeting summary</p>
<p style="text-align: center;">2. Data policy</p> <p>• 6. 4) Data sharing policy (revised)</p> <p>– First, all forcing data and model outputs will be shared by the participants of ADMIP. A key paper will be published with all participants as well as the data provider as the authors. At 1 year after the publication of the key paper, forcing and input data that were used to run the models and the model outputs will be available to the public, under the condition that it complies the data policy of the data sources.</p>	<p style="text-align: center;">3. Data group</p> <div style="background-color: #f08080; padding: 10px;"> <p>• The data working group will be organized to prepare necessary data. All of submission of the data will be directed to the data working group at its data focal point to check its format and quality before being archived at PALS (see below)</p> <p>– Jun Asanuma(convener & KBU site), Kazu Yorozu (data focal point, analyses), Ailikun (Tongyu site), Jun Wen (Pingliang site), Jason Evans (liaison to PALS), William Parton (adviser)</p> </div> <p>Need more volunteer for data group such as satellite analyses: CAREERI group soil properties at Tongyu: Yang K. at KBU: J. A.</p>
<p style="text-align: center;">4. Data & information needed</p> <ul style="list-style-type: none"> • 1. Period of gap filled <ul style="list-style-type: none"> – in text Dr. Yorozu and J. Asanuma – Grazing history: heavy, moderate, seasonal change <ul style="list-style-type: none"> • Tongyu: Aili • KBU: J. Asanuma – root depth: <ul style="list-style-type: none"> • KBU : Dr. Kondo • Tongyu: • 2. CO₂ concentration and flux at Tongyu: N/A <ul style="list-style-type: none"> – flux for validation – Ailikun in charge. 	<p style="text-align: center;">5. Misc</p> <ul style="list-style-type: none"> • Who is the contact at Tongyu with respect to CEOP policy? <ul style="list-style-type: none"> – Dr. huizhi Liu, Prof. Congbin Fu



Discussion 1 (2011/07/13)

J. Asanuma



Discussion on data issue

- 1. Target Period Change
- 2. Data Policy
- 3. Data Group
- 4. Data and information needed
- 5. Miscellaneous

1. Target Period Change

1) Tongyu (China)

target period: ~~2003-2009~~
2003-2004

2) Kherlen Bayan Ulaan (KBU, Mongolia)

target period: ~~2003-2007~~
2003-2009 or possibly 2010

2. Data policy

• 6. 4) Data sharing policy (as of Jul. 2010)

- First, all forcing data and model outputs will be shared by the participants of ADMIP. A key paper will be published with all participants as well as the data provider as the authors. At 1 year after the publication of the key paper, **forcing and input data that were used to run the models** and the model outputs will be available to the public.

from Kickoff meeting summary

2. Data policy

• 6. 4) Data sharing policy (revised)

- First, all forcing data and model outputs will be shared by the participants of ADMIP. A key paper will be published with all participants as well as the data provider as the authors. At 1 year after the publication of the key paper, forcing and input data that were used to run the models and the model outputs will be available to the public, under the condition that it complies the data policy of the data sources.

3. Data group

- The data working group will be organized to prepare necessary data. All of submission of the data will be directed to the data working group at its data focal point to check its format and quality before being archived at PALS (see below)
 - Jun Asanuma (convener & KBU site), Kazu Yorozu (data focal point, analyses), Ailikun (Tongyu site), Jun Wen (Pingliang site), Jason Evans (liaison to PALS), William Parton (adviser)

Need more volunteer for data group
such as satellite analyses: CAREERI group
soil properties at Tongyu: Yang K.
at KBU: J. A.

4. Data & information needed

- 1. Period of gap filled in text: Dr. Yorozu and J. Asanuma
 - Grazing history: heavy, moderate, seasonal change
 - Tongyu: Aili
 - KBU: J. Asanuma
 - root depth:
 - KBU: Dr. Kondo
 - Tongyu:
- 2. CO₂ concentration and flux at Tongyu: N/A
 - flux for validation
 - Ailikun in charge.



5. Misc

- Who is the contact at Tongyu with respect to CEOP policy?
 - Dr. huizhi Liu, Prof. Congbin Fu

Discussion 2 (2011/07/14)

J. Asanuma
Terrestrial Env. Research Ctr, U.
Tsukuba

Note:

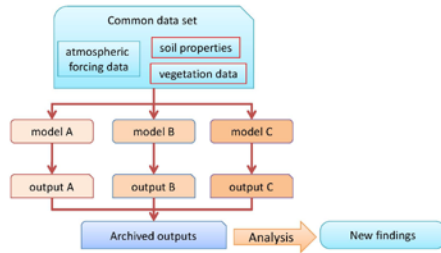
- Powerpoint slides will be shared **among the participants of ADMIP**. If this is not preferable, remove those slides that should be confidential.
- Powerpoint slides will be converted into PDF, and will be placed on ADMIP web page with password protection.

BLIND POLICY

- Now, I am preaching to the converted, or teaching a fish to swim.
- 班门弄斧
- 釈迦に説法



Model Intercomparison



2010/7/11

ADMIP Kickoff, Beijing

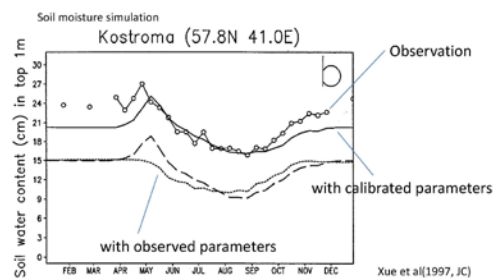
Model Parameters

- The value of model parameter sets significantly influence model results.
- The meaning of a model intercomparison differ largely with the parameter sets we give to the model

Model parameters

- There are 2 kinds of parameter sets:
 - 1) “derived” or “uncalibrated” parameter sets
 - derived independently from the model, such as observation, manually or remotely, or estimated
 - every (or some) model has a “default” parameter sets when no information is available.
 - 2) “calibrated” or “optimized” parameter sets
 - derived so that the model performs well, through a optimization

Example: SSiB application in Russia



Model intercomparison differ with different parameter set

- with a “default” parameter set
 - intercomparison between models evaluates
 - how the models perform when there is no information
 - roughly, “model performance in GCM”, or large scale application
- with a “calibrated” parameter set
 - intercomparison between models evaluates
 - how the models perform when they are “tuned”, or whey they knows “the answer”
 - Even when tuned, models may differ.

Blind policy

- Blind policy defines
 - to what degree each model was given information about soil/vegetation properties and output variables
 - or “how the model operators are kept blind”

Blind policy in ADMIP proposed at Kickoff meeting in 2010

Stage 1: “observed” parameter set

- Parameters in the models will be determined from observations if reliable information is available. Otherwise, the decision is left to the model operators. No calibration will be done in this stage. All of the model operators will be requested to report their parameter selection.

Stage 2: “calibrated” parameter set

- Calibration will be performed by tuning the model results with the observations. All of the model operators will be requested to report their parameter selection.

Blind policy in ADMIP being proposed, Jul. 2011

please refer to [admip:53]

a) Stage 0.5: “default” parameter set(totally blind, new): Oct 30th

- No observed information is available. using the parameters that are normally used without using measured parameters. the decision is left to the model operators.
 - climatology LAI derived from MODIS (only for LSMs)



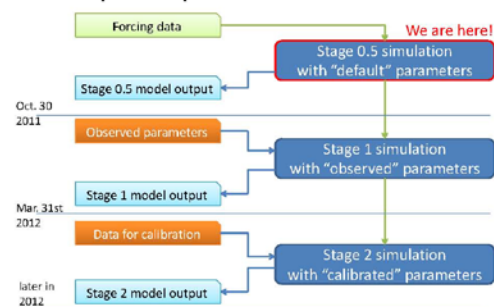
- b) Stage 1: “observed” parameter set (half blind)
Parameters in the models will be determined from observations if reliable information is available. Otherwise, the decision is left to the model operators. **No calibration** will be done in this stage. All of the model operators will be requested to report their parameter selection.
- LAI, albedo through MODIS, at every 500m(LAI) and 5km(albedo), 8days, time series
 - At KBU, LAI will be calibrated with insitu obs.
 - LAI only for LSMs
 - Root profile

- c) Stage 2: “calibrated” parameter set (calibrated)
Calibration will be performed by tuning the model results with the observations. All of the model operators will be requested to report their parameter selection, or how they calibrated the parameters
- what is objective function for the calibration?

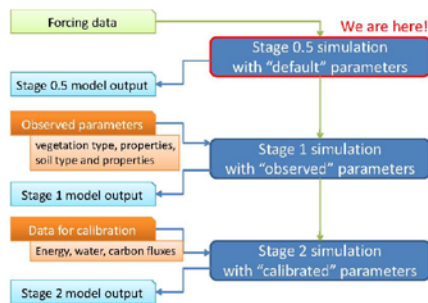
LAI data preparation

- MODIS LAI, at 500m, 8days will be used
- 2 data sets
 - Climatological LAI from data of 10 years
 - Time series
- by J. Wen group and Kondo group

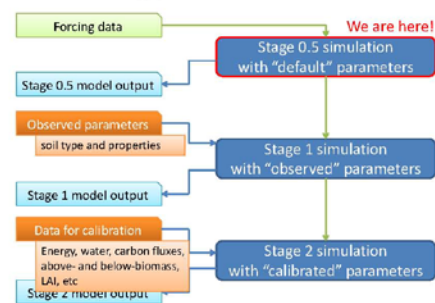
Proposed procedures & deadline



Data being provided for LSMs



Data being provided for TEMs



Blind policy procedures

- Procedures will be controlled through password of the homepage
 - after submitting the output, password for the parameters at the next stage will be given.
- Important deadlines
 - Total blind results: Oct 30th, 2011
 - Half blind results: Mar. 31, 2012

Participating models: unique IDs

- currently 18 models has been registered to the project.
 - the model should be uniquely identified with an acronym
 - different versions need to be distinguished
 - modifications should be distinguished from the original



Acronyms as a unique ID for a model

- acronym will be assigned to each model, and used to identify the model output files.
 - acronyms are in a form:
<model name><version>[_<modified organization>] examples)
 - “CLM40”:
 - Community Land Model ver. 4.0, original version as developed at NCAR
 - “CLM40_UTKB”:
 - Community Land Model modified at U. Tsukuba, after the original version of ver. 4.0

Plan for 2nd year activities 1

- Keep going, and keep going.
 - more data to be opened
 - analyses to be done
- Interim meeting with closed members
 - Somewhere north in Japan, around Nov. or early Dec. 2011
 - hosted by Dr. Miyazaki
 - To check the progress of the project
 - Visa process will be started in Sep.



Plan for 2nd year activities 2

- 3rd ADMIP WS
 - Colorado, USA in last half of May, 2012
 - hosted by Dr Ojima
 - To summarize the results from the 2 year project.
 - To scope the next dimension
 - Visa process will start Feb.
- Possibly, proposing a session at fall AGU, 2012, or somewhere.



3.4 Meeting Summaries

2nd International Workshop of Asian Dryland Model Intercomparison Project

NingWoZhuang Hotel, Lanzhou, China
July, 13-14, 2011
meeting summary

The 2nd WS of ADMIP was held at NingWoZhuang Hotel, Lanzhou with the support from the local host, CARREERI, and MAIRS-IPO.

1. Current status of the project

Events since the kick-off workshop at Beijing on July 2010 were as below.

Jul. 2010: Kickoff meeting (1st workshop) at Beijing

Nov. 2010: Soil samples were collected at Tongyu by Dr. Yang K.

Feb. 2011: Data working group meeting (J. Asanuma and Dr. Yorozu) at Kyoto.
The plan for data opening was decided.

Jun. 2011: Atmospheric forcing data became available at the web site. (Start of Stage 0.5, described below)

Jul. 2011: 2nd workshop at Lanzhou

In addition, uncertainties regarding data availability at the Tongyu site were removed, thanks to Dr. Ailikun at MAIRS-IPO.

2. Data issues reviewed

In light of the updated availability of the observed data, it was agreed to revise the data policy decided during the kickoff meeting.

a) Target sites and periods revised.

1) Tongyu (China)

target period: 2003-2004

2) Kherlen Bayan Ulaan (KBU, Mongolia)

target period: 2003-2009, or possibly 2010 (heavy draught in 2010 summer).

Pinglian remains as a future candidate.

b) Data sharing policy revised

The data sharing policy was revised, as Tongyu data is only available through the



CEOP database, which has its own data policy.

First, all forcing data and model outputs will be shared by the participants of ADMIP. A key paper will be published with all participants as well as the data provider as the authors. At 1 year after the publication of the key paper, forcing and input data that were used to run the models and the model outputs will be available to the public, under the condition that it complies the data policy of the data sources. (the underlined text was added)

3. Reinforced Data Working Group and its tasks.

As the data working group is a bottle neck of the whole project, it will be reinforced for each of the tasks below.

a) LAI and albedo from MODIS:

CAREERI group (Dr. J. Wen) and Mr. Kondo (Fukushima U.) will be in charge. LAI (at 500 m resolution) and albedo (at 5km resolution) will be derived from MODIS products. 2 data sets, namely

- 1) climatological values from 10-years data for blind stage 0.5, and
- 2) time series at every 8 days,

will be constructed. At KBU, LAI will be calibrated with in situ observations.

b) Soil properties

Dr. Yang K.(Tongyu) and Dr. Asanuma (KBU) will derive parameters needed.

c) Grazing history:

Dr. Aili (Tongyu) and Dr. Asanuma (KBU), Intensity, either of heavy, moderate or none, and its seasonal change in a descriptive form should be specified.

d) Root profile (depth) from the literature (at KBU)

Mr. Kondo.



4. New Blind policy

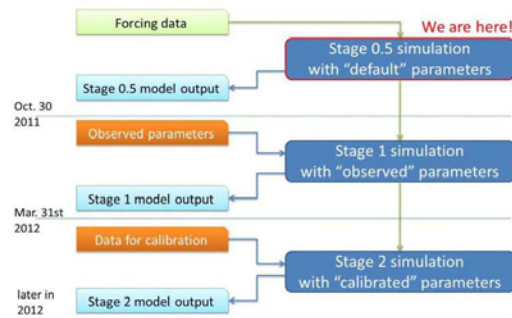


Fig. 1: Schematic diagram for the new blind policy in ADMIP

A new and refined blind policy was adopted in this meeting, and a new stage 0.5 was added in the beginning. The data that are being provided at each stage are now specified in detail. Procedures of the blind policy will be controlled through passwords of the homepage.

a) Stage 0.5: w/"default" parameter set (**NEW**):

currently open, deadline Oct.30th, 2011

"Totally-blind" stage. No observed information is available at this stage. The parameter values without any measured knowledge will be used. The decision is left to the model operators. Data being provided will be

- for LSMs, climatological values of LAI derived from MODIS
- for TEMs, none

b) Stage 1: w/"observed" parameter set:

will be open Nov. 1st, 2012- Mar 31st, 2012

"Half-blind" stage. Parameters in the models will be determined from observations if reliable information is available. Otherwise, the decision is left to the model operators. All of the model operators will be requested to report their parameter selection. Data being provided will be:

- for LSMs, LAI (from MODIS); albedo; vegetation type & properties including root profiles; soil type and properties
- for TEMs, soil type and properties

c) Stage 2: w/"calibrated" parameter set:

will be open Apr 1st, 2012 to later 2012

"Calibration" stage. Calibration will be performed by tuning the model results with



the observations. All of the model operators will be requested to report their parameter selection, or how they calibrated the parameters. The objective function for this calibration will be a matter of future discussion. Data being provided for calibration will be, in addition to those listed in stage 1,

for LSMs: energy, water and carbon fluxes

for TEMs: energy, water and carbon fluxes; above- and below-ground biomass;
LAI

5. Miscellaneous

a) Unique IDs to participating models

An acronym will be given to each of participating models as an unique ID. This ID will be used in the file name for the model outputs. It is in the form,

<model name><version>[_<modified organization>][<modified organization>]

ex1) "CLM40"

Community Land Model ver. 4.0, original version as developed at NCAR

ex2) "CLM40_UTKB13"

Community Land Model modified at Univ. Tsukuba after the original version of ver. 4.0. Its modified version of 1.3.

A list of specific ID for each model will be presented later.

6. Plan for the 2nd year and the future.

a) Interim meeting in late 2011 at Hokkaido Univ., Sapporo, Japan

A closed interim meeting is planned at Hokkaido University, Sapporo, Japan in late Nov. or early Dec, 2011, to check the progress of the project. Kindly hosted by Dr. Miyazaki at Hokkaido Univ.

b) 3rd international workshop of ADMIP in May, 2012 at Colorado, USA

This is planned in the last half of May, 2012, at Colorado, USA, to summarize the outcome of the 2 year project, and to scope the next moves. Kindly hosted by Prof. Ojima.

Proposing a session at the fall AGU meeting in 2012 is also under consideration.



3.5 Pictures





4.0 Asian Dryland Model Intercomparison Project Interim meeting
Dec 1st-2nd, 2011, No.3 meeting room in "Conference Hall", Hokkaido Univ., Sapporo, Japan

4.1 Agenda

Asian Dryland Model Intercomparison Project	
Interim meeting	
Dec 1st-2nd, 2011, No.3 meeting room in "Conference Hall", Hokkaido Univ., Sapporo, Japan	
hosted by IFES-GCOE Program "Establishment of Center for Integrated Field Environmental Science", Hokkaido University	
live broadcast at: http://www.ustream.tv/channel/admip	
Tentative Agenda	
<u>1 Dec, 2011 (Thursday)</u>	
DAY 1- morning	
JST	(UTC)
9:30 - 10:00 (0:30 - 1:00)	Registration
	Opening (Chair: Ailikun)
10:00 - 10:10 (1:00 - 1:10)	Welcome address and the scope of the meeting (Jun
10:10 - 10:40 (1:10 - 1:40)	Shin MIYAZAKI: Introduction of activities by IFES-GCOE Program, Hokkaido University
10:40 - 11:30 (1:40 - 2:30)	Tomonori SATO: Summertime moisture source in Northeast Asia -implications to continental water recycling-
11:30 - 12:20 (2:30 - 3:20)	Byambakhuu Ishgaldan: Simulation of ecohydrological processes in Mongolian semiarid region.
12:20 - 13:50 (3:20 - 4:50)	Lunch (group photo)
DAY 1- afternoon	
	Status of ADMIP and its first results (Chair: Mabuchi)
13:50 - 14:10 (4:50 - 5:10)	Jun ASANUMA: Status of the Project
14:10 - 14:25 (5:10 - 5:25)	Dennis OJIMA:
14:25 - 14:45 (5:25 - 5:45)	Kaoru TACHIIRI
14:45 - 15:05 (5:45 - 6:05)	YANG Kun
15:05 - 15:25 (6:05 - 6:25)	Kazuaki YOROZU: Preliminary comparison of model results on ADMIP stage 0.5
15:40 - 15:55 (6:40 - 6:55)	Coffee break
	Discussion 1: Preparation for the next Stage (Chair:
16:10 - (7:10 -)	Masayuki KONDO & Jun WEN: LAI data preparation
	TBA
18:00 - 20:00	Dinner
<u>2nd Dec, 2011 (Friday)</u>	
DAY 2- morning	
	Discussion 2: Way forward (Chair: J. Asanuma)
10:00 - 12:00 (1:00 - 3:00)	Discussion

GSEES Seminar, hosted by Hokkaido University	
14:45 - 16:15	Dr. Ailikun (Institute of Atmospheric Science, Chinese Academy of Science) : From Physical Processes to Social Sustainability -Introduction of MAIRS Research Activity in East
Place:D102, Graduate School of Environmental Earth Science, Hokkaido University	



4.2 List of Participants

ADMIP (Asian Dryland Model Intercomparison Project) interim meeting
2011

1st and 2nd, Dec., 2011 at Hokkaido University in Sapporo city, Japan

List of Participants

Prof. Jun ASANUMA
Terrestrial Environment Research
Center, University of Tsukuba Tsukuba,
305-8577, JAPAN
Tel: +81-298-53-6704
Fax: +81-298-53-2530
Email: asanuma@suiri.tsukuba.ac.jp

Dr. Ailikun
International Project Office, Monsoon
Asia Integrated Regional Study
(MAIRS), Institute of Atmospheric
Physics, Chinese Academy of Sciences
40# Hua Yan Li, Qi Jia Huo Zi, Chao
Yang District, Beijing 100029, China
Tel: +86-10-82995264
Fax: +86-10-82995161
Email: aili@mairs-essp.org
ailikun@hotmail.com

Dr. Byambakhuu Ishgaldan
Mongolyn Alt (MAK), Industrial Policy
Department
Email: ibyamba@gmail.com

Dr. Masayuki KONDO
Faculty of Symbiotic Systems Science,
Fukushima University
1 Kanayagawa, Fukushima, 960-1296,
Japan
Email: masa@sss.fukushima-u.ac.jp

Prof. Yang KUN
Institute of Tibetan Plateau Research,
Chinese Academy of Sciences
Tel: +86-10-62849727
Email: yangk@itpcas.ac.cn

Dr. Kazuo MABUCHI
Geochemical Research Department,
Meteorological Research Institute 1-1
Nagamine, Tsukuba-city, Ibaraki
305-0052 Japan
Tel: +81-29-853-8722
Fax: +81-29-853-8728
Email: kmabuchi@mri-jma.go.jp

Prof. Shin MIYAZAKI
Graduate School of Environmental
Science & Division of Environmental
Resources, Graduate School of
Agriculture, Hokkaido University
N10W5, Sapporo, 060-0810, Japan
Tel: +81-11-706-3356
Fax: +81-11-706-4867
Email: miyashin@ees.hokudai.ac.jp

Prof. Denis OJIMA
Natural Resource Ecology Laboratory,
Colorado State University
Fort Collins, CO 80523
Tel: 970 491 1976
Email: ojima@heinzctr.org



Prof. Tomonori SATO
Faculty of Environmental Earth
Science, Hokkaido University
North 10, West 5, Sapporo 060-0810
JAPAN
Email: t_sato@ees.hokudai.ac.jp

Terrestrial Environment Research
Center, University of Tsukuba
Email: y_sawa@suiiri.tsukuba.ac.jp

Dr. Kaoru TACHIIRI
Frontier Research Center for Global
Change, Japan Agency for
Marine-Earth Science and Technology
3173-25 Showamachi, Kanazawa-ku,
Yokohama City, Kanagawa 236-0001,
Japan
Tel: +81-45-778-5497
Fax: +81-45-778-5701
Email: tachiiri@jamstec.go.jp



Prof. Jun WEN
Cold and Arid Regions Environmental
and Engineering Research Institute,
Chinese Academy of Sciences
Tel: +86-931-4967109
Email: jwen@lzb.ac.cn

Prof. Kazuaki YOROZU
Hydrology and Water Resources
Engineering Laboratory, Department
of Urban and Environmental
Engineering, Kyoto University
C1, Nishikyo-ku, Kyoto 615-8540,
Japan
Tel: +81-75-383-3363
Fax: +81-75-383-3360
Email: yorozu@hywr.kuciv.kyoto-u.ac.
jp

Ms. Yasuyo SAWAGUCHI



4.3 Presentation slides

  <p>ADMIP (Asian Dryland Model Intercomparison Project) interim meeting SCOPE of the meeting</p> <p>Terrestrial Environment Research Center, U. Tsukuba J. Asanuma</p>	<p>Purpose of this workshop: to confirm progress of the project</p> <ul style="list-style-type: none">• To review submitted results for Stage 0.5• To confirm the Project protocol• To discuss and confirm data preparation and other issues for stage 1.0
<p>Structure of the meeting</p> <ul style="list-style-type: none">• 3 invited talks in the morning today – 2 from local host• Presentations and discussions on the results for Stage 0.5 in the afternoon today• Discussion on future plan and data preparation issues in the late afternoon today and tomorrow.• Seminar by Dr. Ailikun in day 2 afternoon, 14:45 at different building. Open to the local research community	<p>Note:</p> <ul style="list-style-type: none">• Powerpoint slides will be shared among the participants. If this is not preferable, remove those slides that should be confidential.• Meeting is being live-broadcasted through Internet• Please contact with Ms. Sawaguchi for necessary documents before she leaves in this afternoon.



Current Status of ADMIP

Jun Asanuma
Terrestrial Environment Research Center
University of Tsukuba

APN 2007/11 ADMIP interim Report, 2011 JSPS 日本学術振興会

Asian Drylands: characteristics

- ✳ smaller precipitation
- ✳ larger interannual variation of precip.

APHRODITE precip. - clim. Precipitation Climatology (mm/year)

Interannual variation of precipitation

Coeff. variation

APHRODITE precip. - APHRODITE precip. C.Var. Annual mean

APHRODITE precip. - monthly C.Var. Monthly

larger variation in winter snowfall and albedo

Model performance at the region

ex1) GSWP2 (annual)

correlation between GSWP2 (multi model analysis) and GSMDDB(observation) (annual means)

figures made by Dr. Yorozi at Kyoto Univ.
in reference to Guo, and Dirmeyer (2002) JGR, Vol.111, D22S02

Model performance at the region

ex2) recent models & data sets

- ✳ GLDAS (Global Land Data Assimilation System)
 - ✳ Synthesized atmospheric forcing data:
 - merge of surface & satellite observation and reanalysis
- ✳ Model-to-model variation is large even with the same forcing data

Goals of ADMIP

- ✳ Evaluating & improving land surface models (LSM, energy and water) & terrestrial ecosystem model (TEM, carbon) through **offline model intercomparison** using data obtained at **Asian dryland**
 - but not to improve a model/models specific to Asian dryland
- ✳ Better reproduction/prediction of landsurface state at Asian dryland using improved models
- ✳ and capacity building for them.

2013/7/11 ADMIP Kickoff, Beijing

Target Phenomena and Scientific Questions



- 1. Target Phenomena**
 - ✳ Energy, water and carbon exchange at the land surface under water-limited environment, and/or under the control of vegetation growth, at the temporal scales of 3-hourly, daily, and monthly.
- 2. Scientific Questions**
 - ✳ What is the ability of models to reproduce energy, water, and carbon exchange at dryland surface (reproducibility)?
 - ✳ Do current ecosystem models, developed mainly for forest, reproduce energy, water and carbon exchange?
 - ✳ Does the **current complexity** of models effectively simulate land surface processes? (complexity issues, also related to benchmarking)
 - ✳ Does the **multi-model ensemble** of LSMs outperform single model output? (ensemble)

2 013/7/11 ADMIP Kickoff, Beijing

Target sites, periods

- ✳ Tongyu (contact: Ailikun)
 - period: 2003-2004
- ✳ KBU (contact: J. Asanuma)
 - period: 2003-2009
- ✳ Pingliang(contact: Jun WEN)
 - period: 2004-2008
 - candidate in the future stage

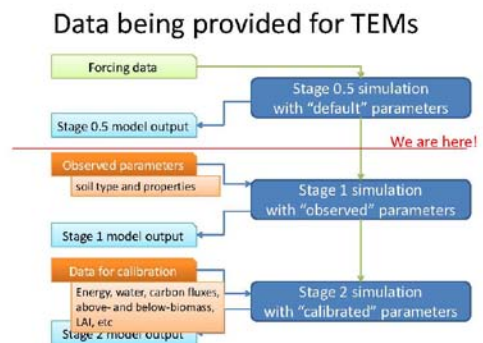
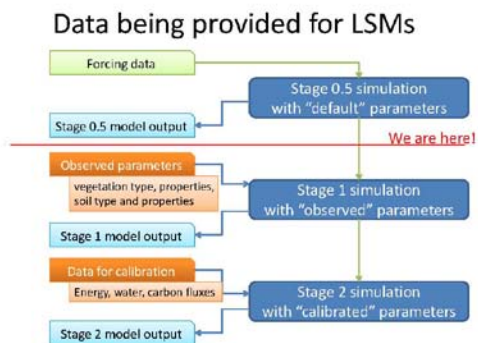
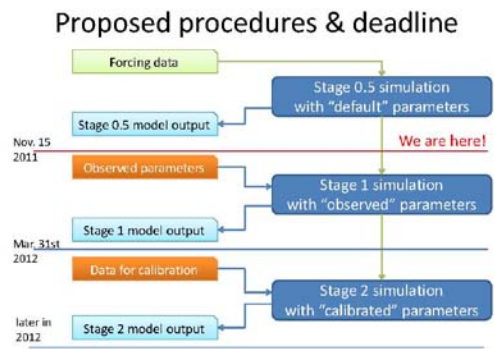


Model Runs (Blind policy)

- ★ Stage 0.5: w/ "default" parameter set :
 - "Totally-blind" stage
 - deadline Nov. 15th, 2011 (just finished)
- ★ Stage 1: w/ "observed" parameter set:
 - "Half-blind" stage
 - to be Nov. 15th, 2011- Mar 31st, 2012
- ★ Stage 2: w/ "calibrated" parameter set:
 - "Calibration" stage.
 - to be open Apr 1st, 2012 to later 2012

2010/7/11 ADMIP Kickoff, Beijing



Stage 0.5:
w/"default" parameter set

- * - Nov. 15, 2011 (Just finished)
- * "Totally-blind" stage.
 - * No observed information is available.
 - * The parameter values without any measured knowledge. The decision is left to the model operators.
- * Data provided
 - * for LSMs,
 - * climatological values of LAI derived from MODIS
 - * for TEMs
 - * none

Stage 1:
w/"observed" parameter set

- * to be Nov. 1st, 2012- Mar 31st, 2012
- * "Half-blind" stage.
 - * Parameters from observations if reliable information is available.
 - * Otherwise, the decision is left to the model operators. All of the model operators will be requested to report their parameter selection.
- * Data provided:
 - * for LSMs
 - * LAI (from MODIS), albedo, vegetation type & properties including root profiles; soil type and properties
 - * for TEMs
 - * soil type and properties

Stage 2:
w/"calibrated" parameter set

- * to be open Apr 1st, 2012 to later 2012
- * "Calibration" stage.
 - * Calibration by tuning the model results with the observations.
 - * Model operators will be requested to report their parameter selection, or their calibration.
 - * The objective function for this calibration will be a matter of future discussion.
- * Data provided for calibration, in addition to stage 1,
 - * for LSMs: energy, water and carbon fluxes
 - * for TEMs: energy, water and carbon fluxes; above- and below-ground biomass; LAI

Output variables

- * Variable names are ALMA compliant.
 - * - from LSMs (Hourly)
 - * Mandatory
 - * [SWUp](#), [LWUp](#), [Qh](#), [Qc](#), [Qg](#), [SoilTemp](#), [Evap](#), [Qs](#), [Qst](#), [SWE](#) and/or [SnowDepth](#), [SoilMoist](#), [WaterTableD](#)
 - * Mandatory if applicable
 - * [GPP](#), [NPP](#), [NEE](#), [AutoResp](#), [HeteroResp](#)
 - * - from TEMs (Daily)
 - * Mandatory
 - * [Qc](#), [SoilTemp](#)
 - * [SoilMoist](#)
 - * [GPP](#), [NPP](#), [NEE](#), [AutoResp](#), [HeteroResp](#), [TotLivBiom](#)
 - * Mandatory if applicable
 - * [SWUp](#), [LWUp](#), [Qh](#), [Qc](#)
 - * [Evap](#), [Qs](#), [Qst](#), [SWE](#) and/or [SnowDepth](#), [WaterTableD](#)

Data working group

- * a. LAI and albedo from MODIS
 - * CAREERI group (Dr. J. Wen) & Fukushima U (Mr. Kondo)
 - * LAI (at 500 m res.) and albedo (at 5km res.) from MODIS products
 - * 1) climatological (LAI) from 10-years data for stage 0.5
 - * 2) time series at every 8 days.
 - * At KBU, LAI will be calibrated with in situ observations.
- * b. Soil properties
 - * Dr. Yang K. (Tongyu) and Dr. Asanuma (KBU)
 - * Parameters derived: b-parameter
- * c. Grazing history
 - * Dr. Aili (Tongyu) and Dr. Asanuma (KBU)
 - * Intensity, either of heavy, moderate or none, and its seasonal change in a descriptive form should be specified.
- * d. Root profile (depth) from the literature (at KBU)
 - * Mr. Kondo.



Things to discuss

- 1. format and convention issues: done
 - will be included into the new version of the protocol
- 2. Spun-up issues
- 3. data preparation
- 4. new schedule
- 5. analyses policy of the model results
- 5. robust protocol?
- 6. next meeting

1. FORMAT AND CONVENTION ISSUE

format and convention issues

- need to specify sign convention
 - upward positive
- soil temperature (skin?, layer?) > Dr. Yorozu
- leap year
 - data at Feb 29th will be dropped after receiving the data.
- when a day starts?
 - UTC 00Z
- need info. on soil layer depth
 - better to be included in netCDF
 - text is also acceptable.
- some outlaws !

2. SPUN-UP ISSUES

Year 2003 is anomalous, globally

Solar radiation anomaly

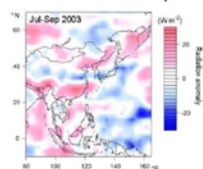


Fig. 5 – The spatial distribution of the anomaly in summer solar radiation (3-month average of July-September) in 2003 obtained from NCEP/NCAR reanalysis data (base period: 1979–2000).

Saigusa et al.(2008)

Temperature anomaly

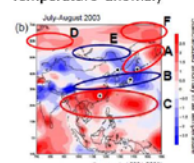
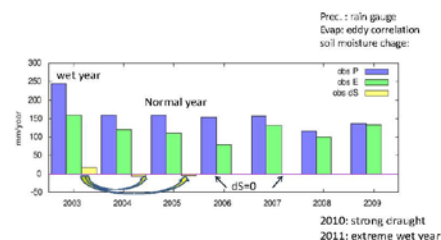


Fig. 4. Spatial distribution of (a) the anomaly in the maximum (Tmax) and (b) the minimum (Tmin) temperature in the summer of 2003 obtained from NCEP/NCAR reanalysis data (base period: 1979–2000). Black circles show the location of the study sites.

Saigusa et al.(2010)

In KBU as well: Observed interannual variation



Proc.: rain gauge
Evap: eddy correlation
soil moisture: chage:

2010: strong drought
2011: extreme wet year

Spun-up taken in the models

- Century:
 - monthly, or daily, 500-1000 years simulation until stabilizes
 - half degree monthly data set.
 - simulation started from Jan. 2004
- Seib-DGVM
 - blended obs and reanalysis,
 - spin-up for few thousand years
 - transient ?? years
- Noah and SiB2
 - first 1 year of the data set
 - Tongyu: 2003, KBU: 2003

Spun-up issue

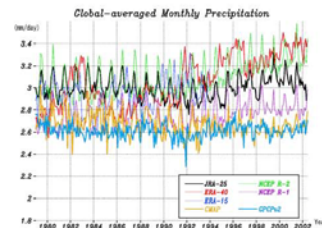
- Long-term data for spun-up forcing will be provided (preferably before Christmas)
 - 1hr or 30min interval, more than 30 yrs
 - LSM:
 - same as forcing
 - TEMs
 - monthly enough, daily is better.
 - few ten years
 - Tmas, Tmin, monthly average.
- method of spun-up should be documented?
 - data should be detrended? left to the model operator?
 - need to create randomized data set
 - who?
- already-ran model results spun-up-ed with 2003 data ?
 - all of models will be re-run for stage 0.5
 - new deadline Jan. 15



Possible data set for spun-up

- Reanalysis (JRA, NCEP, ECMWF et al.)
 - know to be problematic
- Gauge based data set
 - APHRODITE (gauge-based precipitation grid data set)
 - CMAP, GPCP et al
- Some global data sets for LSM forcing
 - by T. Qian @ NCAR (now at U. Michigan)
 - by Hirabayashi @ U. Tokyo
 - by Princeton Univ. (current GLDAS)
 - ITP, temp, wind, hum. prec., down long & short wave

Precipitation in reanalysis



Spun-up data set

- should be consistent with observation (especially for precip.)
- > will be discussed over the lunch this afternoon between J.A, YangKun, and others
- for Tongyu:
 - ITP data set: 1981-2008
 - 3hour to 0.5 hour with linear interpolation,
- for KBU:
 - 1979, by older GLDAS, 3 hourly (data I have? if not contact Y.K.)
 - radiation and prec. corrected with obs at KBU
 - J.A. will do it.

DATA PREPARATION

Stage 0.5: w/"default" parameter set

- - Nov. 15, 2011 (Just finished)
- "Totally-blind" stage.
 - No observed information is available.
 - The parameter values without any measured knowledge. The decision is left to the model operators.
- Data provided
 - for LSMs,
 - climatological values of LAI derived from MODIS >done by U. Fukushima Group
 - for TEMs
 - none

Stage 1: w/"observed" parameter set

- to be Nov. 1st, 2012- Mar 31st, 2012
- "Half-blind" stage.
 - Parameters from observations if reliable information is available.
 - Otherwise, the decision is left to the model operators. All of the model operators will be requested to report their parameter selection.
- Data provided:
 - for LSMs
 - LAI (from MODIS); albedo; vegetation type & properties including root profiles; soil type and properties
 - for TEMs
 - soil type and properties

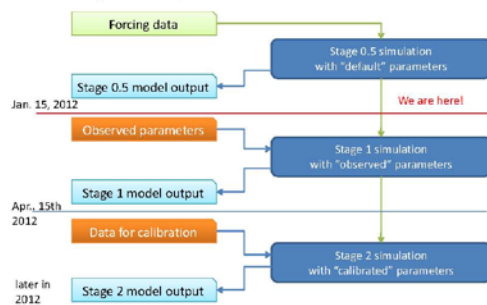
Data preparation (deadline Jan. 15, 2012)

- a. LAI and albedo from MODIS
 - CAREERI group (Dr. J. Wen) & Fukushima U. (Mr. Kondo)
 - LAI (at 500 m res.) and albedo (at 5 km res.) from MODIS products.
 - 1) climatological (LAI) from 10 years data for stage 0.5 (new data set by U. Fukushima)
 - 2) time series at every 8 days for stage 1, (LAI, U. Fukushima; albedo, CAREERI)
 - At KBU, LAI will be calibrated with in situ observations (only with KBU1)
- b. Soil properties
 - Dr. Yang K. (Tongyu, reevaluation) and Dr. Asanuma (KBU, not yet)
 - Parameters derived: b-parameter etc (need someone to compile)
- c. Grazing history:
 - Dr. Aili (Tongyu, done) and Dr. Asanuma & Byanhuu (KBU)
 - Intensity, either of heavy, moderate or none, and its seasonal change in a descriptive form should be specified.
 - Tongyu: heavily overgrazed since 1960's, with increasing grazing pressure upto now. No grazing since 2003 when it was fenced
 - KBU: heavily overgrazed, especially in winter, in summer moderately overgrazed (Byanhuu will investigate)
- d. Root profile (rooting depth) from the literature (at KBU)
 - Tongyu (10-35cm, no root below 35cm), KBU (upto 50 cm)

NEW SCHEDULE



Proposed procedures & deadline



ANALYSIS POLICY ISSUES

2010/7/11

/ADMIP/ICoM/Bojng

4) Data sharing policy

- First, all forcing data and model outputs will be shared by the **participants** of ADMIP.
- A key paper will be published with all participants as well as the data provider as the authors.
- At 1 year after the publication of the key paper, forcing and input data that were used to run the models and the model outputs will be available **to the public**.

Who is the participants?

BAIM	SM
CLM	BBGC
CoLM	SEIB
MATSIRO	VISIT
NCASA	HAL
SCASA	DAYCENT
SiBUC	SSiB2
and data providers	
(CAREERI, Tongyu,	expected
Mongolian operational agency)	Jinkyu Hong

need for new analysis policy

- model outputs will be shared by the participants
- a key paper will be written
- need to promote few scientists into intercomparison analyses
- need to avoid conflict in between scientists
- Before May meeting, develop
 - outline of the key paper
 - clearly objectives of in intercomparison
 - through e-mail
 - somewhere in Feb., we will start to discuss
 - Colorado meeting organization, meeting focus
- At May meeting

New Analysis policy

- Model output is shared
- Analysis proposal made
 - method
 - method of analysis
 - expected results
- Then, produce the results
- Authorship?
 - all of the authors.
- Key analyses in key papers
 - key papers
- > Proposal coordinating session in Colorado meeting.
- need some ways to avoid conflict.

May meeting

- At Colorado, near CSU, Fort Collins.
 - May 17(thr)-18(fri) (19 for excursion)
- Jan., list up participants
- Visa application
 - need to dispatch invitation in the mid Jan.
 - Dennis will write up the draft.
 - visa process will start then

organizing a specific experiment

- Scheme-specific
 - grazing scheme
- will be proposed at mailing list
- keep in track with the main stream schedule
- Need revised model code



recomputation

- Stage0.5 computation
 - year 2003 problem
 - new LAI data
 - new Spin-up data set especially for TEMs
 - deadline is Jan 15th

Appendix 4.4 Meeting Summaries



Asian Dryland Model Intercomparison Project

Interim Meeting

Conference Hall, Hokkaido Univ., Sapporo, Japan
Dec 1st-2nd, 2011

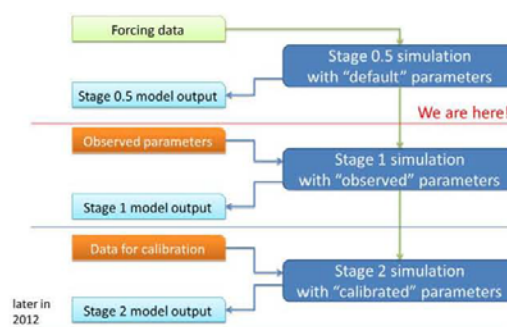
meeting summary

ADMIP interim meeting was held at Hokkaido University, Sapporo, Japan, with the support from IFES-GCOE Program "Establishment of Center for Integrated Field Environmental Science" at Hokkaido University. The below is the summary of the meeting.

1. Current status of the project

Events since the kick-off workshop at Beijing on July 2010 were as below.

Jul. 2010	Kickoff meeting (1st workshop) at Beijing
Nov. 2010	Soil samples were collected at Tongyu by Dr. Yang K.
Feb. 2011	Data working group meeting (J. Asanuma and Dr. Yorozu) at Kyoto. The plan for data opening was decided.
Jun. 2011	Atmospheric forcing data became available at the web site. (Start of Stage 0.5, described below)
Jul. 13-14th, 2011	2nd workshop at Lanzhou
Nov. 15th, 2011	Submission deadline of Stage 0.5
Dec. 1st-2nd, 2011	Interim meeting at Sapporo



At the time of the meeting, output from the following models were submitted.

BAIM, CLM, CoLM, MATSIRO, NCASA, SCASA, SiBUC, SM, Biome-BGC,
SEIB, VISIT, HAL, DAYCENT, SSiB2

and more to be expected.

2. Data format and convention issues

As discussed in the earlier meeting, the format of the forcing/output data should follow the ALMA convention. In addition, additional specifications were agreed as below.

sign convention	flux values should be upward positive
leap year issues	365-day calendar should be used, and, therefore, data at Feb 29th will be dropped after the data submission.
the start of the day	00UTC (0 o'clock in UTC)
soil layer depth	information should be submitted in the netCDF file or in an separate text.
soil temperature	SoilTemp in ALMA convention is larger-averaged soil temperature.

It should be NOTED by all of the participants that ALMA-compliance should be followed strictly.

3. Year 2003 and Spun-up issues

It was found that the warm season precipitation at KBU and Tongyu was uncommonly larger in the year 2003, that is the beginning year of the forcing at KBU and Tongyu. According to some of the participants, when the data in this year was used for the spin-up process of LSMs, anomalous initial condition was brought up. The same can be applied with long-term spin-up of TEMs.

In order to cope with this issue, as well as to provide a common and forcing data set with TEMs, it was agreed that the project will prepare long-term data both at KBU and Tongyu for spun-up, and that the participants will be requested to rerun their model for stage 0.5 by using this data set in its spun-up. The data prepared will be as follows.

At Tongyu:

date set composed by Institute of Tibetan Plateau Research
period 1981-2008, 3 hourly



At KBU,

data set composed by NCAR (Qian et al., 2007, J.Hydro.Met) for
Community Land Model.

period 1948-2004, 3 hourly

In addition, the model participants are requested to document their spin-up
processes.

4. Data preparation

1) for Stage 0.5

- Climatological LAI at 500m resolution is being derived by U. Fukushima group from MODIS-derive LAI over 10 years. It will be updated by new information given during the meeting. Calibration at KBU1, not KBU2, will be done.

2) for Stage 1.0

- Time series of LAI (at 500m horiz. resolution) and albedo (at 5km horizontal resolution) will be derived at every 8 days by U. Fukushima and CAREERI group, respectively.
- Soil properties will be derived by Dr. Yang K. (Tongyu), and Asanuma (KBU). Someone is needed to compile this.
- Grazing history is as follows,
 - ✧ Tongyu: heavily overgrazed since 1960's, with increasing grazing pressure upto now. No grazing since 2003 when it was fenced
 - ✧ KBU: heavily overgrazed, especially in winter, in summer moderately overgrazed (Dr. Byanbahuu will investigate further)
- Root profile (rooting depth) from the literature
 - ✧ Tongyu (10-35cm, no root below 35cm), KBU(upto 50 cm)

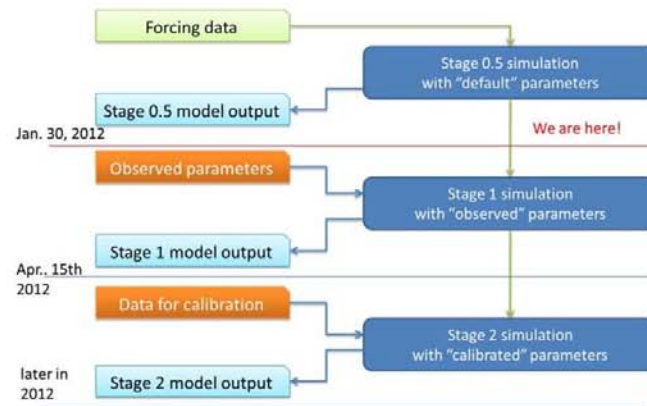
5. New schedule

New schedule was decided as follows:

- Resubmission of stage 0.5: Jan 30, 2012
- Submission of Stage 1 output: Apr, 15th, 2012



- Submission of Stage 2 output: later in 2012



6. Next meeting

- At Colorado State University, in Fort Collins, Colorado, USA
- Meeting during May 17th (Thr) and 18th (Fri), and 19th for excursion.
- In Jan, 2012, the invitation letters will be dispatched by Dr. Ojima, and participants will be listed up. Visa acquisition process will be followed.



Appendix 4.5 Pictures



5.0 The 3rd International Workshop on Asian Dryland Model Intercomparison Project, Colorado, USA 16-18, May, 2012

5.1 List of Participants

The 3rd International Workshop on Asian Dryland Model
Intercomparison Project
16-18, May, 2012, Colorado, USA

List of Participants

Prof. Jun ASANUMA
Terrestrial Environment Research
Center, University of Tsukuba
Tsukuba, 305-8577, JAPAN
Tel: +81-298-53-6704
Fax: +81-298-53-2530
Email: asanuma@suiri.tsukuba.ac.jp

Dr. Ailikun
International Project Office, Monsoon
Asia Integrated Regional Study
(MAIRS)
Institute of Atmospheric Physics,
Chinese Academy of Sciences
40# Hua Yan Li, Qi Jia Huo Zi, Chao
Yang District, Beijing 100029, China
Tel: +86-10-82995264
Fax: +86-10-82995161
Email: aili@mairs-essp.org
ailikun@hotmail.com

Dr. Byambakhuu Ishgaldan
Sustainable development institute at
National University of Mongolia
Main building, University Street 1,
P.O.Box 46/120, Ulaanbaatar 14201,
Mongolia
Tel: 976-11-320159, 77307730
(1138)
Email: i_byambakhuu@num.edu.mn

Dr. Xiaodan Guan
College of Atmospheric Science,
Lanzhou University
222# Tianshui Southroad, Lanzhou,
730000, P.R.China
Tel: +86-931-8913139
Email: guanxd05@lzu.cn

Prof. Kazuhito Ichii
Faculty of Symbiotic Systems Science,
Fukushima University
1 Kanayagawa, Fukushima, 960-1296,
Japan
Tel: +81-24-548-5256
Fax: +81-24-548-5256
Email: ichii@sss.fukushima-u.ac.jp

Dr. Akihiko Ito
Center for Global Environmental
Research, National Institute for
Environmental Studies
16-2 Onogawa, Tsukuba-City, Ibaraki,
305-8506, Japan
Tel: +81-29-850-2981
Fax: +81-29-850-2960
Email: itoh@nies.go.jp

Prof. Yang KUN
Institute of Tibetan Plateau Research,
Chinese Academy of Sciences
Tel: +86-10-62849727
Email: yangk@itpcas.ac.cn



Dr. Qian LI
Institute of Atmospheric Physics
Chinese Academy of Sciences
Email:qian@mail.iap.ac.cn

Dr. Kazuo MABUCHI
Geochemical Research Department,
Meteorological Research Institute
1-1 Nagamine, Tsukuba-city, Ibaraki
305-0052 Japan
Tel: +81-29-853-8722
Fax: +81-29-853-8728
Email: kmabuchi@mri-jma.go.jp

Prof. Shin MIYAZAKI
Graduate School of Environmental
Science & Division of Environmental
Resources, Graduate School of
Agriculture, Hokkaido University
N10W5, Sapporo, 060-0810, Japan
Tel: +81-11-706-3356
Fax: +81-11-706-4867
Email: miyashin@ees.hokudai.ac.jp

Prof. Dennis OJIMA
Natural Resource Ecology Laboratory,
Colorado State University
Fort Collins, CO 80523
Tel: 970 491 1976
Email: ojima@heinzctr.org

Prof. William Parton
Natural Resource Ecology Laboratory,
Colorado State University
Fort Collins, CO 80523
Tel: (970) 491-1987
Fax: (970) 491-1965
Email: billp@nrel.colostate.edu

Dr. Purevjav Gomboluudev
Climate Sec. Institute of Meteorology
and Hydrology
Juulchiny gudamj 5, Ulaanbaatar 46,
Mongolia
Tel: +976-11-326606
Fax: +976-11-326614
Email: p_gombo@hotmail.c

Dr. J. Straube
Natural Resource Ecology Laboratory,
Colorado State University

Dr. Hesong WANG

Prof. Yongkang XUE
1255 Bunche Hall
Department of Geography
Department of Atmospheric & Oceanic
Sciences
University of California, Los Angeles
(UCLA)
Los Angeles, CA 90095-1524, U.S.A.
Tel: (310) 825-1122
Fax: (310) 206-5976
E-Mail: yxue@geog.ucla.edu

Prof. Zong-Liang Yang
Department of Geological Sciences,
The University of Texas at Austin
1 University Station #C1100, Austin,
TX 78712-0254
Tel: 512-471-3824
Fax: 512-471-9425
Email: liang@mail.utexas.edu

Dr. Chen Yingying
Institute of Tibetan Plateau Research,



Chinese Academy of Sciences

Prof. Kazuaki YOROZU
Hydrology and Water Resources
Engineering Laboratory, Department
of Urban and Environmental
Engineering, Kyoto University
C1, Nishikyo-ku, Kyoto 615-8540,
Japan
Tel: +81-75-383-3363
Fax: +81-75-383-3360
Email: yorozu@hywr.kuciv.kyoto-u.ac.
jp

Dr. xia ZHANG
Institute of Atmospheric Physics
Chinese Academy of Sciences
P.O.box9804 TEA
BeiJing 100029 China
Email: zhangx@tea.ac.cn



5.2 Presentation slides

Current Status of ADMIP

Jun Asanuma
Terrestrial Environment Research Center
University of Tsukuba

APN 2007/11 ADMIP Interim Report, 2011 JSPS 日本学術振興会

Asian Drylands: characteristics

- ✦ smaller precipitation
- ✦ larger interannual variation of precip.

Precipitation Climatology (mm/year)

Interannual variation of precipitation

Coeff. variation

Annual mean

Monthly

APHRODITE precip. - monthly C.Var.

larger variation in winter

Model performance at the region

ex1) GSWP2 (annual)

correlation between GSWP2 (multi model analysis) and

figures made by Dr. Yorozu at Kyoto Univ.

Model performance at the region

ex2) recent models & data sets

- ✦ GLDAS (Global Land Data Assimilation System)
- ✦ Synthesized atmospheric forcing data:
 - merge of surface & satellite observation and reanalysis
- ✦ Model-to-model variation is large even with the same forcing data

Goals of ADMIP

- ✦ Evaluating & improving land surface models (LSM, energy and water) & terrestrial ecosystem model (TEM, carbon) through **offline model intercomparison** using data obtained **at Asian dryland**
- ✦ but not to improve a model/models specific to Asian dryland
- ✦ Better reproduction/prediction of landsurface state at Asian dryland using improved models
- ✦ and capacity building for them.

2010/7/11 ADMIP Kickoff, Beijing

Target Phenomena and Scientific Questions

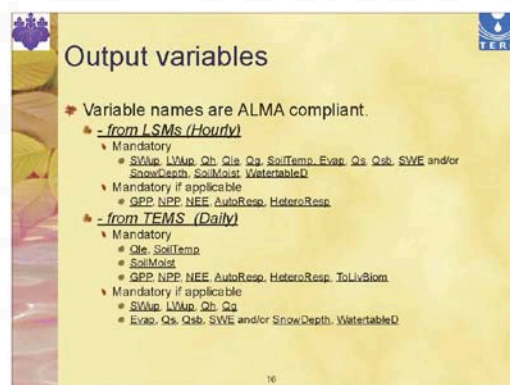
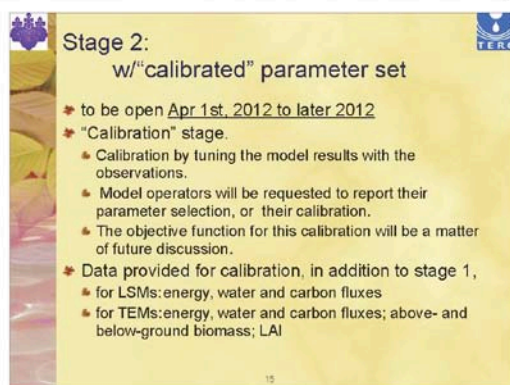
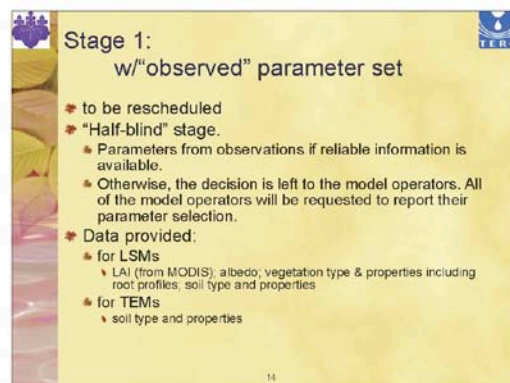
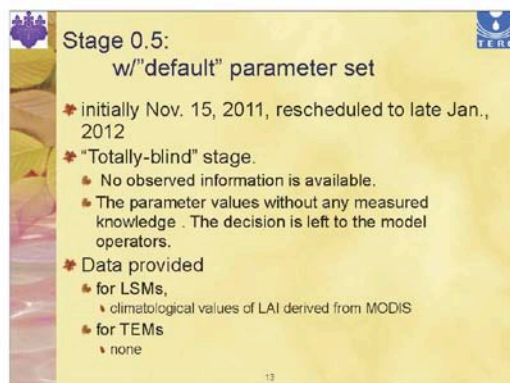
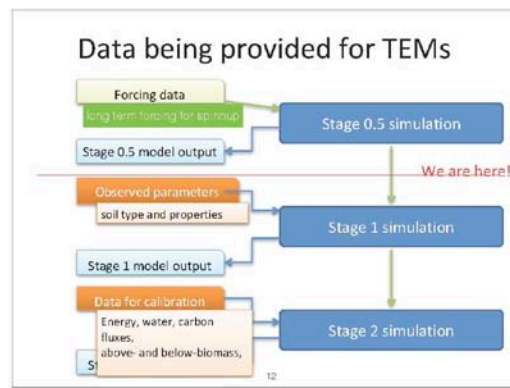
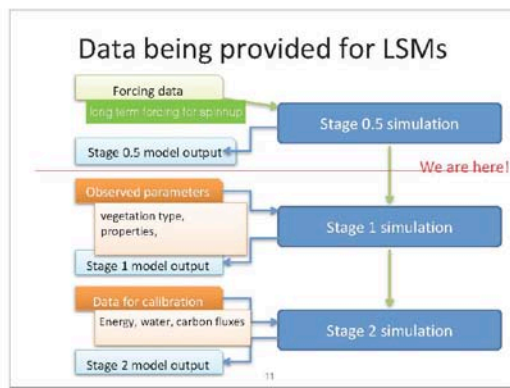
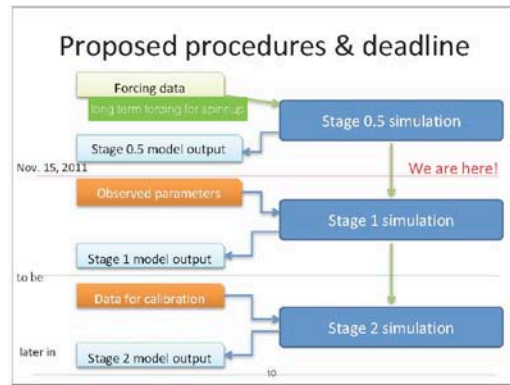
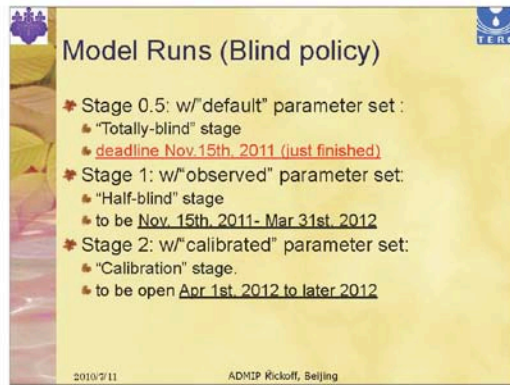
- 1. Target Phenomena**
 - ✦ Energy, water and carbon exchange at the land surface under water-limited environment, and/or under the control of vegetation growth, at the temporal scales of 3-hourly, daily, and monthly.
- 2. Scientific Questions**
 - ✦ What is the ability of models to reproduce energy, water, and carbon exchange at dryland surface (reproducibility)?
 - ✦ Do current ecosystem models, developed mainly for forest, reproduce energy, water and carbon exchange?
 - ✦ Does the **current complexity** of models effectively simulate land surface processes? (complexity issues, also related to benchmarking)
 - ✦ Does the **multi-model ensemble** of LSMs outperform single model output? (ensemble)


010/7/11 ADMIP Kickoff, Beijing

Target sites, periods

- ✦ Tongyu (contact: Ailikun)
 - ✦ period: 2003-2004
- ✦ KBU (contact: J. Asanuma)
 - ✦ period: 2003-2009
- ✦ Pingliang (contact: Jun WEN)
 - ✦ period: 2004-2008
 - ✦ candidate in the future stage







Data working group

- a. LAI and albedo from MODIS**
 - CAREERI group (Dr. J. Wen) & Fukushima U (Mr. Kondo)
 - LAI (at 500 m res.) and albedo (at 5km res.) from MODIS products.
 - 1) climatological (LAI) from 10-years data for stage 0.5.
 - 2) time series at every 8 days.
 - At KBU, LAI will be calibrated with in situ observations.
- b. Soil properties**
 - Dr. Yang K. (Tongyu) and Dr. Asanuma (KBU)
 - Parameters derived: b-parameter
- c. Grazing history**
 - Dr. Aili (Tongyu) and Dr. Asanuma (KBU)
 - Intensity, either of heavy, moderate or none, and its seasonal change in a descriptive form should be specified.
- d. Root profile (depth) from the literature (at KBU)**
 - Mr. Kondo.

17
17

Discussions 1

J. Asanuma

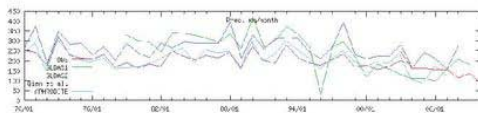
1

Things to discuss

- Rescheduled deadlines of the projects
- Outline of a prospective key paper
- Promotion of intercomparison study
- Funding Issues
- Format and convention issues, revisited

2

Consistency of the long-term spin up forcing data set with observation



red: forcing data (obs.)
purple: forcing for spin-up (w/NCAR CLM)
green: GLDAS1
dark blue: GLDAS2
light blue: APHRODITE

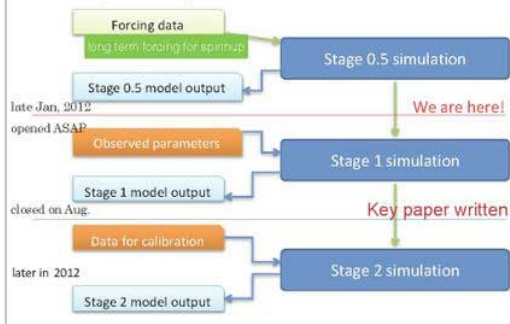
3

RESCHEDULED DEADLINES

4



Proposed procedures & deadline



need Gap-filling works

Gapfil

done with KBU, except for CO₂ flux

with Reichstein's statistical method on Web.

will be done with tongyu, need especially with CO₂ flux

same as KBU

due on the end of AUG.

OUTLINE OF A KEY PAPER

Contents

1. Introduction: what and how is dryland?
2. ADMIP: the first MIP initiative in Asia
3. General intercomparison results, W, E & C results
4. More focus on water limited processes
5. Conclusions

Intended for JGR, J. Hydromet., WRR, J. Arid Land, etc..

1. Introduction: what and how is dryland?

"While one third of the earth's land surface is classified as 'dry', our knowledge of and land is far less than this amount. bla bla bla ..."

Why dryland research is important?

major dust sources, vulnerability in society

What is the key phenomena in water(W), energy(E) & carbon(C) exchange at landsurface?

- Water scarcity is the major controlling factor for WEC exchange
 - W stress limits E, W&C processes, and then C processes (plant growth), in turn, controls W & E processes, at a longer time scale
 - Surges of W supply (precip.) enable plants to accelerate its growth, and then enhances W & E processes
 - Biotic control vs Physical control
- Highly-variable drivers in W (precip.) generates complicated interaction between W-E-C process
- not too general

Introduction

The purpose of the paper is

to evaluate the current skill of LSMs and TEMs to predict W, E & C exchange process at water-controlled land surface

to key parameters and processes to further improvement of LSMs and TEMs

2. ADMIP: the first landsurface MIP at Asian Dryland

a) MIP reviews

PILPS, ALMIP etc

b) Motives of ADMIP

less observed, less tested with and poorly performed by models

b) Objectives

To intensively test LSMs and TEMs at Asian drylands to identify key phenomena to model performances

c) Models: bunch of TEMs and LSMs

d) Target sites and observed data

e) Experimental design

Stage 0.5 (no info.) and Stage 1 (w/ obs. pars.)



3. General intercomparisons

- a) monthly GPP, AR, HR, ET, Bo
- b) effects of spin-up (stage 0.5 vs 0.5rerun)
- c) effects of added information, i.e. observed pars. (stage 0.5 vs 1)
- d) intersites comparison

13

4. Water-limited processes (optional)

More focused intercomparison w.r.t. intra seasonal variations

Are we able to reproduce W, E, and C interaction during water-stressed period, and the recovering period after it.

less P > less ET and less GPP

leads to less LAI after few months, then less ET

14

5. Conclusions

The current models can do what, and cannot do what.

Spin-up has a effect on what to what extent

Observed information improves what and does not improve what.

Model results are sensitive to a parameter/parameters X, which needs to be cautiously given to the models

15

PROMOTING INTERCOMPARISON ANALYSES

2018/7/11

ADMIP Kick-off, Beijing

16

4) Data sharing policy

- First, all forcing data and model outputs will be shared by the **participants** of ADMIP.
- A key paper will be published with all participants as well as the data provider as the authors.
- At 1 year after the publication of the key paper, forcing and input data that were used to run the models and the model outputs will be available **to the public**.

17

toward more elaborated works in model intercomparisons

- model outputs will be shared by the participants after a key paper is published in a near future
- need to promote few scientists into intercomparison analyses
- need to avoid conflict in between scientists
- there are ideas of organized specific experiments
 - e.g. with modified codes, or modified forcings
-

18

organizing a specific experiment partially discussed in Dec. meeting

- Examples
 - Theme specific
 - drought experiment (with actual or virtual drought)
 - Scheme-specific
 - grazing scheme
 - soil water
 - May need revised model codes
- In parallel with or after the main stream schedule
- Analyses of the multi-model outputs may required for the planning

19

A proposal

- Model outputs will be shared among the participants of the project
- Proposals on analysis of the multi-model outputs will be made
 - during the tomorrow morning session
 - on mailing list during a limited time period for those absent in this meeting

20



FUNDING ISSUES

21

Funding Issues: possible sources 1

- APN: extension or new project
 - JA will inquire the sci. director about possibilities
- Bilateral funds (supplemented by other funds)
 - JSPS-CAS(国际合作部): (1,500K JPY/year from JSPS, 1 to 2yrs 9mo)
 - mainly on meetings (travel, accommodation etc)
 - proposal due in mid. Sep (JA and Aili)
 - JSPS (Core-to-core program, "Asia Africa Science Platform")
 - 8,000K JPY/year
 - probably due in Sep. (JA and Aili)
 - JST-MOST (???)

22

Funding Issues: possible sources 2

- Tri-lateral funds (JST, JSPS)
 - Japan, China, Korea: need a strong group in Korea
- Belmont forum

23

Funding Issues: possible sources 3: Belmont forum

- What is Belmont forum?

The Belmont Forum is a high level group of the world's major and emerging funders of global environmental change research and international science councils. It aims to accelerate delivery of the international environmental research most urgently needed to remove critical barriers to sustainability by aligning and mobilising international resources. The aims of the Belmont Forum are detailed in a White Paper, and encapsulated as the Belmont Challenge. To deliver knowledge needed for action to mitigate and adapt to detrimental environmental change and extreme hazardous events.
- the G8 Heads of Research Councils (HORCs)
 - Australia, Brazil, Canada, France, Germany, India*, Japan, Russia, South Africa, the United Kingdom, and the United States of America.

24

Funding Issues: possible sources 3: Belmont forum

- Calling proposals for Core Research Activities(CRA): multilateral
 - Coastal vulnerability
 - Fresh water security
- 1. Identification and characterization of the interactions between natural processes (physical and biological, including ecological processes) and human (including cultural, social, economic, technological, adaptation, transfer and water re-use) practices that govern water budgeting in selected regions. This will include establishing baseline data on water use, storage and distribution, including climate events and global impact, their impacts, and determining which are most important in governing the vulnerability of socio-economic and environmental systems to water scarcity.
- 2. Development of approaches that support the evolution of resilient communities/regions through improved seasonal (months to multi-year) forecasting of droughts, taking into account natural (hydro-meteorological) and socio-economic drivers identified in the above work package. Research should clearly couple the complex system science of water stress at multiple scales to the structure and protocols for decision making. Development of these approaches is expected to involve both model-based and process-based research. That will use of existing observations and existing modelling approaches, and where possible identify key leading local observations, it will explore utilization of forerunner science, and will consider determining how individuals, communities, businesses and governments etc., as well, their habits and practices on the scale of improved research.

25

FORMAT AND CONVENTION ISSUE, REVISED

26

format and convention issues

- submitted netCDF's have variable format w.r.t
 - dimensional structures
 - variable names (those not specified in ALMA)
 - etc
- Stick to netCDF?
 - need more strict specification?
 - or surrender to text format?
- Conclusion: will use CSV.
 - need a specification of header comments and a sample for it

27

Analysis proposals

Yang Kun:
need stomatal resistance

28



Final discussion

J. Asanuma

1

Soil freezing/thawing process and its control to vegetation phenology (By Kun Yang and Qian Li)

- Observed soil moisture and temperature relationship
- Simulated soil moisture and temperature in models
- Explore Low temperature and physiology relationship
- Soil freezing/thawing control to vegetation phenology
- Add a Tibet site (Nam Co grassland)

2

Possible topics

- How soil moisture control ET and CO₂ (By Bill Parton)

3

Dennis,

Carbon and water dynamics: Soil moisture-Land-atmosphere flux relationship

Compute WUE (net carbon uptake relative ET) relationships relative to soil moisture levels

Soil moisture relative to NEE (GPP and Respiration) and LE

Example analyses of observations (model comparisons can will be made in a comparable manner)

4

Discussion

data submission

- all of the submission from now on will be in CSV
- Spin-up data set for Precipitation, consistent with forcing?
 - need to check PDF, timeseries
 - Fukushima G will do, also for Tongyu
- Data opening for Stage 1
 - Tongyu: need albedo, check with J. Wen (Yorozu will send mail)
 - KBU: need soil properties, and J. A. with help of Bynabahu will do
 - Tongyu will be open, then KBU will follow

9

0.5 and 0.5 rerun submission

BAIM

CLM

CoLM(Aili wil contact)

MATSIRO

NCASA

SCASA

SiBUC

NOAH-MP(will be submitted in 1 month,)

CSV, will be used

SM

BBGC

SEIB

VISIT

HAL(J.A. will)

DAYCENT

SSiB2

expected

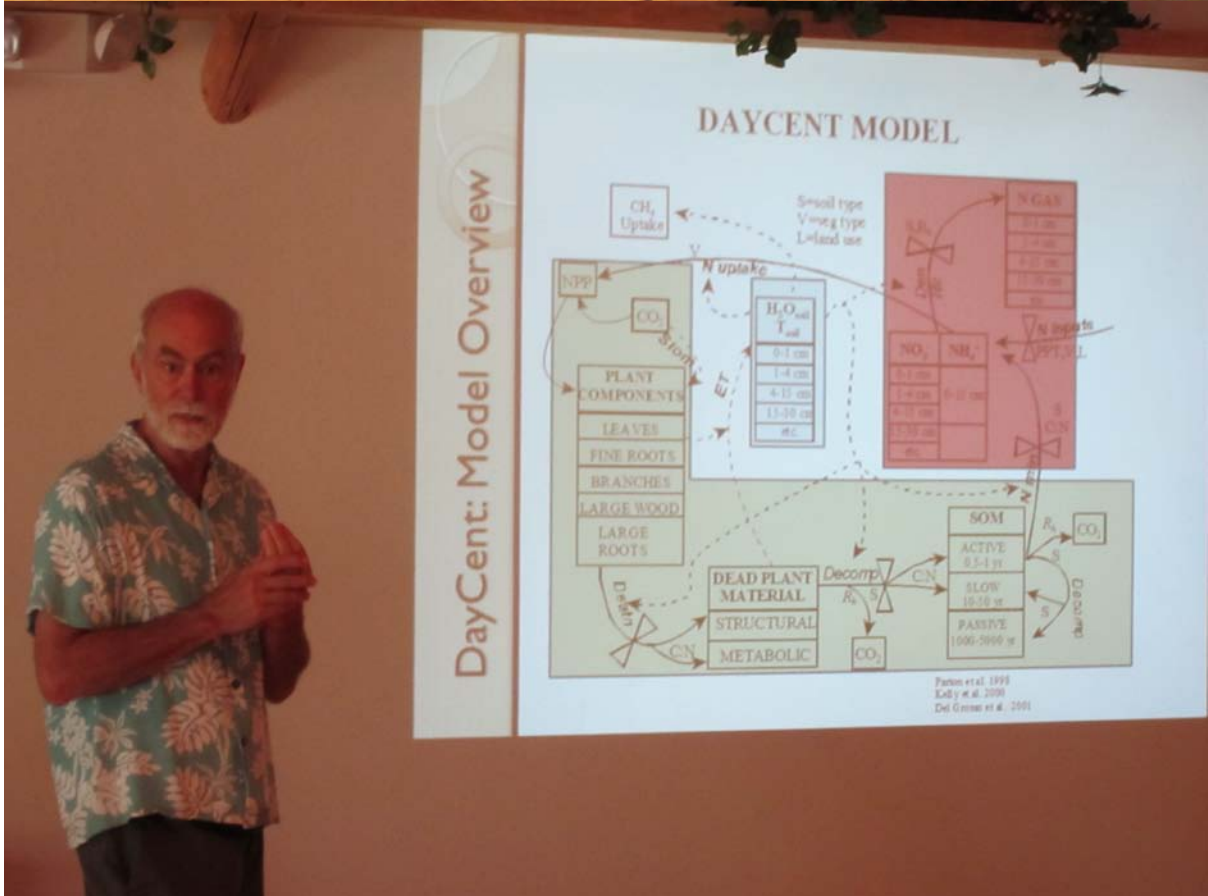
Jinkyu

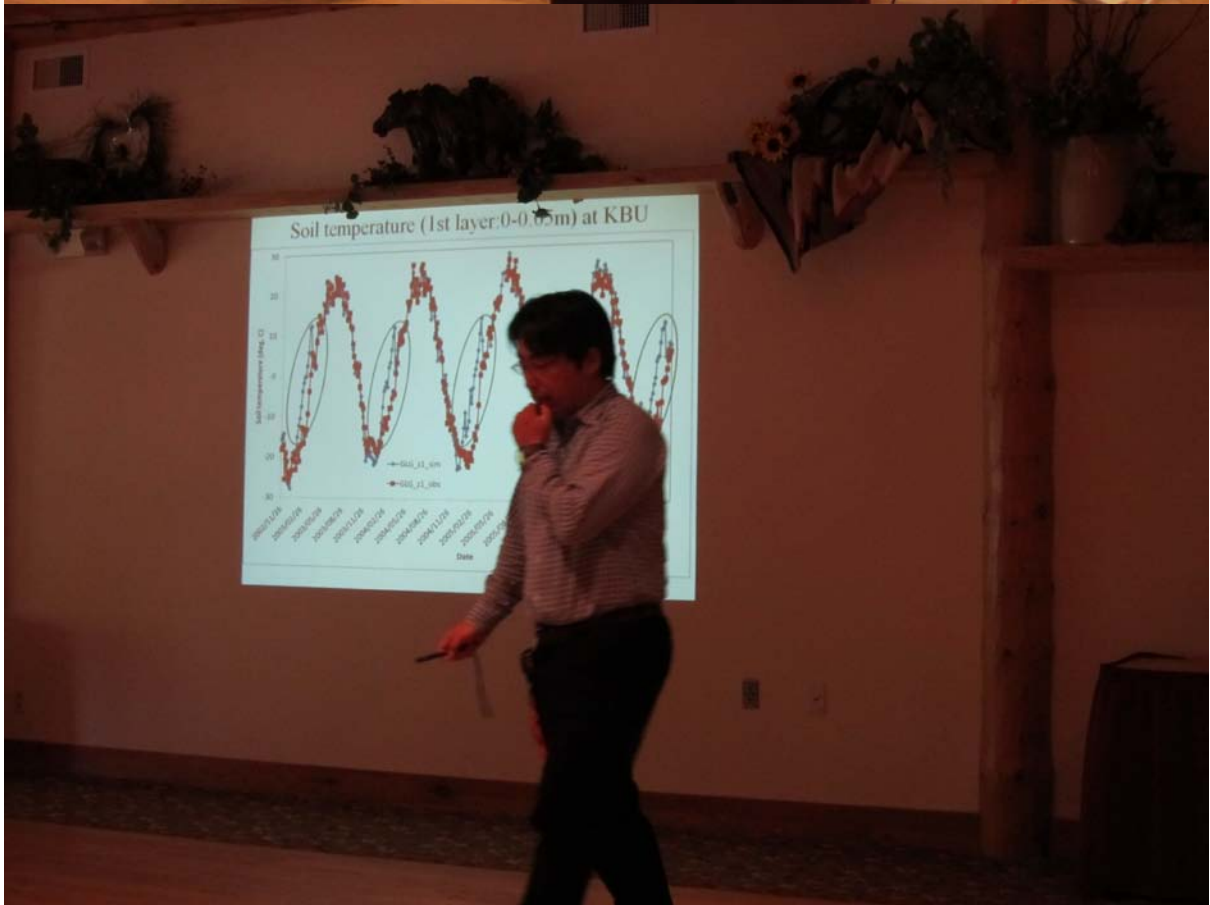
Hong(JULES, J.A>

10



5.3 Pictures









**The 2nd Summer School on Land Surface Observing,
Modeling and Data Assimilation**

13-16 July 2010, Beijing Normal University, Beijing, China

The 2nd Summer School on Land Surface Observing, Modeling and Data Assimilation

13-16 July 2010, Beijing Normal University, Beijing, China

Program Schedule

Registration	
8:30 -20:00, July 12 (Monday), the main hall of the Science and Technology Building	
8:00-12:00, July 13, RM502, Zeng Xianzi Building (Teaching Building 9), Registration	
Opening Ceremony, Chaired by Prof. S. Liang	
8:30 – 9:30am, July 13 (Tuesday), RM502, Teaching Building 9	
Welcome remarks by Prof. Jianping Ge, Executive vice president of Beijing Normal University	
Guest speeches	
Prof. Guoping Zhang, Director of National Center for Remote Sensing, China	
Dr. Ramasamy Jayakumar, UNESCO Office in Beijing	
Dr. Ailikun, Monsoon Asia Integrated Regional Study (MAIRS)	
Prof. Wenjie Dong, Executive Associate Dean, College of Global Change and Earth System Sciences, Beijing Normal University	
Logistics of the Summer School, Ms. Bo Jiang, Beijing Normal University	
9:30 – 10:00 Group picture taking	
PART I: Data Sources, Chaired by Dr. J. Shi, July 13, 2010 (Tuesday):	
10:00 - 11:00	Microwave Modeling and Parameterization for Data Assimilation, Dr. Jiancheng Shi
11:00 - 11:10	Break
11:10 - 12:10	Atmosphere Forcing Data in China Region Based on FY2 Satellite, Dr. Chunxiang Shi
12:10 -14:00	Lunch
14:00 - 15:20	Second Generation Polar-Orbiting Meteorological Satellites of China: Fengyun 3 and its application on global monitoring, Prof. Peng Zhang

15:20 - 15:30	Break
15:30 - 16:50	NASA Satellite Remote Sensing and Assimilation Data and Services: A Tutorial of Data Accessing, Dr. Suhung Shen
16:50 - 17:00	Break
17:00-18:20	Overview of Satellite Remote Sensing Products for Land Data Assimilation, Prof. S. Liang
PART II: Land Surface Modeling, Chaired by Dr. Ailikun, July 14, 2010 (Wednesday):	
8:30 - 9:50	Modeling Land and Atmosphere Interactions, Prof. Yongkang Xue
9:50 - 10:00	Break
10:00 - 11:20	Land Surface Modeling in Arid and Semi-arid Regions, Dr. Kun Yang
11:20 - 14:00	Lunch
14:00 - 15:20	Integration of MODIS Products and a Crop Simulation Model for Crop Yield Estimation, Prof. Hongliang Fang
15:20 - 15:30	Break
15:30 - 16:50	Uncertainty Quantification of Large Complex Dynamical Systems, Prof. Qingyun Duan
16:50 - 17:00	Break
17:00-18:20	Parameter Estimation and Error Analysis of Ecosystem Models: Component Analysis of Ecosystem Models, Dr. Dennis Ojima and Prof. William Parton
PART III: Data Assimilation Methods, Chaired by Dr. Xin Li, July 15, 2010 (Thursday):	
8:30 - 9:50	Assimilating Remote Sensing Data into Land Surface Models: Theory, Method and a Software Tool I, Dr. Xin Li
9:50 - 10:00	Break
10:00 - 11:20	Cost-function Based Methods and Particle Filter Method: Theory and Application in Soil Moisture Estimation, Dr. Jun Qin
11:20 - 14:00	Lunch
14:00 - 15:20	Review of Parameterization and Parameter Estimation for Hydrologic models, Prof. Soroosh Sorooshian
15:20-15:30	Break
15:30 - 16:50	Forecast Error Covariance Inflation For Data Assimilation, Prof. Xiaogu Zheng
16:50 - 17:00	Break

17:00 - 18:20	Assimilating Remote Sensing Data into Land Surface Models: Theory, Method and a Software Tool II , Dr. Xin Li and Liangxu Wang
PART IV: Data assimilation applications, Chaired by Dr. Xin Li or Dr. Xianhong Xie, July 16, 2010 (Friday):	
8:30 - 9:40	Towards Operational Soil Moisture Monitoring, Prof. Jeffrey Walker
9:40 - 9:50	Break
9:50 - 11:00	The Application of Remotely Sensed Land Surface Properties in Regional Climate Model, Dr. Lixin Lu
11:00 - 11:10	Break
11:10-12:20	Land Surface Modeling and Its Applications, Prof. Xu Liang
12:20 - 14:00	Lunch
14:00 - 15:10	North American Land Data Assimilation System, Dr. Youlong Xia
15:20 - 15:30	Break
15:30 - 16:30	Distributed Hydrological Modeling and Data Assimilation, Dr. Xianhong Xie
16:30-16:40	Break
Special session on UNESCO-IHP, Chaired by Dr. Xin Li	
16:40 - 17:40	Role of UNESCO-IHP (International Hydrological Programme) in Global Hydrological System and Modelling and IHP Programme in East Asia, Dr. Ramasamy Jayakumar and Dr. Anil Mishra
17:40 - 18:20	Climate Change Adaptation: Two Case Studies of IHP from Indian Sub Continent, Dr. Bhanu Neupane
18:20 - 19:00	Returning earphones and University Temporary card Distribution of CD containing all lecture notes, slides, and pictures

b

CONTENT

PART I: Data Sources

- 1 **Shunlin Liang:** [Liang, S., J. Qin (2008), Chapter 12: Data assimilation methods for land surface variable estimation, *Advances in Land Remote Sensing: System, Modeling, Inversion and Application*, Springer]**2 – 28**
- 2 **Chunxiang Shi:** [Shi, C., Z. Xie, et al. (2010), China land soil moisture EnKF data assimilation based on satellite remote sensing data]**29 – 40**
- 3 **Peng Zhang:** [Zhang, P., J. Yang, et al. (2009) General introduction on payloads, ground segment and data application of Fengyun 3A, *Front. Earth Sci. China*] **41 – 46**
- 4 **Suhung Shen:** [Shen, S., G. Leptoukh and H. Fang (2010), NASA satellite and model land data services: data access tutorial]**47 – 60**
- 5 **Jiancheng Shi:** [Jiang, L., J. Shi, et al. (2007), A parameterized multiple-scattering model for microwave emission from dry snow, *Remote Sensing of Environment*]**61 – 70**

PART II: Land Surface Modeling and Uncertainties

- 6 **YongKang Xue:** [Xue, Y., H. Juang, et al. (2004), Role of land surface processes in monsoon development: East Asia and West Africa, *Journal of Geophysical Research*]**72 – 95**
- 7 **Kun Yang:** [Chen, Y., K. Yang, et al. (2010), Improving the Noah land surface model in arid regions with an appropriate parameterization of the thermal roughness length]**96 – 107**
- 8 **Qingyun Duan:** [Ajami, N., Q. Duan, and S. Sorooshian (2007), An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction, *Water Resources Research*]**108 – 126**
- 9 **Dennis Ojima:** [Raupach, M., P. Rayner, et al., (2005), Model-data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications, *Global Change Biology*]**127 – 146**

PART III: Data Assimilation Methods

- 10 **Xin Li:** [Talagrand, O. (1997), Assimilation of observations, an introduction, *Journal of the Meteorological Society of Japan*]**148 – 166**
- 11 **Jun Qin:** [Qin, J., S. Liang, et al. (2009), Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal, *Journal of Geophysical Research*]**167 – 179**
- 12 **Soroosh Sorooshian:** [Liu, Y., H. Gupta (2007), Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, *Water Resources Research*] **180 – 197**
- 13 **Xiaogu Zheng:** [Zheng, X. (2009), An adaptive estimation of forecast error covariance parameters for Kalman filtering data assimilation, *Advances in Atmospheric Sciences*].....**198 – 204**

PART IV: Applications of Data Assimilation

- 14 **Jeffrey Walker:** [Draper, C., J. Mahfouf, and J. Walker (2009), An EKF assimilation of AMSR-E soil moisture into the ISBA land surface scheme, *Journal of Geophysical Research*]**206 – 218**
- 15 **Hongliang Fang:** [Fang, H., S. Liang, et al. (2008), Corn-yield estimation through assimilation of remotely sensed data into the CSM-CERES-Maize model, *International Journal of Remote Sensing*]**219 – 240**
- 16 **Youlong Xia:** [Mitchell, K., D. Lohmann, et al. (2004), The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *Journal of Geophysical Research*]**241 – 272**
- 17 **Xianhong Xie:** [Xie, X., D. Zhang (2010), Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter, *Advances In Water Resources*] **273 – 285**

Special Session on UNESCO-IHP

- 18 **Ramasamy Jayakumar:** [Jayakumar, R., X. Duan, et al. (2009), The role of United Nations Educational, Scientific and Cultural Organization–International Hydrological Programme in sustainable water resources management in East Asian countries, *Journal of Geographical Sciences*] **287 – 300**

863 Program—Generation & Application of Global Products of Essential Land Variables 301 – 305

List of Invited Experts 306 – 306

PART I: DATA SOURCES

This part addresses scientific problems about data sources, focuses on the Satellite Remote Sensing Products, and Interprets how to couple these data with land surface models based on data assimilation frameworks. Five invited experts contributed their lecture notes or recommended some helpful reference papers.

Chapter 12

Data Assimilation Methods for Land Surface Variable Estimation

Shunlin Liang and Jun Qin

Abstract Estimating land surface variables from remote sensing data is an ill-posed problem. Integration of observations from multiple satellite sensors with different spectral, spatial, temporal and angular signatures is now an important research frontier. Data assimilation (DA), integrating not only remotely sensed data products, but also other measurements and land dynamic models, is an advanced set of techniques for innovative parameter estimation. After a brief introduction, we describe the basic principles of DA, and then provide in-depth discussions of some relevant issues while using DA. The latest applications of DA for estimation of soil moisture, energy balance, carbon cycle and agricultural productivity are summarized.

12.1 Introduction

Despite the abundance and variety of remote sensing measurements, land surface characterization from satellite observations is still very challenging. There are multiple sources of surface information, such as remote sensing data and derived products, in situ measurements, and land surface model outputs. Innovative techniques are needed to merge these information sources and optimize the use of satellite measurements for robust surface products and greater predictability. Data assimilation (DA) is a mathematical approach that enables use of all available information within a given time window to estimate various unknowns. The information that can be incorporated includes observational data, existing *a priori* information, and, very importantly, a dynamic model that describes our system and encapsulates current theoretical understanding. The model brings consistency to the observational

Shunlin Liang
Department of Geography, University of Maryland, College Park, USA
sliang@umd.edu

Jun Qin
Institute for Geographical Science and Natural Resource Research, Beijing, China

S. Liang (ed.), *Advances in Land Remote Sensing*, 313–339.
© Springer Science + Business Media B.V., 2008

313

data, and interpolates or extrapolates data into data-devoid regions in space and time. The observational data, representing the actual state of the system, corrects the trajectory of the imperfect model through adjusting model parameters. DA is also called model–data synthesis or data–model fusion in different disciplines.

The benefits of DA for maximizing the scientific and economic value of remote sensing data is summarized as follows (O'Neill et al., 2004; Raupach et al., 2005):

1. *Forecasting and error tracking.* By regularly comparing forecasts with observations, extremely valuable error statistics can be built up, which in turn can be used to improve the quality of the observations (e.g., by revealing biases in instrument calibration) as well as the quality of the models. It serves as a model testing and data quality control procedure.
2. *Combining multiple data sources.* Different observing systems (both in situ and remote sensing data) have varying virtues and deficiencies. Such variety can be preferentially exploited or contrasted to optimize the value of the resulting data set.
3. *Interpolating spatially and temporally sparse observations.* The model provides a way to propagate information in a consistent manner in space and time from data-rich regions to data-poor regions. This capability is vital to successfully utilize satellite observations, which due to limited and sequential sampling provide only an incomplete picture of the Earth. DA fills in “missing pieces” to achieve a full global picture.
4. *Inferring, from available observations, quantities not directly observable.* Through relationships expressed in the model's governing equations, measured parameters convey knowledge of those that are inadequately measured or completely lacking. For example, soil moisture vertical profile can be inferred from the surface skin temperature or surface top-layer soil moisture content.
5. *Forecasting.* Predicting forward in time on the basis of past and current observations.
6. *Designing observing systems.* Decisions to deploy new satellite-borne instruments require critical assessment of the incremental value or benefit of the data to be acquired by the new sensors. With careful design, DA experiments provide an objective, quantitative way to contribute to such assessment. In addition, DA can optimize the sampling pattern from an observing system, and can target observations to capture features of concern, such as a rapidly developing storm.

DA methods exploit data streams not only to validate model outputs, or directly to infer fluxes, but principally to constrain internal model parameters to optimize values (i.e., parameter estimation). Different data sets constrain different components of a model which is able to assimilate data across a range of space and time scales. Another distinctive characteristic of DA is that uncertainties associated with the observation, techniques, processing, representation, and accuracy are as important in determining the final outcome as the measured values themselves. Thus, uncertainty estimates, for both measurements and model parameters, take on even greater importance (Canadell et al., 2004).

The meteorological and oceanographic communities have been at the forefront in developing and using DA methods. In recent years, meteorologists and

oceanographers have tended to view DA as a model state estimation problem. The land community has aggressively endeavored to catch up and apply DA methods in recent years, and some example applications are shown in Section 12.4.

The remainder of this chapter presents a brief introduction to the basic principles of land DA as this is a relatively new field. Following the introduction is a discussion of some critical issues pertinent to land DA. Section 12.4 introduces recent DA applications in several major disciplines.

12.2 Principles of Land DA

In this chapter, land DA is considered as essentially an estimation problem: that is to acquire optimal estimates of model state variables of the land surface and its parameters given a set of remote sensing products, a land surface process model, and any available *a priori* information. Various land DA schemes have different characteristics, but they may have the following common features: (1) a forward land dynamic model that describes the time evolution of state variables such as surface temperature, soil moisture and carbon stocks; (2) an observation model that relates the model estimates of state variables to satellite observations and vice versa; (3) an objective function that combines model estimates and observations along with any associated prior information and error structure; (4) an optimization scheme that adjusts forward model parameters or state variables to minimize the discrepancy between model estimates and satellite observations; and (5) error matrices that specify the uncertainty of the observations, model and any background information (these are usually included in the objective function).

12.2.1 Dynamic Model

Land surface process models are often structured as a discrete-time nonlinear state-space model with additive noise as:

$$x_{t+1} = f(x_t, u_t, \theta) + w_t \quad (12.1)$$

where x denotes the model state vector (e.g., soil moisture), u the external forcing data (e.g., meteorological data), θ the model parameter vector (e.g., soil texture), w_t the model noise, and $f(\cdot)$ the model operator mapping the previous state x_t to the next state x_{t+1} . The differences among various land surface models are reflected in different specifications of x , u and $f(\cdot)$.

12.2.2 Measurement Process

DA methods assimilate the measured values into a dynamic model, so the measured quantities have to be linked with model variables. The measurement process

is a mathematical model that relates the model state vector (x_t) to the observation vector (y_t):

$$y_t = h(x_t, \theta) + e_t \quad (12.2)$$

where e_t the observation noise, $h(\cdot)$ the observation operator, and θ is the parameter set of observation operator. If remote sensing data products are used as the observations, the remote sensing reflectance, or emittance model of land surfaces, and the coupled land and atmosphere system are the observation operator and θ is its parameters.

12.2.3 Objective Function

The underlying principle of the DA method is to estimate the parameters and variables of the dynamic models by minimizing the differences between the predictions of the model and the assimilated data products. The difference is often characterized as the objective (or cost) function, and the common measure is the least squares.

DA methods usually attempt to minimize the objective function J

$$J(x) = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b) + \frac{1}{2}(H(x) - y)^T R^{-1}(H(x) - y) = J_b + J_o \quad (12.3)$$

where y is the observation vector, x is the extended model state variables, x_b is the background field (or first guess), H is the model operator, and R is the observation-error covariance matrix and B is background-error covariance matrix.

The DA problem now becomes: vary x to minimize $J(x)$, subject to the constraint that the state variables must satisfy the dynamic model. The value of x at the minimum is the a posteriori estimate of x , including information from the observations as well as the background. In Eq. (12.3), the first term J_b is to force the optimal parameters as closely as possible to background fields, and the second term J_o is to adjust parameters so that model outputs will be as close to the observations as possible. Specifying R and B depends on the relative accuracy of background information and remote sensing data products. In extreme cases, if the errors of the “first-guess” values are extremely large, the final estimates will be decided from the fitting of the observations and will be close to the “first-guess” values.

12.2.4 Assimilation Algorithm

Since land DA is considered to be an estimation problem, an assimilation or estimation algorithm is needed to estimate model parameters or states by assimilating observations into the dynamic model. This process provides three kinds of output: optimal estimates for the model properties to be adjusted, uncertainty statements about these estimates, and an assessment of how well the model fits the data, given

the data uncertainties. There are two types of assimilation algorithms popular in current research and applications, namely cost function-based methods and sequential methods. These two methods have many different forms and will be discussed in Section 12.3.5.

12.2.5 Error Model

Both dynamic models and measurements are not perfect and contain uncertainties to some degree. These uncertainties or errors have to be characterized in the DA system. Consequently statistical properties of w_t in Eq. (12.1) and e_t in Eq. (12.2) or R and B in Eq. (12.3) need to be estimated. The error terms in the dynamic and observation models can, in principle, be quite general in form, including biases, drifts, temporal correlations, extreme outliers and so on. Many extant methods assume these error terms (e.g., w_t and e_t) have Gaussian distributions with zero mean and no temporal correlation. Detailed discussions of observational errors are given in Section 12.3.6.

12.3 Critical Issues in Land DA

12.3.1 Dynamic Models

Land DA uses land surface process models that describe the exchanges of momentum, energy and mass between soil, vegetation and atmosphere. These models, often called soil–vegetation–atmosphere transfer (SVAT) models, have similar structures, whether land surface, hydrological, ecological or crop growth models.

How to select an appropriate dynamic model is largely related to the objectives of the DA study. For the same objective, multiple models might be available. For example, if we aim to estimate land surface energy balance components such as latent and sensible heat fluxes by assimilating remotely sensed land surface temperature (LST) products, almost any land surface process models may be appropriate. However, some models are rather complicated since they may describe other processes such as those in the carbon cycle in addition to water and energy budgets and thus contain more model parameters. Often a good strategy is to use a relatively simple land surface model if possible.

A more general and important issue is the uncertainty of the model. All mathematical models are abstractions and approximations of a complex physical environment. When dealing with ecosystem models, for example, keep in mind that their designs are not completely derived from rigorous natural laws and hence the mathematical descriptions of the biological processes are not universal. Thus, we are uncertain not only about values of the numerous model parameters, but also about the model parameterizations and errors.

DA aims to incorporate measured observations into a dynamic system model to produce accurate estimates of the system's current (and future) state variables. The neglected model uncertainties are interpreted, in the course of strongly constraining parameter estimation, as variations in model parameters and this may result either in unrealistic parameter estimates or in solutions deviating far from the data. In reality, the model does not exactly reproduce system behavior. Significant errors can arise due to a lack of resolution, and inaccuracies in physical parameters, boundary conditions and forcing terms.

Many studies have been reported to address these issues in terms of weak constraints (Liaqat et al., 2003a; Losa et al., 2004; Natvik et al., 2001; Qin et al., 2007b). As a rule, weakly constrained DA has been used for state estimation. However, if the model parameters are poorly known, state estimation with fixed model parameters can produce unacceptable results. Weakly constrained parameter estimation results from a combination of the adjoint method and the generalized inversion; consequently it can take into account errors inherent in the model and data, as well as find optimal values of the poorly known model parameters. In addition, weakly constrained DA makes it possible to derive valuable supplementary information about the model itself. This information is completely lost when the strong-constraint scheme is used.

Chepurin et al. (2005) recently summarized four solutions to correct biases due to model error in the meteorological community. The most straightforward approaches to handling time-mean bias involve computing climatologies, and then introducing correction terms into the equations of motion. The second type of approach is to correct for rapidly changing bias in a data-rich environment. It involves examination of the previous few updating cycles for a systematic trend, which is then corrected. A third class of approaches, useful in linear one-dimensional problems, involves pre-whitening the errors so that their frequency spectrum resembles white noise. A fourth class of approaches is referred to as "two-stage estimation." The two-stage estimation algorithm begins with the assumption that a reasonable estimate of the bias may be made prior to estimating the state of the system itself, thus allowing the estimation procedures for bias and state to be carried out successively.

12.3.2 Observation Vector

Which remotely sensed data products and other measurements are the most valuable for DA? There are two major schemes for assimilating remotely sensed data products into dynamic models (Liang, 2004). The first approach assimilates the derived high-level land products (e.g., leaf area index (LAI), fractional photosynthetically active radiation absorbed by green vegetation (FPAR), LST, gross primary production (GPP), biomass) from remote sensing observations (see Fig. 12.1). It is equivalent to simplification of measurement process through static remote sensing inversion. So the observation vector is directly the components of the model state and $h(\cdot)$ is simplified as a matrix H with non-diagonal elements equal to 0:

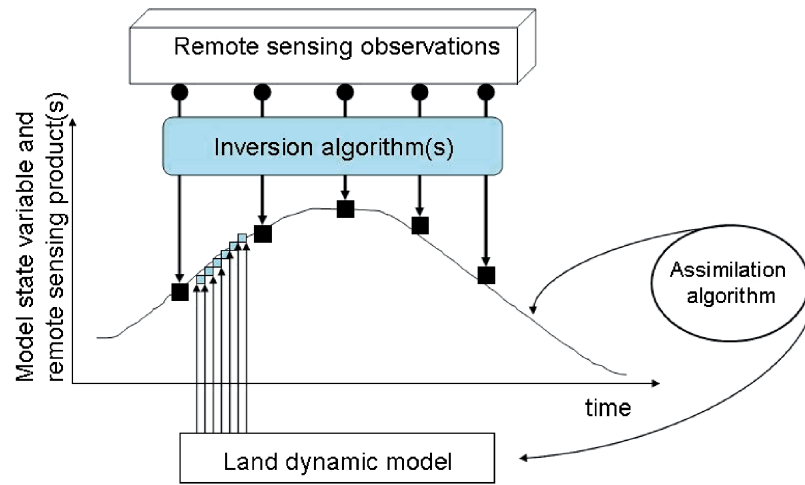


Fig. 12.1 Illustration of assimilating remotely sensed high-level products into a land dynamic model

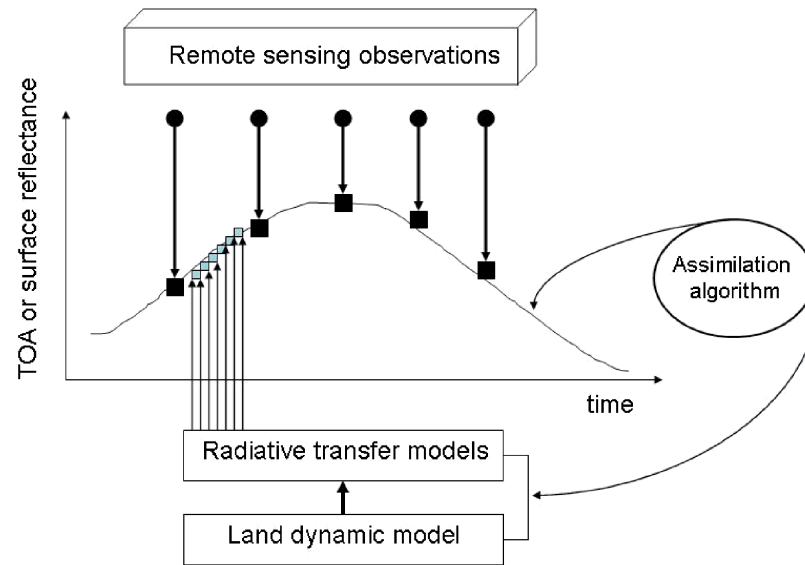


Fig. 12.2 Illustration of assimilating the direct remote sensing observations to the coupled radiative transfer and land surface dynamic model

$$y_t = Hx_t + e_t \quad (12.4)$$

The second approach assimilates the direct observations (radiance, reflectance, brightness temperature, or vegetation indices) (see Fig. 12.2). In this approach, $h(\cdot)$ is a radiative transfer (RT) model relating the state vector x to the remote sensing signal (e.g., radiance) or its simple transformation (e.g., reflectance).

Both approaches have strengths and weaknesses. If the high-level product retrieval is very accurate and its uncertainties well-characterized, then, use of such data is preferable. However, most inversion processes are ill-posed (see Chapters 7 and 8), and the resulting products may be poor. This is particularly so if their error characteristics cannot be specified accurately for DA (see Section 12.3.6). Another issue is that most current land products are not continuous both spatially and temporally due to cloud contamination or the inversion failure due to an inadequate number of good observations. Since many biogeophysical variables are “continuous” spatially and temporally, “gaps” need to be filled using various spatiotemporal modeling techniques (Fang et al., 2007a; Fang et al., 2007c; Julien et al., 2006; Moody et al., 2005).

On the other hand, there is no need for an inversion if direct observations are used, but the forward RT models must be coupled with land surface process models. With physically based RT models, an additional advantage of the second approach is that consistency can be maintained between the direct physical meaning of model parameters in both the RT and land surface process models. This is not necessarily so for high-level products that may be derived under a different set of assumptions (e.g., spatial distribution of foliage) to those assumed in the process model.

For using high-level products, not only those from optical and thermal observations (e.g., LAI/FPAR, albedo, LST) are considered but also those obtained from LIDAR (e.g., canopy structure parameters and above-ground biomass) and RADAR (above-ground biomass and moisture conditions) observations. Other products, such as forest biomass inventories, soil carbon survey, and Eddy covariance flux measurements, atmospheric CO₂ and tracers concentrations, nutrient fluxes and stocks (particularly N) can also be used for assimilation.

A highly related issue is the compatibility of remotely sensed products and variables in land surface process models. Although substantial efforts have been made from both remote sensing and modeling communities, the gaps still exist. For example, most land surface models separate broadband albedos into direct and diffuse components (Dai et al., 2003), but remote sensing albedo products are usually total broadband albedo (see Chapter 9). Most land surface models normally partition surface temperature into sunlit/shadow leaf and soil temperatures, but remote sensing skin temperature is an effective temperature for each pixel (see Chapters 4 and 10).

A particular advantage of DA schemes is that they can be used to evaluate the “advantage gained” (assessed through reduction in uncertainty) by the addition of new observations. This will be of great use in assessing the value of different types and sample availability of observations, as well as performing synthetic experiments to evaluate new types of observations (e.g., from proposed monitoring networks or space missions).

12.3.3 Target Variables

The target variables are the properties of the model to be adjusted or estimated in the optimization process. They include any model property considered to be sufficiently

uncertain as to benefit from constraint by the data. Model properties which can be target variables include: (1) model parameters (θ); (2) forcing variables, if there is substantial uncertainty about them; (3) initial conditions on the state variables; and (4) time-dependent components of the state vector. Land surface models are becoming more and more sophisticated with numerous parameters. Some parameters may be easily estimated, while others may not be very sensitive to the cost (objective) function. These target variables must be identified based on extensive sensitivity experiments.

Selecting the target variables depends on the land surface model used in the DA system, along with how many and which data products are assimilated. For example, Kaminski et al. (2002) optimized 24 parameters (light use efficiency and Q10 for heterotrophic respiration for each of 12 biomes) in a terrestrial biosphere model coupled with an atmospheric transport model using CO₂ data. Barrett (2002) optimized a set of parameters (turnover times, C allocation ratios, humification ratios, and light use efficiency) in a terrestrial C cycle model. Rayner et al. (2005) estimated 56 process parameters plus an initial condition through terrestrial carbon cycle DA system with a coupled ecosystem and atmospheric transport model. Williams et al. (2005) estimated nine unknown parameter constants in the C box model and the initial values of the five C pools using the ensemble Kalman filter.

For a given land surface model, a rigorous sensitivity study is absolutely required, which enables determination of the parameters/state variables that are sensitive to the assimilated data. The automatic differentiation (AD) tool called the Tangent linear and Adjoint Model Compiler (Giering and Kaminski, 1998) as well as its successor, Transformation of Algorithms in Fortran (Giering and Kaminski, 2002) used to generate the adjoint code of the process model may provide an effective way for sensitivity studies.

12.3.4 Multi-Objectives Optimization

Practical experience suggests that any single objective (cost) function, no matter how carefully chosen, is often inadequate to properly measure all of the characteristics of the assimilated data sets deemed to be important (Demarty et al., 2005; Vrugt et al., 2003). One strategy to circumvent this problem is to define several objective functions that measure different (complementary) aspects of the system behavior and to use multi-objective optimization to identify the set of non-dominated, efficient, or Pareto optimal solutions. Such an approach makes it more efficient to assimilate multiple data products simultaneously.

The choice of fitting criterion or “estimator” is crucial. The very popular weighted least-squares estimator is the simplest, and is amenable to very efficient mathematical algorithms. However it may be seriously inappropriate in certain circumstances. For example, the estimator of maximum likelihood may be a much more reasonable choice when errors in the data or dynamical information are non-Gaussian.

Generally speaking, the multi-objective problem can be stated as the following minimizing formulation:

$$\text{Min}\{J_1(\theta), J_2(\theta), \dots, J_m(\theta)\} \quad (12.5)$$

where J_i symbolizes a single objective function with $i = 1, \dots, m$ and $\theta = \{\theta_1, \theta_2, \dots, \theta_p\}$ a particular set of p parameters included in the feasible parameter space.

This issue is highly related to those discussed in Section 12.3.2. A single objective function corresponds to each high-level product or directly observed variable, thus, we can combine both high-level products and direct observations. Even for the same product, a single objective function can be a measure of errors (Gupta et al., 1998), such as root mean square error, mean absolute error, maximum absolute error, and so on.

The solution of the multi-objective formulation given in Eq. (12.5) does not lead to a unique solution, but to a set of solutions, generally named Pareto set or “behavioral” set (Gupta et al., 1998). There are many different algorithms to address this problem (Marler and Arora, 2004), such as the Multiobjective Shuffled Complex Evolution Metropolis (MOSCEM) algorithm that is capable of solving the multi-objective optimization problem for hydrologic models (Vrugt et al., 2003). MOSCEM is available to the public. Other assimilation algorithms will be discussed in the following Section 12.3.5.

12.3.5 Assimilation Algorithms

The target variables characterizing the land surface properties are estimated through the assimilation algorithms. On one hand, there are many different assimilation algorithms available in the literature from meteorological and oceanographic DA communities, such as three- or four-dimensional variational algorithms (see Section 12.3.5.1), and Kalman filters (Section 12.3.5.2). On the other hand, land DA has different characteristics, such as handling only dozens of unknowns, while the meteorological DA system typically manages 10^6 – 10^9 unknowns. Neural network and particle filtering techniques are briefly introduced in Sections 12.3.5.3 and 12.3.5.4.

Before moving on, the measurement sequence is defined as $y_{1:t} \equiv \{y_i\}_{i=1}^t$ and the state sequence as $x_{0:t} \equiv \{x_i\}_{i=0}^t$. With these definitions, the whole DA process as an estimation problem can be stated as:

$$\hat{x}_{0:t} = g(y_{1:t}) \quad (12.6)$$

where $\hat{x}_{0:t}$ represents estimated state sequence, and $g(\cdot)$ the estimator, which can be viewed as the assimilation algorithm. When a measurement sequence $y_{1:t}$ is inserted into the estimator, we obtain a realization of the estimator ($\hat{x}_{0:t}$).

12.3.5.1 Cost Function-Based Methods

This type of method is also called non-sequential or batch one. The most general form of cost function-based methods can be expressed as the following constrained optimization problem:

$$\begin{aligned} \min_{x_{0:t}} J &= \|x_0 - \bar{x}_0\|_{B^{-1}}^2 + \sum_{i=0}^{t-1} \|w_i\|_{Q_i^{-1}}^2 + \sum_{i=0}^t \|e_i\|_{R_i^{-1}}^2 \\ \text{s.t. } x_{i+1} &= f(x_i, u_i, \theta) + w_i, i = 0, \dots, t-1 \\ y_i &= h(x_i, u_i, \theta) + e_i \quad i = 1, \dots, t \end{aligned} \quad (12.7)$$

where the weighting matrices Q_i , R_i , and B can be regarded as covariances of the distribution functions $p(w_i)$, $p(e_i)$, and $p(x_0)$. If the parameter vector θ needs to be estimated simultaneously, a term $\|\theta - \bar{\theta}\|_{P_\theta^{-1}}^2$ can be added into Eq. (12.7) and

the control vector becomes $\begin{bmatrix} x_{0:t} \\ \theta \end{bmatrix}$.

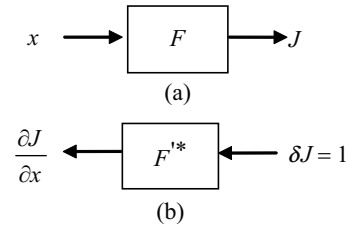
This problem is the so-called weak constraint DA approach (see Section 12.3.2) when the dynamic model is considered to be imperfect. The traditional strong constraint DA problem is just its special case in which the model is assumed to be perfect and thus the term $\sum_{i=0}^{t-1} \|w_i\|_{Q_i^{-1}}^2$ is removed in Eq. (12.7). The formulation statements can then be expressed as:

$$\begin{aligned} \min_{x_{0:t}} J &= \|x_0 - \bar{x}_0\|_{B^{-1}}^2 + \sum_{i=0}^t \|e_i\|_{R_i^{-1}}^2 \\ \text{s.t. } x_{i+1} &= f(x_i, u_i, \theta) \quad i = 0, \dots, t-1 \\ y_i &= h(x_i, u_i, \theta) + e_i \quad i = 1, \dots, t \end{aligned} \quad (12.8)$$

In the cost function-based problems, all data are treated simultaneously and the minimization problem is solved only once, as presented in Eq. (12.7) or (12.8). The problem becomes the usual optimization problem. It is computationally expensive when the number of unknowns is large. There are many non-sequential algorithms depending on whether the derivative information is used, but the key issue is to incorporate an effective global searching algorithm. The typical algorithms include the shuffled complex evolutionary (SCE) method (Duan et al., 1993; Duan et al., 1992; Duan et al., 1994), a very fast simulated annealing (SA) algorithm (Ingber, 1989; Li et al., 2004), the differential evolutionary (DE) method (Storn and Price, 1997; Storn and Price, 1996), and the genetic algorithm (GA) (Goldberg, 1989). Codes for these methods are available to the public. The common weakness of these methods is their slow computational speed.

In order to solve both weak and strong constraint problems stated above, some descent algorithms, such as the conjugate gradient method, can be used. These approaches require the first-order derivative or even Hessian matrix of the cost function. To this end, the adjoint of the dynamic model has to be developed. However, this process is tedious. It is encouraging that automatic differentiation (AD) techniques have been developed and applied to automatically generate the adjoint model

Fig. 12.3 Illustrations of evaluating the derivatives of the cost function with AD techniques. F denotes the whole codes to evaluate the cost function. F'^* represents the adjoint codes of the original codes, which can be used to evaluate derivatives easily.



at the level of computer codes (Bischof et al., 1996; Bischof et al., 2002; Carmichael et al., 1997; Dobmanin et al., 1995; Giering and Kaminski, 2002; Verma, 2000). AD is very effective and easy to use. This dramatically conserves time and energy of DA practitioners. The principle of AD is simple and based on two facts. First, any computer code statement can be regarded as a composition of elementary functions. Second, chain rule can be used to differentiate this composition of elementary functions. Many software packages have been developed in accordance with the principles described above for FORTRAN and C computer languages. They are given at the web site www.autodiff.org. The illustration of AD running process is presented in Fig. 12.3. We have recently applied this method in estimating LAI from satellite data (Qin et al., 2007a).

The advantages of cost function-based method are twofold. First, all data in a batch window are used to estimate the state. Second, inequality, equality, and bound constraints can be included explicitly. Its disadvantages are also apparent. First, the adjoint model is generated to evaluate the derivative of the cost function if the highly efficient descent optimization method is used. However, the development of the adjoint model requires that the dynamic model should be differentiable. This condition can not usually be met in the land surface process modeling because of many discontinuous parameterizations. Instead, SCE, SA, DE, or GA could be used, but they are computationally very slow. Second, the cost function-based method just uses the inverses of covariance matrices as the weights, as seen in Eqs. (12.3) or (12.7) and (12.8). Since the covariance is the second moment of one distribution, more information included in the distributions $p(w_i)$, $p(e_i)$, and $p(x_0)$ is not used and therefore wasted. If $p(w_i)$, $p(e_i)$, and $p(x_0)$ are normal distributions, no information is discarded since Gaussian distributions are completely characterized with the first and second moments.

12.3.5.2 Sequential-Based Methods

To derive the “optimal” sequential assimilation scheme, assume that the background states are represented by prior estimates. The sequential approaches are based on the Bayesian inference that combines the prior knowledge of the state vector and the measurement to obtain the posterior distribution of the state vector. Once obtaining the posterior distribution, everything is known about the state vector. This process can be expressed as follows.

$$p(x_{0:t}|y_{1:t}) \propto p(y_{1:t}|x_{0:t})p(x_{0:t}) \quad (12.9)$$

where $p(x_{0:t})$ represents the prior distribution of the state vector, $p(y_{1:t}|x_{0:t})$ the measurement distribution, and $p(x_{0:t}|y_{1:t})$ the posterior distribution. Typically, a Markov assumption is applied to the prior. So the state vector at time t only depends on the state vector at time $t - 1$:

$$p(x_{0:t}) = p(x_0) \prod_{i=1}^t p(x_i|x_{i-1}) \quad (12.10)$$

where $p(x_t|x_{t-1})$ is the evolution distribution, and $p(x_0)$ is the distribution of the initial state vector (background or “first guesses”). Another important assumption is that measurements are independent given the true state:

$$p(y_{1:t}|x_{0:t}) = \prod_{i=1}^t p(y_i|x_i) \quad (12.11)$$

Substituting Eqs. (12.9) and (12.10) into (12.8), we obtain:

$$p(x_{0:t}|y_{1:t}) \propto p(x_0) \prod_{i=1}^t p(y_i|x_i)p(x_i|x_{i-1}) \quad (12.12)$$

This equation implies that once new data is available, the previous estimate of the state process could be sequentially updated without having to calculate from scratch. However, this also means we have to store all state vectors up to time t , and the size of $x_{0:t}$ will expand as time goes by, becoming very large. In fact, there often is an interest in the filtering distribution $p(x_t|y_{1:t})$, that is to estimate the probability density of the current state vector conditioned on the measurements up to now. The whole filtering process is straightforward.

The filtering density, $p(x_t|y_{1:t})$, and the one step prediction, $p(x_{t+1}|y_{1:t})$, density are recursively given by a measurement update according to

$$p(x_t|y_{1:t}) = \frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{p(y_t|y_{1:t-1})} \quad (12.13)$$

$$p(y_t|y_{1:t-1}) = \int p(y_t|x_t)p(x_t|y_{1:t-1})dx_t \quad (12.14)$$

and a time update according to

$$p(x_{t+1}|y_{1:t}) = \int p(x_{t+1}|x_t)p(x_t|y_{1:t})dx_t \quad (12.15)$$

and the recursion is initiated by

$$p(x_0|y_{0:-1}) = p(x_0) \quad (12.16)$$

In the general case, one is normally unable to obtain an analytical expression of the filtering density except under the assumption of linear model and observation, and Gaussian error distributions. This leads to the prominent Kalman filter (KF). If the dynamical model and measurement process are characterized as follows

$$\begin{aligned}x_{t+1} &= Fx_t + w_t, \quad w_t \sim N(0, Q_t) \\y_t &= Hx_t + e_t, \quad e_t \sim N(0, R_t)\end{aligned}\tag{12.17}$$

the KF can be expressed as

$$\begin{aligned}\hat{x}_{t+1|t} &= F\hat{x}_{t|t} \\P_{t+1|t} &= FP_{t|t}F^T + Q_t \\ \hat{x}_{t+1|t+1} &= \hat{x}_{t+1|t} + G(y_{t+1} - H\hat{x}_{t+1|t}) \\P_{t+1|t+1} &= (I - GH)P_{t+1|t} \\G &= P_{t|t-1}H^T(H^TP_{t|t-1}H + R_t)^{-1}\end{aligned}\tag{12.18}$$

where \hat{x} denotes the estimate of the state vector.

There are two popular variants of Kalman filter. One is the Extended Kalman filter (EKF) and the other is the Ensemble Kalman filter (EnKF). EKF handles the cases where either H or F is nonlinear. Note that most land surface models (H) are nonlinear. In this case, they can be defined as its tangent linear. The EnKF is a sophisticated sequential DA method. The EnKF applies an ensemble of model states to represent the error statistics of the model estimate, uses ensemble integrations to predict the error statistics forward in time, and employs an analysis scheme which operates directly on the ensemble of model states when observations are assimilated. The EnKF efficiently manages strongly nonlinear dynamics and large state spaces and is now used in realistic applications with primitive equation models for the ocean and atmosphere. Originally proposed by Evensen (1994), the EnKF is more recently reviewed by Evensen (2003), providing detailed information on the formulation, interpretation and implementation of the EnKF.

Other sequential assimilation algorithms widely used in meteorological or oceanographic DA communities include successive correction, optimal or statistical interpolation, analysis correction, 3DVAR and 3DPSAS.

All these methods have been widely applied in a variety of fields (Evensen, 2003; Houtekamer and Mitchell, 1998; Jones et al., 2004; Kumar and Kaleita, 2003; Qin et al., 2006; Reichle et al., 2002; Wade and Eric, 2003).

12.3.5.3 Neural Network-Based Methods

In recent years, people have attempted to combine artificial neural network (ANN) and DA method (Liaqat et al., 2003a, b; Tang and Hsieh, 2001; Wu et al., 2005; Yu et al., 1997). The primary objective of these investigations is to complete or approximate the dynamic models from measurements by estimating weights and biases of ANN using DA method. The advent of the feed-forward neural network model opens the possibility of hybrid neural-dynamical models via variational DA. Such a hybrid model may be used in situations where some variables, difficult to model dynamically, have sufficient data for empirical modeling with ANN. Liaqat et al. (2001) used a neural network for constructing an arbitrary mapping function.

A neural network is trained by optimizing an object function composed of squared residuals of differential equations at collocation points and squared deviations of the observation data from the computed values. An assimilation problem is solved even if the model differential equations do not express the observed phenomena exactly. Since the dynamic model can be constructed directly from measurements and can be approximated, the computational speed greatly improves.

12.3.5.4 Particle Filtering

Particle filtering (PF) is also called sequential Monte Carlo filtering based on point mass (or “particle”) representations of probability densities, which can be applied to any state-space model and which generalize the traditional KF methods (Ristic et al., 2004; Smith et al., 2005). It is an important technique to manage the DA with elements of nonlinearity and non-Gaussianity such that the underlying dynamics of a physical system are modeled accurately. It has been applied in many engineering fields and attracted attention from some DA practitioners since the posterior distribution of state vector can be represented with Monte Carlo samples. However, KF and its variants just evaluate the mean and covariance of the posterior distribution. PF better grasps the filtering density evolution of the nonlinear system in time than KF and its variants do. PF itself also has many variants, among which the bootstrap filter, also called sampling importance resampling filter (SIR), is the most easily used. Its main steps include:

Step 1: for $t = 0$, sample $\{x_0^{(i)}\}_{i=1}^N \sim p(x_0)$;

Step 2: draw $\{\tilde{x}_{t+1}^{(i)}\}_{i=1}^N \sim p(x_{t+1}|x_t^{(i)})$. That is $\tilde{x}_{t+1}^{(i)} = f(x_t^{(i)}, u_t, \theta) + w_t^{(i)}$, $w_t^{(i)} \sim p(w_t)$;

Step 3: compute weights $u_{t+1}^{(i)} = \frac{p(y_{t+1}|\tilde{x}_{t+1}^{(i)})}{\sum_{i=1}^N p(y_{t+1}|\tilde{x}_{t+1}^{(i)})}$

where $p(y_{t+1}|\tilde{x}_{t+1}^{(i)})$ denotes the value of $p(y_{t+1} - h(\tilde{x}_{t+1}^{(i)}, \theta))$;

Step 4: resample $\{\tilde{x}_{t+1}^{(i)}\}_{i=1}^N$ with replacement according to weights $\{u_{t+1}^{(i)}\}_{i=1}^N$ in order to get $\{\tilde{x}_{t+1}^{(i)}\}_{i=1}^N$ with weights $\{1/N\}_{i=1}^N$;

Step 5: set $t = t + 1$ and go to step 2.

Observe that SIR is easy to be implemented in DA practice.

12.3.6 Observation Error Matrix

Characterizing the errors in DA is extremely important since it significantly affects the estimates (Daley, 1991). The observation error covariance matrix R in Eqs. (12.3), (12.6) and (12.16) can in principle be quite general in form, including biases, temporal correlations, extreme outliers and so on. Many extant methods

assume the error to be Gaussian with zero mean and no temporal correlation. However, more general error structures are very common, and the development of methods for dealing with such errors is an active area of current research (Evensen, 2003; Raupach et al., 2005).

All NASA Earth Observing System (EOS) land surface high-level products have been claimed to be “validated,” but most were based on limited “ground truths” (Morissette et al., 2006). The error magnitudes, and their spatial and temporal distributions, have never been well specified. Assigning uncertainties to these high level products is complex. There are a few studies reporting the accuracies of individual products, but comprehensive modeling of spatial and temporal error structures, and correlation among errors of different products has not yet been done. It raises the challenge of evaluating the uncertainty properties of major products, and it is evident that this is an enormous goal. A range of issues identified by Raupach et al. (2005) as needing to be addressed include:

- The error magnitude r_{ii} for each high-level product, inclusive of all error sources (in other words, the diagonal elements of the covariance matrix R)
- The correlations among errors in different products, quantified by the off-diagonal elements of the covariance matrix R
- The temporal structure of the errors: whether they are random in time or temporally correlated, and the possible presence of unknown long-term drifts or biases
- The spatial structure of errors (random, slowly varying or bias as for temporal structure)
- The error distribution: normal (Gaussian), lognormal, skewed or the sum of multiple error sources with different distributions, such as a small Gaussian noise together with occasional large outliers because of measurement corruption events
- Possible mismatches between the spatial and temporal averaging implicit in the model and the measurements (the “scaling problem”).

This same set of challenging issues (to define and specify the error models) pertains to other data sets, such as the error properties of direct flux measurements and direct measurements of carbon stores in addition to remote sensing of land surface properties. It is more straightforward to assign uncertainties to lower-level products than to higher-level ones. This is a major factor for considering the assimilation of direct observations (radiance/reflectance or vegetation indices), as noted above.

12.4 Recent DA Applications

DA is a powerful way to consistently combine measurements and dynamical models for the accurate estimation of model parameters and the state vector. Since DA has been explored and implemented in many applications, progress in the following selected fields is summarized.

12.4.1 Soil Moisture Estimation

While microwave remote sensing provides the opportunity to map global soil content, the revisiting frequency is limited. Moreover, the L band brightness temperature is related to only the surface soil moisture (top 5 cm) and yields little information about the root zone. A land surface process model forced by atmospheric data can produce soil moisture and temperature profiles at the model time resolution. It is obvious that just running the model without any constraints can lead to large errors due to uncertainties in the model structure, model parameters, and external forcing data. DA offers a means to consistently take advantage of both modeling and observations. Two DA algorithms have been applied to estimate not only soil moisture, but also latent heat flux (tightly associated with the root-zone moisture) using L band microwave remote sensing data.

Entekhabi et al. (1994, 1998) first applied EKF to assimilate microwave remote sensing data into a land surface model to subsequently retrieve soil water moisture and temperature profiles. Although synthetic data was used and the land surface model was relatively simple, their research opened up prospects for land surface DA. Many improvements have been made since then. Walker and Houser (2001) and Walker et al. (2001a, b) compared direct insertion and EKF assimilation methods, conducted retrieval experiments with field data, and then applied their algorithm on the regional scale for initialization of climate and hydrological models. But surface soil moisture data rather than satellite observations was assimilated in their studies.

Reichle et al. (2002) used EnKF to assimilate L-band (1.4 GHz) microwave radiobrightness observations into a land surface model, compared it with the variational method, and investigated the influence of the ensemble size on the retrieved results. Their research indicates that the EnKF is a flexible and robust DA option that gives satisfactory estimates even for moderate ensemble sizes although the updating process is suboptimal. Crow (2003) and Crow and Wood (2003) applied EnKF to assimilate L-band microwave data to correct for the impact of poorly sampled rainfall on land surface model predictions of root-zone soil moisture and surface energy fluxes. The results suggest the EnKF-based assimilation system is capable of correcting a substantial fraction of model error in root zone (40 cm) soil moisture and latent heat flux predictions associated with the use of temporally sparse rainfall measurements as forcing data. The recent studies also applied EnKF to estimation of soil moisture profiles (e.g., Merlin et al., 2006; Zhou et al., 2006). Margulis et al. (2002) explained why the EnKF technique is so appealing for soil moisture estimation: (i) its sequential nature is well suited to real-time data streams and forecasting; (ii) it not only provides an estimate of surface and profile soil moisture, but information about the statistical confidence of estimation; (iii) it is sufficiently straightforward to use with “off-the-shelf” models; and (iv) it is relatively efficient, making its application to large-scale problems feasible.

The variational cost function-based approach has also been successfully applied to hydrological studies in recent years. The key point in using the variational assimilation method is to develop the adjoint of a dynamic model. This requires the model should be differentiable, but land surface schemes usually are not.

Consequently some approximations have to be made. Reichle (2000) and Reichle et al. (2001) first used the variational method to assimilate L-band microwave data into a complicated land surface model to retrieve the soil moisture profile. The results showed that the state estimates obtained from the assimilation algorithm improve significantly over prior model predictions derived without assimilating radio-brightness data.

Soil moisture estimation is also one of the major goals in the regional and global land surface DA systems, such as the global land DA system (Rodell et al., 2004), the North American Land DA system (Mitchell et al., 2004), and the South America Land DA system (de Goncalves et al., 2006).

12.4.2 Energy Balance Fluxes Estimation

Accurate estimation of energy and momentum fluxes, especially sensible and latent heat fluxes, between the land surface and the atmospheric boundary layer, is required in a wide variety of agricultural, hydrological, and meteorological applications (Courault et al., 2005; Su, 2002). Many methods, such as eddy correlation and Bowen ratio, can be used to measure these fluxes at the field level. Their applicability, however, is limited on the fine spatial scale. Currently, the only way to achieve this goal of mapping fluxes regionally is to use remote sensing techniques that provide various spatial and temporal imageries covering large areas.

Estimation of surface energy balance components using remote sensing data can be roughly divided into three categories: empirical, residual and DA. *Empirical methods* directly build on the relationship between remote sensing products, such as various vegetation indices and retrieved LST for the estimation. *Residual methods* of the energy budget couple some empirical formulas and physical mechanisms to realize the estimation of evapotranspiration (ET or LE) and sensible heat flux (H) by using remote sensing to directly estimate input parameters, such as the surface energy balance system (SEBS) (Su, 2002). In recent years, *DA methods* have integrated remote sensing data and soil–vegetation–atmosphere transfer models for estimating surface fluxes, and have achieved encouraging results. This new method has been receiving considerable attention from researchers in recent years (Boni et al., 2001a, b; Caparrini and Castelli, 2004; Caparrini et al., 2003; Castelli et al., 1999) since it combines dynamic models and temporal remote sensing data based on the control theory in order to accurately retrieve critical parameters for flux estimation. The first two methods more or less use only the instantaneous data and empirical relationships.

Castelli et al. (1999) developed a simple land surface scheme and its adjoint, defined a moisture index as control variables, used radiometric surface temperature as observations, and performed DA experiments. Results indicated that this algorithm can retrieve land surface energy balance components effectively. Boni et al. (2001a, b) investigated the impact of land surface temperature sampling frequency on assimilation results and suggested that satellite remote sensing of land surface

temperature may be used to provide estimates of components of the surface energy balance and land surface control on evaporation. Caparrini et al. (2004, 2003) replaced a moisture index with evaporative fraction as control variables, made many assimilation experiments, and applied this algorithm on the regional scale using AVHRR land surface temperature as observations.

12.4.3 Carbon Fluxes Estimation

It is increasingly recognized that DA (model–data synthesis, data–model fusion or many other names) is one of the methods that can meet such a challenge. The US Climate Change Science Program strategic plan and the North American Carbon Program (NACP) plan repeatedly call for the development of DA methods for carbon cycle studies. A near-term priority identified by the US Carbon DA Program Workshop Report (Fung et al., 2002) was urgent support “for interdisciplinary teams to develop component and coupled prototype carbon assimilation models.” Canadell et al. (2004) identified three fundamental research areas that require major development in order to provide policy relevant knowledge for managing the carbon-climate system over the next few decades, with the first area being “carbon observations and multiple constraint data assimilation.”

For carbon cycle studies, some efforts have been made to develop DA systems, but most of these are based on “top-down” strategies that begin with measured changes in atmospheric carbon concentrations and attempts to infer the spatial distribution and magnitude of the net exchange. For example, the Carbon Data-Model Assimilation (C-DAS) project at NCAR (National Center for Atmospheric Research) is developing a carbon DA system based on their atmospheric transport model (<http://dataportal.ucar.edu/CDAS>).

An equally important approach in carbon cycle science is the “bottom-up” approach (Cihlar et al., 2000) that starts with a specific land parcel to account for the various pathways of carbon exchange between the ecosystem and the atmosphere, and then scales up to much larger regions. It relies on both ecosystem process models and spatial data sets. The model can be developed at the local level and validated using conventional measurements. Satellite observations provide spatial distribution and frequent up-to-date information on the rate of change of the variables driving the model. These models can in turn be used to estimate spatial and temporal variations in CO₂ uptake and release over large areas, if appropriate “input” data sets (e.g., vegetation and soils maps, weather data) are available.

Several obstacles stand in the way of extensive use of ecosystem process models for extrapolation. The major issue is the requirement of much more information on vegetation characteristics than is readily available or even known for many areas of the globe. Many parameters are not measurable at earth system scales, because of high small-scale spatial variability (e.g., any soil property), high temporal variability (e.g., stomatal conductance), and physical inaccessibility (e.g., most roots, deep soil). Moreover, from the modeling perspective, some processes are well understood

(e.g., photosynthesis, decomposition of litter and soil organic matter), while other processes are not (C-allocation among plant tissues, T-sensitivity of humus decomposition). As a result, large uncertainties arise in calculating terrestrial carbon cycle using the “bottom-up” approach (Barrett, 2002).

Wang and Barrett (2003) developed a modeling framework that synthesizes various types of field measurements at different spatial and temporal scales to estimate monthly means (and their standard deviations) of gross photosynthesis, total ecosystem production, net primary production (NPP), and net ecosystem production (NEP) for eight regions of the Australian continent between 1990 and 1998. Williams et al. (2005) developed a DA approach combining stock and flux observations with a dynamic model to improve estimates of ecosystem carbon exchanges. Rayner et al. (2005) developed a terrestrial carbon cycle DA system (CCDAS) for determining the space-time distribution of terrestrial carbon fluxes for the period 1979–1999. Hazarika et al. (2005) integrated the MODIS LAI product with an ecosystem model for accurate estimation of NPP. Validations of results in Australia and the USA show that NPP estimated using the DA method to be more accurate than that generated by the data “forcing” method. Their research demonstrates the utility of combining satellite observations with an ecosystem process model to achieve improved accuracy in estimates and monitoring global net primary productivity.

Barrett et al. (2005) demonstrated that a “multiple-constraints” model–data assimilation scheme using a diverse range of data types offers improved predictions of carbon and water budgets at regional scales. Xu et al. (2006) applied the Bayesian probability inversion and a Markov chain Monte Carlo (MCMC) technique to a terrestrial ecosystem model to analyze uncertainties of estimated carbon (C) transfer coefficients and simulated C pool sizes. Their study shows that the combination of a Bayesian approach and MCMC inversion technique effectively synthesizes information from various sources for assessment of ecosystem responses to elevated CO₂. Sacks et al. (2006) used a model–data synthesis approach with a simplified carbon flux model to extract process-level information from 5 years of eddy covariance data at an evergreen forest in the Colorado Rocky Mountains. Including water vapor fluxes, in addition to carbon fluxes, in the parameter optimization did not yield significantly more information about the partitioning of the net ecosystem exchange of CO₂ into gross photosynthesis and ecosystem respiration. Sacks et al. (2007) used the model–data synthesis method to address fundamental questions about climate effects on terrestrial ecosystem net CO₂ exchange.

12.4.4 Crop Yield Estimation

Advance information on crop yield during the crop growing season is vital for effective crop management and for national food security policy. Agricultural harvested grain yield is a reliable way to estimate crop yields by sampling field measurements of standing crops. However, this method is both time consuming and costly, with results that are not available until after harvest. In the last three decades, satellite

remote sensing data have been used to estimate crop yields over large areas because these methods are more cost effective and more timely than traditional survey procedures (MacDonald and Hall, 1980).

Earlier studies were mostly based on empirical regression methods that relate crop yield to remotely sensed surface reflectance and their combinations (i.e., vegetation indices). These relationships could be described with linear, cubic polynomial, or exponential regression (Jiang et al., 2003). Essentially a statistical model, this method cannot predict the time-dependent processes of crop growth. Additionally, the relationship between yield and NDVI may not be accurate under extreme weather conditions.

Mathematical crop growth models simulate fundamental processes such as photosynthesis, respiration, biomass partitioning, and water and nitrogen transfers (Baret et al., 2000). This allows researchers to evaluate a wide array of alternatives, and to assemble processes in an integrated package. Along with supporting better crop management decisions, mechanical crop growth models can simulate the dynamics of LAI and other structural properties of the crop fields (e.g., height and biomass). The combination of remote sensing and crop growth simulation models is increasingly recognized as a promising approach for monitoring growth and estimating yield (Bauman, 1992).

The use of crop models is often limited by uncertainties in their input parameters such as soil conditions, sowing date, planting density and initial field conditions. Except in some controlled experimental fields, many of these parameters are poorly known. Remote sensing can play a critical role in helping identify the field and crop status from estimated biophysical parameters (Clevers and Leeuwen, 1996). Remote sensing data, therefore, can be assimilated with crop growth models to improve their overall performance.

Several assimilation schemes, of various degrees of complexity and integration, have been developed in the last 10 years (Moulin et al., 1998). Various methods for integrating a crop growth model with remote sensing data were described by Mass (1988a, b) and were also reviewed by Fischer et al. (1997) and Moulin et al. (1998). Mass (1993) compared the results of calibrating a crop simulation model using LAI observations obtained either from field sampling or remote sensing. Winter wheat yields were modeled more accurately using remotely sensed LAI observations than field-sampled LAI observations (Mass, 1993). This difference appeared to result from the apparent ability of the remotely sensed LAI observations to better represent the photosynthetically active plant area in the crop canopy.

Bach et al. (2003) experimented with coupling a raster-based PROMET-V model with the radiative transfer model GeoSAIL to predict biomass and yield. In their study, LAI, fraction of brown leaves, and surface soil moisture were used as free variables; surface reflectance was used as the control variable. Their assimilation procedure produced improved biomass and yield results. Guérif and Duke (2000) combined the SUCROS crop model with the SAIL canopy reflectance model for accurate estimation of sugar beet yield. Ground measured reflectance was used to match the predicted reflectance. One limitation of their study is that many crop and soil parameters need to be obtained from field measurements. The SAIL model

has also been integrated with the EPIC crop model to estimate the yield of spring wheat in North Dakota (Doraiswamy et al., 2003). Planting date is the only adjustable variable in this model. The estimated yields are mostly within 10% of the NASS (National Agricultural Statistical Service) reports. In this work, climate data are based on interpolation of weather station measurements. The crop area mask is based on the 1 km AVHRR classification. NDVI is also calculated from AVHRR. Although the AVHRR data set is easily accessible, using it will compromise the precision of analyses owing to outdated calibrations and the application of partial atmospheric corrections. Doraiswamy et al. (2004) used a look-up table (LUT) method to estimate LAI from 250 m MODIS reflectance data. The crop modeled LAI was adjusted to fit the MODIS simulated LAI by changing planting time, time when maximum LAI is attained, and the beginning of leaf senescence.

Jongschaap and Schouten (2005) estimated the regional wheat yield by assimilating SPOT data into a crop model. Microwave remote sensing data (ERS SAR C-band) were used to estimate regional wheat flowering dates to calibrate a wheat growth simulation model. Pauwels et al. (2007) assessed to what extent the results of a fully coupled hydrology–crop growth model can be optimized through the assimilation of observed LAI and soil moisture values using the EnKF. A practical procedure using the variational optimization method to predict crop yield at the regional scale from MODIS data was recently developed by Fang et al. (2007b). This method outputs agronomic variables (yield, planting, emergence and maturation dates) and biophysical parameters (e.g., LAI).

12.5 Summary

Though DA has reached maturity in meteorological and oceanographic applications, the land community has just begun to employ it for estimation of land surface variables. Herein, basic DA principles have been described, critical issues in land DA have been identified, and many of the latest applications in hydrology, carbon cycle, and agriculture have been introduced. Because of the continuous improvement in DA methods and computational technology, along with an available wealth of remote sensing observations and extensive ground observation networks, DA is likely to become the best technique to monitor and map land surface environments by integrating *a priori* knowledge with an enormous variety and sheer volume of data.

Revisiting the key issues addressed in this chapter, the following questions are put forth:

1. Which remotely sensed data products and other measurements are the most valuable for land DA?
2. How are the differences between the constrained data sets and predictions of a dynamic model characterized?
3. How can the model parameters/variables be estimated effectively?
4. How will the errors of the assimilated data be specified?
5. Which land surface properties can be estimated?
6. What types of dynamic models are suitable for DA?

To answer these questions, “explorers” from the community are needed to traverse this new frontier, and we hope they will do so.

The computational issues of DA have not been broached since DA algorithms are usually computationally expensive. However, the fast pace of computer science advances promises to minimize such obstacles. Community efforts are needed to build the practical tools so that few researchers have to start from scratch. A good example of such endeavor is the land information system developed by NASA (Kumar et al., 2006). Ideally, more educational tools for this enterprise will follow.

Acknowledgements We are very grateful for Dr. John Townshend for providing valuable comments. S. Liang is partially funded by NASA under grant NNG04GL85G.

References

- Bach H, Mauser W, Schneider K (2003) The use of radiative transfer models for remote sensing data assimilation in crop growth models. In: J Stafford, A Werner (eds), Precision agriculture: Papers from the 4th European Conference on Precision Agriculture, Berlin, Germany, 15–19 June 2003
- Baret F, Weiss M, Troufleau D, Prevo L, Combal B (2000) Maximum information exploitation for canopy characterisation by remote sensing. *Asp. Appl. Biol.* 60(71–82)
- Barrett DJ (2002) Steady state turnover time of carbon in the Australian terrestrial biosphere. *Glob. Biogeochem. Cycles* 16(4):Art. No. 1108
- Barrett DJ, Hill MJ, Hutley LB, Beringer J, Xu JH, Cook GD, Carter JO, Williams RJ (2005) Prospects for improving savanna biophysical models by using multiple-constraints model-data assimilation methods. *Aust. J. Bot.* 53(7):689–714
- Bauman BAM (1992) Linking physical remote sensing models with crop growth simulation models applied for sugarbeet. *Int. J. Remote Sens.* 14:2565–2581
- Bischof C, Khademi P, Mauer A, Carle A (1996) Adifor 2.0: Automatic Differentiation of Fortran 77 Programs. *IEEE Comp. Sci. Eng.* 3(3):18–32
- Bischof CH, Hovland PD, Norris B (2002) Implementation of automatic differentiation tools. Proceedings of the 2002 ACM SIGPLAN Workshop on Partial Evaluation and Semantics-Based Program Manipulation (PEPM-02), vol. 37, pp. 98–107
- Boni G, Castelli F, Entekhabi D (2001a) Sampling Strategies and Assimilation of Ground Temperature for the Estimation of Surface Energy Balance Components. *IEEE Trans. Geosci. Remote Sens.* 39(1):165–172
- Boni G, Entekhabi D, Castelli F (2001b) Land data assimilation with satellite measurements for the estimation of surface energy balance components and surface control on evaporation. *Water Resour. Res.* 37(6):1713–1722
- Canadell JG, Ciais P, Cox P, Heimann M (2004) Quantifying, understanding and managing the carbon cycle in the next decades. *Climatic Change* 67(2–3):147–160
- Caparrini F, Castelli F (2004) Estimation of surface turbulent fluxes through assimilation of radiometric surface temperature sequences. *J. Hydrometeorol.* 5:145–159
- Caparrini F, Castelli F, Entekhabi D (2003) Mapping of land-atmosphere heat fluxes and surface parameters with remote sensing data. *Bound. Lay. Meteorol.* 107:605–633
- Carmichael GR, Sandu A, Potra FA (1997) Sensitivity analysis for atmospheric chemistry models via automatic differentiation. *Atmos. Environ.* 31(3):475–489
- Castelli F, Entekhabi D, Caporali E (1999) Estimation of surface heat flux and an index of soil moisture using adjoint-state surface energy balance. *Water Resour. Res.* 35(10):3115–3125

- Chepurin GA, Carton JA, Dee D (2005) Forecast model bias correction in ocean data assimilation. *Monthly Weather Rev.* 133(5):1328–1342
- Cihlar J, Denning AS, Gosz J (2000) Global Terrestrial Carbon Observation: Requirements, Present Status, and Next Steps, Report of a Synthesis Workshop, 8–11 February 2000, Ottawa, Canada
- Clevers JPGW, Leeuwen V (1996) Combined use of optical and microwave remote sensing data for crop growth monitoring. *Remote Sens. Environ.* 56:42–51
- Courault D, Seguin B, Olioso A (2005) Review on estimation of evapotranspiration from remote sensing data: from empirical to numerical modeling approaches. *Irrig. Drain. Syst.* 19:223–249
- Crow WT (2003) Correcting land surface model predictions for the impact of temporally sparse rainfall rate measurements using an ensemble Kalman filter and surface brightness temperature observations. *J. Hydrometeorol.* 4(5):960–973
- Crow WT, Wood EF (2003) The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using Ensemble Kalman filtering: a case study based on ESTAR measurements during SGP97. *Adv. Water Resour.* 26(2):137–149
- Dai Y, Zeng X, Dickinson RE, Baker I, Bonan GB, Bosilovich MG, Denning AS, Dirmeyer PA, Houser PR, Niu G, Oleson KW, Schlosser CA, Yang Z.L (2003) The Common Land Model (CLM). *Bull. Am. Meteorol. Soc.*, 84(8):1013
- Daley R (1991) Atmospheric data analysis. Cambridge University Press, New York.
- de Goncalves LGG, Shuttleworth WJ, Burke EJ, Houser P, Toll DL, Rodell M, Arsenault K (2006) Toward a South America Land Data Assimilation System: Aspects of land surface model spin-up using the simplified simple biosphere. *J. Geophys. Res. -Atmos.* 111(D17)
- Demarty J, Ottle C, Braud I, Olioso A, Frangi JP, Gupta HV, Bastidas LA (2005) Constraining a physically based Soil-Vegetation-Atmosphere Transfer model with surface water content and thermal infrared brightness temperature measurements using a multiobjective approach. *Water Resour. Res.* 41(1):Art. No. W01011
- Dobmanin M, Liepelt M, Schittkowski K (1995) Algorithm 746: PCOMP: a Fortran Code for Automatic Differentiation. *ACM Trans. Math. Software* 21(3):233–266
- Doraiswamy PC, Hatfield JL, Jackson TJ, Akhmedov B, Prueger J, Stern A (2004) Crop condition and yield simulation using Landsat and MODIS. *Remote Sens. Environ.* 92(4):548–559
- Doraiswamy PC, Moulin S, Cook PW, Stern A (2003) Crop yield assessment from remote sensing. *Photogramm. Eng. Remote Sens.* 69(6):665–674
- Duan QY, Gupta VK, Sorooshian S (1993) Shuffled complex evolution approach for effective and efficient global minimization. *J. Optim. Theor. Appl.* 76(3):501–521
- Duan QY, Sorooshian S, Gupta VK (1992) Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.* 28(4):1015–1031
- Duan QY, Sorooshian S, Gupta VK (1994) Optimal use of sce-ua global optimization method for calibrating watershed models. *J. Hydrol.* 158:265–284
- Entekhabi D, Nakamura H, Njoku EG (1994) Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely sensed observations. *IEEE Trans. Geosci. Remote Sens.* 32(2):438–448
- Evensen G (1994) Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast statistics. *J. Geophys. Res.* 99(C5):10143–10162
- Evensen G (2003) The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dyn.* 53:343–367
- Fang H, Kim H, Liang S, Schaaf C, Strahler A, Townshend GRG, Dickinson R (2007a) Spatially and temporally continuous land surface albedo fields and validation. *J. Geophys. Res.* 112, doi: 10.1029/2006JD008377
- Fang H, Liang S, Hoogenboom G, Teasdale J, Cavigelli M (2007b) Crop yield estimation through assimilation of remotely sensed data into DSSAT-CERES. *Int. J. Remote Sens.* doi: 10.1080/01431160701408386
- Fang H, Liang S, Townshend J, Dickinson R (2007c) Spatially and temporally continuous LAI data sets based on a new filtering method: examples from North America. *Remote Sens. Environ.* doi:10.1016/j.rse.2006.07.026

- Fischer A, Kergoat L, Dedieu G (1997) Coupling satellite data with vegetation functional models: review of different approaches and perspectives suggested by the assimilation strategy. *Remote Sens. Rev.* 15:283–303
- Fung I, Kalnay E, Schimel D, Denning S, Doney S, Pawson S (2002) A US Carbon Data Assimilation Program Workshop Report, University of Maryland, College Park
- Giering R, Kaminski T (1998) Recipes for Adjoint Code Construction. *ACM Trans. Math. Software* 24(4):437–474
- Giering R, Kaminski T (2002) Generating recomputations in reverse mode AD. In: G Corliss, A Griewank, FCL Hascoet, U Naumann (eds), *Automatic differentiation of algorithms: from simulation to optimization*. Springer, Heidelberg, pp 283–291
- Goldberg DE (1989) *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley, Reading, MA
- Guérif M, Duke CL (2000) Adjustment procedures of a crop model to the site specific characteristics of soil and crop using remote sensing data assimilation. *Agric. Ecosyst. Environ.* 81:57–69
- Gupta HV, Sorooshian S, Yapo PO (1998) Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information. *Water Resour. Res.* 34(4):751–763
- Hazarika MK, Yasuoka Y, Ito A, Dye D (2005) Estimation of net primary productivity by integrating remote sensing data with an ecosystem model. *Remote Sens. Environ.* 94(3):298–310
- Houtekamer PL, Mitchell HL (1998) Data assimilation using an Ensemble Kalman Filter technique. *Monthly Weather Rev.* 126(3):796–811
- Ingber L (1989) Very fast simulated re-annealing. *J. Math. Comput. Model.* 12(8):967–973
- Jiang D, Wang NB, Yang XH, Wang JH (2003) Study on the interaction between NDVI profile and the growing status of crops. *Chin. Geogr. Sci.* 13(1):62–65
- Jones JW, Graham WD, Wallach D, Bostick WM, Koo J (2004) Estimating soil carbon levels using an Ensemble Kalman filter. *Trans. ASAE* 47(1):331–339
- Jongschaap REE, Schouten LSM (2005) Predicting wheat production at regional scale by integration of remote sensing data with a simulation model. *Agron. Sustain. Dev.* 25(4):481–489
- Julien Y, Sobrino JA, Verhoef W (2006) Changes in land surface temperatures and NDVI values over Europe between 1982 and 1999. *Remote Sens. Environ.* 103:43–55
- Kaminski T, Knorr W, Rayner PJ, Heimann M (2002) Assimilating atmospheric data into a terrestrial biosphere model: a case study of the seasonal cycle. *Glob. Biogeochem. Cycles* 16(4):Art No. 1066
- Kumar P., Kaleita AL (2003) Assimilation of near-surface temperature using Extended Kalman Filter. *Adv. Water Resour.* 26:79–93
- Kumar SV, Peters-Lidard CD, Tian Y, Houser PR, Geiger J, Olden S, Lighty L, Eastman JL, Doty B, Dirmeyer P, Adams J, Mitchell K, Wood EF, Sheffield J (2006) Land information system: an interoperable framework for high resolution land surface modeling. *Environ. Model. Software*, 21(10):1402–1415
- Li X, Koike T, Pathmathevan M (2004) A very fast simulated re-annealing (VFSA) approach for land data assimilation. *Comput. Geosci.* 30(3):239–248
- Liang S (2004) *Quantitative remote sensing of land surfaces*. Wiley, New York
- Liaqat A, Fukuhara M, Takeda T (2001) Application of neural network collocation method to data assimilation. *Comput. Phys. Commun.* 141(3):350–364
- Liaqat A, Fukuhara M, Takeda T (2003a) Applying a neural network collocation method to an incompletely known dynamical system via weak constraint data assimilation. *J. Mar. Syst.* 131(8):1696–1714
- Liaqat A, Fukuhara M, Takeda T (2003b) Optimal estimation of parameters of dynamical systems by neural network collocation method. *Comput. Phys. Commun.* 150(3):215–234
- Losa SN, Kivman GA, Ryabchenko VA (2004) Weak constraint parameter estimation for a simple ocean ecosystem model: what can we learn about the model and data? *J. Mar. Syst.* 45(1–2): 1–20
- MacDonald R, Hall F (1980) Global crop forecasting. *Science* 208:670–679
- Margulis S, McLaughlin DB, Entekhabi D, Dunne S (2002) Land data assimilation and soil moisture estimation using measurements from the Southern Great Plains 1997 Field Experiment. *Water Resour. Res.* 38(12):1299, doi:10.1029/2001WR001114

- Marler RT, Arora JS (2004) Survey of multiobjective optimization method for engineering. *Struct. Multidisc. Optim.* 26(6):369–395
- Mass S (1988a) Using satellite data to improve model estimates of crop yield. *Agron. J.* 80:662–665
- Mass SJ (1988b) Use of remotely-sensed information in agricultural crop growth models. *Ecol. Model.* 41:247–268
- Mass SJ (1993) Within-season calibration of modeled wheat growth using remote sensing and field sampling. *Agron. J.* 85(3):669–672
- Merlin O, Chehbouni A, Boulet G, Kerr Y (2006) Assimilation of disaggregated microwave soil moisture into a hydrologic model using coarse-scale meteorological data. *J. Hydrometeorol.* 7(6):1308–1322
- Mitchell KE, Lohmann D, Houser PR, Wood EF, Schaake JC, Robock A, Cosgrove BA, Sheffield J, Duan QY, Luo LF, Higgins RW, Pinker RT, Tarpley JD, Lettenmaier DP, Marshall CH, Entin JK, Pan M, Shi W, Koren V, Meng J, Ramsay BH, Bailey AA (2004) The multi-institution North American Land Data Assimilation System (NLDAS): utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res. -Atmos.* 109(D7):Art. No. D07S90
- Moody EG, King MD, Platnick S, Schaaf CB, Gao F (2005) Spatially complete global spectral surface albedos: value-added datasets derived from terra MODIS land products. *IEEE Trans. Geosci. Remote Sens.* 43(1):144–158
- Morissette JT, Baret F, Privette JL, Myneni RB, Nickeson JE, Garrigues S, Shabanov NV, Weiss M, Fernandes RA, Leblanc SG, Kalacska M, Sanchez-Azofeifa GA, Chubey M, Rivard B, Stenberg P, Rautiainen M, Voipio P, Manninen T, Pilant AN, Lewis TE, Iames JS, Colombo R, Meroni M, Busetto L, Cohen WB, Turner DP, Warner ED, Petersen GW, Seufert G, Cook R (2006) Validation of global moderate-resolution LAI products: a framework proposed within the CEOS Land Product Validation subgroup. *IEEE Trans. Geosci. Remote Sens.* 44(7):1804–1817
- Moulin S, Bondeau A, Delécolle R (1998) Combining agricultural crop models and satellite observations; from field to regional scales. *Int. J. Remote Sens.* 19(6):1021–1036
- Natvik LJ, Eknes M, Evensen G (2001) A weak constraint inverse for a zero-dimensional marine ecosystem model. *J. Mar. Syst.* 28(1–2):19–44
- O'Neill A, Mathieu PP, Zehner C (2004) Making the most of earth observation with data assimilation. *ESA Bulletin* 118:33–38
- Pauwels VRN, Verhoest NEC, De Lannoy GJM, Guissard V, Lucau C, Defourny P (2007) Optimization of a coupled hydrology-crop growth model through the assimilation of observed soil moisture and leaf area index values using an ensemble Kalman filter. *Water Resour. Res.* 43(4)
- Qin J, Liang S, Li X, Wang J (2007a) Development of the adjoint model of a canopy radiative transfer model for sensitivity study and inversion of leaf area index. *IEEE Trans. Geosci. Remote Sens.* (revised)
- Qin J, Liang S, Liu R, Zhang H, Hu B (2007b) A weak-constraint based data assimilation scheme for estimating surface turbulent fluxes. *IEEE Geosci. Remote Sens. Lett.* 4(4):649–653
- Qin J, Yan G, Liu S, Liang S, Zhang H, Wang J, Li X (2006) Application of Ensemble Kalman Filter to geophysical parameter retrieval in remote sensing: a case study of kernel-driven BRDF model inversion. *Sci. China, Series D*, 49(6):632–640
- Raupach MR, Rayner PJ, Barrett DJ, DeFries RS, Heimann M, Ojima DS, Quegan S, Schimmlus CC (2005) Model-data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications. *Glob. Change Biol.*, 11(3):378–397
- Rayner PJ, Scholze M, Knorr W, Kaminski T, Giering R, Widmann H (2005) Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS). *Glob. Biogeochem. Cycles* 19(2):Art. No. GB2026
- Reichle RH (2000) Variational Assimilation of Remote Sensing Data for Land Surface Hydrologic Applications, MIT, Boston, MA
- Reichle RH, McLaughlin DB, Entekhabi D (2001) Variational data assimilation of microwave radiobrightness observations for land surface hydrology applications. *IEEE Trans. Geosci. Remote Sens.* 39(8):1708–1718

- Reichle RH, McLaughlin DB, Entekhabi D (2002) Hydrologic data assimilation with the Ensemble Kalman Filter. *Month. Weather Rev.* 130:103–114
- Ristic B, Aruampalam S, Gordon N (2004) *Beyond the Kalman Filter*. Artech House, Boston/London
- Rodell M, Houser PR, Jambor U, Gottschalk J, Mitchell K, Meng CJ, Arsenault K, Cosgrove B, Radakovich J, Bosilovich M, Entin JK, Walker JP, Lohmann D, Toll D (2004) The global land data assimilation system. *Bull. Am. Meteorol. Soc.* 85(3):381
- Sacks WJ, Schimel DS, Monson RK, 2007. Coupling between carbon cycling and climate in a high-elevation, subalpine forest: a model-data fusion analysis. *Oecologia* 151(1):54–68
- Sacks WJ, Schimel DS, Monson RK, Braswell BH (2006) Model-data synthesis of diurnal and seasonal CO₂ fluxes at Niwot Ridge, Colorado. *Glob. Change Biol.* 12(2):240–259
- Smith A, Doucet A, Freitas ND, Gordon N (eds) (2005) *Sequential Monte Carlo methods in practice*. Springer, New York
- Storn R, Price K (1997) Differential Evolution – a simple and efficient heuristic for global optimization over continuous space. *J. Global Optim.* 11:341–359
- Storn R, Price KV (1996) Minimizing the real function of the ICEC'96 contest by differential evolution, *IEEE conference on evolutionary computation*, pp 842–844
- Su Z (2002) The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. *Hydrol. Earth Syst. Sci.* 6(1):85–99
- Tang YM, Hsieh WW (2001) Coupling neural networks to incomplete dynamical systems via variational data assimilation. *Month. Weather Rev.* 129(4):818–834
- Verma A (2000) An introduction to automatic differentiation. *Curr. Sci.* 78(7):804–807
- Vrugt JA, Gupta HV, Bastidas LA, Bouten W, Sorooshian S (2003) Effective and efficient algorithm for multiobjective optimization of hydrologic models. *Water Resour. Res.* 39(8):Art No. 1214
- Wade TC, Eric FW (2003) The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using Ensemble Kalman Filtering: a case study based on ESTAR measurements during SGP97. *Adv. Water Resour.* 26(2):137–149
- Walker JP, Houser PR (2001) A methodology for initializing soil moisture in a global climate model: assimilation of near-surface soil moisture observations. *J. Geophys. Res. -Atmos* 106(D11):11761–11774
- Walker JP, Willgoose GR, Kalma JD (2001a) One-dimensional soil moisture profile retrieval by assimilation of near-surface measurements: a simplified soil moisture model and field application. *J. Hydrometeorol.* 2(4):356–373
- Walker JP, Willgoose GR, Kalma JD (2001b) One-dimensional soil moisture profile retrieval by assimilation of near-surface observations: a comparison of retrieval algorithms. *Adv. Water Resour.* 24(6):631–650
- Wang YP, Barrett DJ (2003) Estimating regional terrestrial carbon fluxes for the Australian continent using a multiple-constraint approach I. Using remotely sensed data and ecological observations of net primary production. *Tellus Ser B Chem. Phys. Meteorol.* 55:270–289
- Williams M, Schwarz PA, Law BE, Irvine J, Kurpius MR (2005) An improved analysis of forest carbon dynamics using data assimilation. *Glob. Change Biol.* 11(1):89
- Wu J, Fukuhara M, Takeda T (2005) Parameter estimation of an ecological system by a neural network with residual minimization training. *Ecol. Model.* 189(3–4):289–304
- Xu T, White L, Hui DF, Luo YQ (2006) Probabilistic inversion of a terrestrial ecosystem model: analysis of uncertainty in parameter estimation and model prediction. *Glob. Biogeochem. Cycles* 20(2)
- Yu TW, Iredell M, Keyser D (1997) Global data assimilation and forecast experiments using SSM/I wind speed data derived from a neural network algorithm. *Weather Forecast.* 12(4):859–865
- Zhou YH, McLaughlin D, Entekhabi D (2006) Assessing the performance of the ensemble Kalman filter for land surface data assimilation. *Month. Weather Rev.* 134(8):2128–2142

China Land Soil Moisture EnKF Data Assimilation Based on Satellite Remote Sensing Data^{*}

Shi Chunxiang^{1*} Xie Zhenghui² Qian Hui³ Liang Miaoling⁴ & Yang Xiaochun¹

¹National Satellite Meteorological Center, China Meteorological Administration, Beijing 100081;

²Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029;

³Institute of Geology, Chinese Academy of Sciences; Key Geodynamics Laboratory, Ministry of Ministry of Land and Resources, 100037;

⁴National Meteorological Center, China Meteorological Administration, Beijing 100081

Abstract: Soil moisture plays an important role in land-atmosphere interactions. It is an important geophysical parameter in research on climate, hydrology, agriculture, and forestry. Soil moisture has important climatic effects by influencing ground evapotranspiration, runoff, surface reflectivity, surface emissivity, surface sensible heat and latent heat flux. At the global scale, the extent of its influence on the atmosphere is second only to that of sea surface temperature. At the terrestrial scale, its influence is even greater than that of sea surface temperatures. This paper presents a China Land Soil Moisture Data Assimilation System (CLSMDAS) based on EnKF and land process models, and results of the application of this system in the China Land Soil Moisture Data Assimilation tests. CLSMDAS is comprised of the following components: 1) A land process model—Community Land Model Version 3.0 (CLM3.0)—developed by the US National Center for Atmospheric Research (NCAR); 2) Precipitation of atmospheric forcing data and surface-incident solar radiation data come from hourly outputs of the FY2 geostationary meteorological satellite; 3) EnKF (Ensemble Kalman Filter) land data assimilation method; and 4) Observation data including satellite-inverted soil moisture outputs of the AMSR-E satellite and surface soil moisture observation data. Results of soil moisture assimilation tests from June to September 2006 were analyzed with CLSMDAS. Both simulation and assimilation results of the land model reflected reasonably the temporal-spatial distribution of soil moisture. The assimilated soil moisture distribution matches very well with severe summer droughts in Chongqing and Sichuan Province in August 2006, the worst since the foundation of the People's Republic of China in 1949. It also matches drought regions that occurred in eastern Hubei and southern Guangxi in September.

Key words: EnKF land data assimilation, AMSR-E soil moisture, FY2C stationary satellite, high-resolution precipitation, surface incident solar radiation

Soil moisture affects matter and energy balances between land and atmosphere through influencing moisture flux, sensible heat, and latent heat flux. Accurate estimates of the spatial and temporal distribution of soil moisture are of vital importance to further understanding the ecological and physical processes of land and land-atmosphere interactions, and are of great importance in the research and application of meteorology in environmental management, ecology, hydrology, and agriculture. Soil moisture information can currently be acquired with the following methods: regular and field observations, satellite remote sensing, and land hydrological modeling and simulation. Soil moisture data acquired through regular observation has a low temporal frequency and few spatial points, and using some field observations is interim and limited to a small range. Regular and field soil moisture observations provide only point-based data and cannot represent spatial variations in the soil moisture profile. Synchronous field observation is usually expensive. Satellite remote sensing can provide global soil moisture data with a high spatial and temporal resolution, which cannot be achieved via the regular observation network. However, soil moisture observations based on satellite remote sensing are also problematic. Precision of soil moisture inversion is related to soil type, ground surface roughness, and vegetation coverage. The uncertainty in these parameters results in inversion errors and insufficient temporal and spatial resolutions. Soil moisture obtained through land hydrological modeling simulation and calculation has good temporal frequency and spatial distribution. However, its precision is influenced greatly by model structure and input data. Land assimilation technology serves as an effective solution to these problems [1-2]. It is a technical approach that integrates optimally physical process information of the land model.

Data assimilation methods used currently in atmospheric, marine, and land data assimilation systems include mainly the optimal interpolation method, 3-D variation method, 4-D variation method, Kalman filter, and EnKF. EnKF is an assimilation method is used widely in land model assimilation systems, and uses the Monte Carlo method (overall integration method) to calculate predicted error covariance of the state. It was proposed by Evensen on the basis of Epstein's stochastic dynamic prediction theory [3]. The model state prediction is considered an

^{*} Supported by the National High Technology Research and Development Program of China (Grant No. 2007AA12Z144), the Chinese COPES Project (Grant No. GYHY200706005), the China Meteorological Administration New Technology Promotion Project (Grant No. CMATG2008Z04), and the National High Technology Research and Development Program of China (Grant No. 2009AA12Z129)

* E-mail: shicx@cma.gov.cn

approximately stochastic dynamic prediction. An overall state is used to represent the probability density function in a stochastic dynamic prediction. By integrating ahead, we can calculate easily the statistical characteristics (e.g., mean value and covariance) corresponding to probability density functions at various times of the overall state. EnKF key feature is that it does not require linearized model operators or observation operators as the Kalman filter does. Huge and low-efficiency computation requirements are its greatest drawbacks. As all current land models are single-column models, they contain much fewer model state variables than atmospheric and marine models and thus have a higher computational efficiency. Strong non-linear characteristics in land process models are significant.

Starting in 1998, the Department of Hydrological Sciences and Data Assimilation Office of US NASA Goddard Space Flight Center organized many organizations to initiate the research of Land Data Assimilation System (LDAS), and developed the Global Land Data Assimilation System (GLDAS). US National Oceanic and Atmospheric Administration also offers support for hydrology-based LDAS research at the scale of the North American continent (North America Land Data Assimilation System project). Land data assimilation uses many kinds of new satellite and ground-based observation data to generate optimal land state and flux data (<http://www.knmi.nl/samenw/LDAS/>). In 2001, Europe started research on the European Land Data Assimilation System. The main purpose of ELDAS research is to improve forecasting and monitoring of floods and droughts. The ELDAS project conducted to: assemble specialists in soil moisture assimilation to design and produce a unified, flexible, and practical data assimilation framework; use independent observation data to verify the assimilated soil moisture field; evaluate the seasonal water circulation forecast after soil moisture data assimilation; evaluate flood risk; establish an exemplary data set covering at least one season over the European region; use data from new satellite platforms MSG (MTEOSAT Second Generation) and SMOS (ESA Soil Moisture/Ocean Salinity Mission); and make European contributions to the Global Land Data Assimilation System. ELDAS has been applied in four numerical forecasting centers (ECMWF, DWD, CNRM and INM) and already been brought from the research stage into the service operation stage (<http://www.knmi.nl/samenw/eldas/>). The West China Land Data Assimilation System (WCLDAS) was researched and developed by the Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences and Department of Atmospheric Sciences, School of Resources and Environment, Lanzhou University. The institute established a single-point soil moisture assimilation system based on EnKF and SiB2 models, and completed preliminary analysis and evaluation of this assimilation system [4-5]. Yang et. al. [6] developed an automatic rating system used to estimate moisture and energy balances of soil by assimilating AMSR-E vertical polarization 6.9GHz and 18.7GHz luminance temperature data. Tian Xiangjun [7-14], Zhang Shenglei [15-16], and Jia Binghao [17] et. al. researched the improvement of land data assimilation methods and the rating of microwave luminance temperature assimilation observation operator models.

This paper explains the composition of the China Land Soil Moisture Data Assimilation System (CLSMDAS), processing methods and quality inspection of atmospheric driving data, the performance test of CLSMDAS with single-point observation data, error analysis of observed soil moisture data, and design, and analysis of the CLSMDAS experiment.

1. China Land Soil Moisture Data Assimilation System

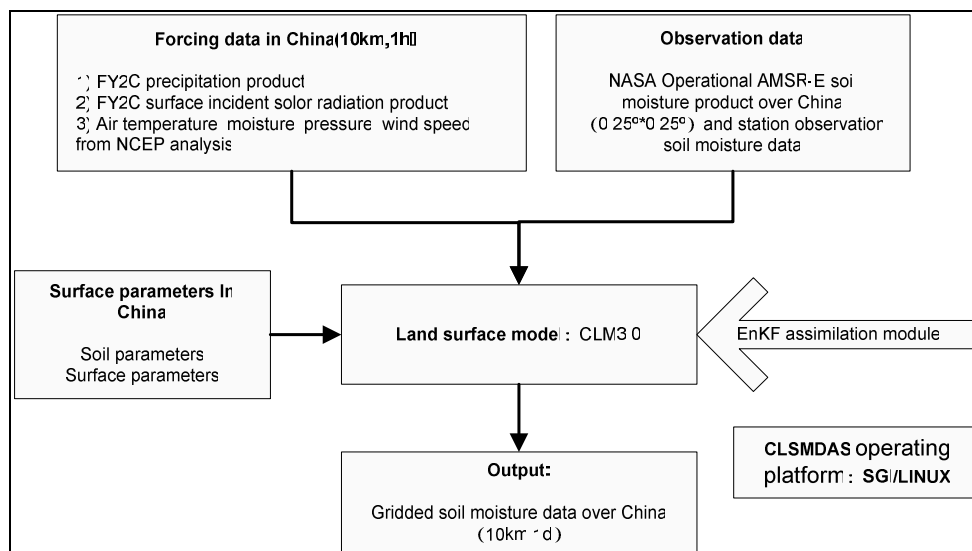


Figure 1 Structure of the China Land Soil Moisture Data Assimilation System (CLSMDAS)

CLSMDAS is comprised mainly of the following (See Fig.1): 1) Land model: A NCAR-CLM3.0 land model used widely at present; 2) Driving data: High spatial and temporal resolution precipitation estimates and ground-incident

solar radiation data acquired from the FY2 geostationary meteorological satellites, and surface air temperatures, humidity, atmospheric pressure and wind speed in the NCEP reanalysis data set are interpolated temporally and spatially to construct atmospheric driving data of the land model; 3) Data assimilation method: We have chosen the EnKF assimilation method which is suitable for land data assimilation; 4) Observation data: Soil moisture inversion data acquired from AMSR-E carried by the AQUA EOS satellite, and soil moisture data observed at ground level, are used; 5) Output data set: Assimilated soil moisture grid point data can be output. CLSMDAS works at present on a SGI-LINUX platform. Its individual components are described below.

1.1. Land process model

The land process model is the core of the land data assimilation system. The extent to which the land process model describes accurately and reasonably the energy and material exchanges at the ground surface has a direct influence on the output of the land assimilation system, forecasting the state of the next iteration. A land process model—Community Land Model Version 3.0 (CLM3.0 for short)—developed by US National Center for Atmospheric Research (NCAR) [18] was used here.

The CLM3.0 model is designed mainly for coupling with the atmospheric numerical model and provides the surface albedo (direct and scattered light within the visible and infrared bands), upward long-wave radiation, sensible heat flux, latent heat flux, water vapor flux, and east-to-west and south-to-north surface stress needed by the atmospheric model. These parameters are controlled by many ecological and hydrological processes. The model simulates the phenology of leaves and physiological and water circulations of pores. Ecological differences between vegetation types and thermal and hydraulic differences between different soil types are also considered. Each grid cell can be covered by several types of land surfaces. The river transport model transfers downward to the sea. Because this land model has to be coupled with the climate model and numerical weather forecast model, compromise has to be made between effectiveness and complexity of the parametric computation of land processes. Land spatial non-uniformity is realized by nesting subgrids where one grid cell contains several land units, snow and soil cylindrical blocks, and different types of vegetation. Each cell contains several land units, each land unit contains a different number of soil and snow cylindrical blocks, and each cylindrical block may contain several types of vegetation functions [19]. Many researchers have tested and evaluated the application of the CLM3.0 model in China and proved the feasibility of its application in China [18].

The EnKF assimilation module is coupled mainly with the soil moisture module. Below is the equation of one-dimensional soil water vertical motion (horizontal flow is ignored) used in CLM3.0:

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q}{\partial z} - E - R_{fm}, \quad (1)$$

where θ is the volumetric water content of soil ($\text{m}^3 \text{ m}^{-3}$), q is vertical soil water flux (mm s^{-1}), E is evaporation rate (mm s^{-1}). R_{fm} is melting (negative) or freezing (positive) rate, and z is vertical distance to the ground surface (positive under q and z directions) [18].

1.2. Atmospheric driving data

One of the difficulties preventing us improving the land model simulation precision is lack of long-term, high-resolution, observational atmospheric driving data. Atmospheric driving data usually requires diurnal precipitation, atmospheric temperature, relative humidity, atmospheric pressure, wind speed and downward solar radiation data, which cannot be observed directly at the global scale. Much work has been done on processing atmospheric driving data for offline land models. Qian's [20] research showed that there are false long-term variations with NCEP's reanalysis of precipitation and terrestrial radiation. The errors in precipitation data will affect inevitably the effects of soil moisture prediction, simulation, and assimilation. Other researchers have also realized these problems, and as a result either directly use observation or adjust the reanalysis data.

In this paper, inverted precipitation and ground-incident solar radiation products acquired through high spatial and temporal resolution geostationary satellite data were introduced and combined with the NCEP reanalysis data to establish a set of high-quality atmospheric driving data which covers the Chinese region.

(1) High spatial- and temporal-resolution precipitation

FY2C/D/E satellite precipitation estimate outputs can be downloaded free of charge from the website of the China Satellite Data Service Center (<http://satellite.cma.gov.cn/>). Precipitation estimate outputs include daily accumulative and 6-hour accumulative precipitation estimate outputs. A new hourly accumulative precipitation estimate output was made available in September 2007. The integral time step of the CLM3 land model is 30 min. Therefore, we need

atmospheric forcing field data with high temporal resolution. Compared with the 6-hour accumulative precipitation data, the hourly precipitation data is more suitable to drive simulation of the land model. To this end, Shi Chunxiang [21] developed a method to perform time-weighted interpolation of accumulative precipitation on the basis of hourly geostationary satellite cloud category information. This was applied to the FY2C 6-hour precipitation estimate output provided by the China Satellite Data Service Center, to derive a precipitation data set with $0.1^\circ \times 0.1^\circ$ spatial resolution and 1-h temporal resolution, and used automatic rainfall observation data from China to verify and evaluate the precipitation estimate data. The result indicated that the method was reasonable. See Fig. 2 for the process flow of geostationary satellite 1-hour precipitation output. See [22] for details on the processing and quality evaluation of precipitation data.

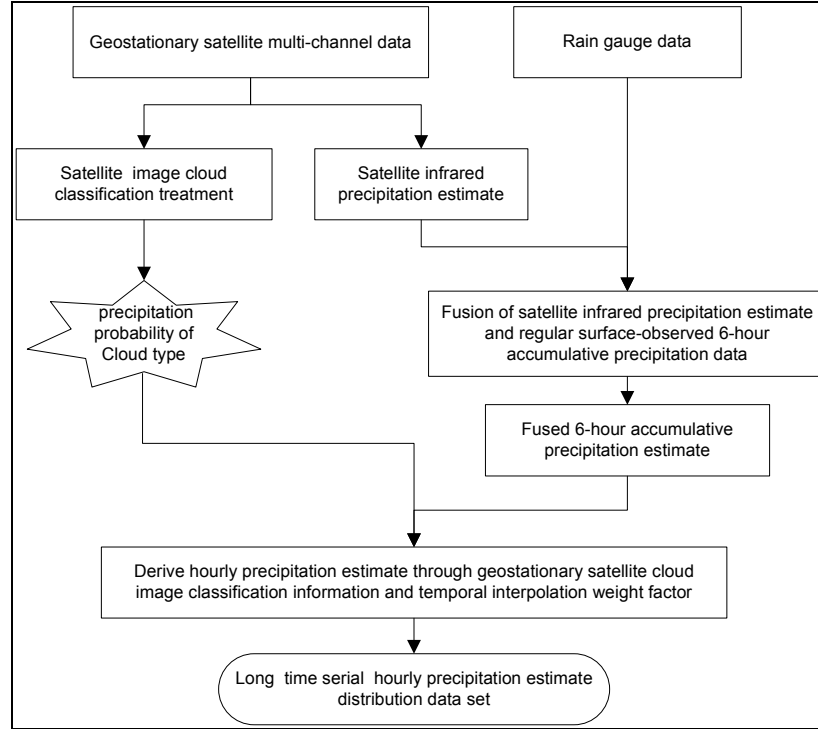


Figure 2 Precipitation process flow with high spatial and temporal distribution

(2) Ground-incident solar radiation data with high spatial and temporal resolution

FY2C/D/E ground-incident solar radiation out can be downloaded free of charge from the website of China Satellite Data Service Center (<http://satellite.cma.gov.cn/>). The temporal resolution of the product is 1 day. The spatial resolution is $0.5^\circ \times 0.5^\circ$. Due to its low spatial and temporal resolution, the output does not meet the requirements of the land model for atmospheric forcing data. Therefore, an inversion algorithm was used here on the FY2C/D/E ground-incident solar radiation output of the China Satellite Data Service Center (<http://satellite.cma.gov.cn/>). Visible-band observations acquired through the FY2C geostationary meteorological satellite are used to generate a ground-incident solar radiation data set with a temporal resolution of 1 hour and a spatial resolution of $0.1^\circ \times 0.1^\circ$. The quality of this data set was verified with the solar radiation data from five China Meteorological Administration climatic observation stations.

The discrete ordinate method proposed by Stamnes [23] was used to calculate radiation transfer in the inversion algorithm for the ground-incident solar radiation output. This algorithm calculates the radiance of any direction and thus gives consideration to the anisotropy of solar radiation reflected at the top of the atmospheric layer. It first calculates the radiance of solar radiation reflected towards the satellite observation direction at the top of the atmospheric layer, and then converts it into the visible-light two-way albedo observed through the visible light channel of the satellite. The period in which the solar radiation incident at the top of the atmospheric layer penetrates the atmosphere and reaches the ground surface contains a series of physical interaction processes with the atmosphere and the ground surface. The inversion model considers mainly: 1) Ozone absorption; 2) Repeated Rayleigh scattering of molecules; 3) Repeated scattering and absorption of cloud droplets; 4) Water vapor absorption; 5) Repeated scattering and absorption of aerosol; and 6) Repeated reflection between the ground surface and the atmosphere. Similar to the work of Stuhlmann [24], we designed a 5-layer planoparallel ideal atmospheric model which is not

uniform in the vertical direction. It is divided into five solar spectral intervals (0.2-0.4, 0.4-0.5, 0.5-0.6, 0.6-0.7, 0.7-4.0 μm) to calculate the scattering, absorption, and reflection of solar radiation that take place within them. Judged by the extents to which the above-mentioned physical processes influence the ground-incident solar radiation, cloud has a greater effect than all other factors by one order of magnitude. Ground-incident solar radiation is determined mainly by cloud, which needs to be stressed in the inversion model.

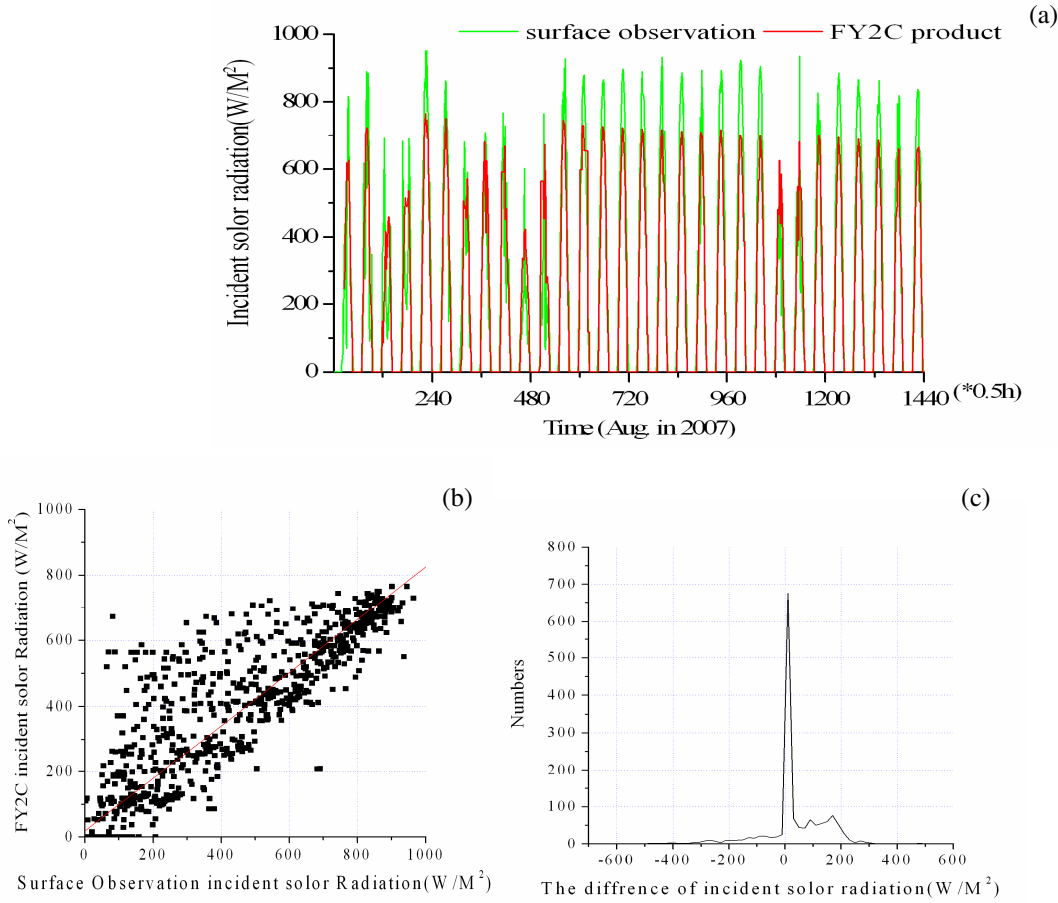


Figure 3 Comparison between FY2C satellite-inverted and station observation surface incident solar radiation data. (a) Variation of surface incident solar radiation data with time. Surface observation is station observation data of surface incident solar radiation (green line). FY2C product is output of FY2C satellite-inverted surface incident solar radiation data (red line). (b) Scatter diagram of FY2C satellite-inverted data and ground-based observation data of surface incident solar radiation. (c) Difference between FY2C satellite-inverted data and station observation data of surface incident solar radiation (W m^{-2}).

Yuan Wanping performed many tests and verifications of the inversion algorithm of ground-incident solar radiation (<http://satellite.cma.gov.cn/>). In this paper, the ground-based observation data provided by newly-built climatic observation stations of China Meteorological Administration was used for a comparative verification of the FY2C ground-incident solar radiation data between July 2007 and December 2007. The five climatic observation stations were located in Xilinhot, Shouxian, Zhangye, Dali, and Dianbai. Fig. 3a shows time-varying ground-incident solar radiation data observed at Xilinhot Observation Station and inverted from FY2C satellite data. The horizontal coordinate is time, the unit of measurement (UOM) is hours; the accumulative total counts from 00:00 August 1, 2007 (universal time). The vertical coordinate is ground-incident solar radiation, the UOM is W m^2 . Fig. 3a shows that when the value of ground-incident solar radiation is high, the value observed at ground level is greater than the satellite-inverted ground-incident solar radiation. This is probably because satellite observation covers a relatively large area while ground-based observation covers a very limited area, which can almost be considered a point. Therefore, the result of satellite inversion is closer to the mean state. Fig. 3b is a scatter diagram of ground-incident solar radiation data observed at ground level and inverted by the FY2C satellite. Fig. 3c is a histogram of the difference between ground-incident solar radiation data inverted by the FY2C satellite and observed at ground level. Both Fig. 3b and Fig. 3c show that satellite-inverted values are higher than ground-observed values. Fig. 3c also shows that under most circumstances the difference between satellite-inverted data and ground-observed data of solar radiation is approximately $<50 \text{ W m}^{-2}$.

3) Atmospheric driving field data set

Surface temperature, humidity, atmospheric pressure, and wind speed that drive operation of the CLM3.0 land model were derived from interpolation of NCEP reanalysis data with $1^\circ \times 1^\circ$ resolution and 6-h interval. These parameters were combined with precipitation and ground-incident solar radiation data, with large spatial and temporal distribution inverted through geostationary satellite to form an atmospheric forcing data set that drives the land model.

In this paper, the relevant data between July 2005 and June 2009 were processed and then subject to quality control and verification. An atmospheric driving data set is generated (UOM of one month). Temporal resolution of the data set is 1 hour, the horizontal resolution is $0.1^\circ \times 0.1^\circ$, the spatial coverage is $15-55^\circ \text{ N}$, $75-135^\circ \text{ E}$, and the data format is NETCDF.* This atmospheric driving field data set was used in the China Soil Moisture Land Data Assimilation experiment described below. A $0.1^\circ \times 0.1^\circ$ atmospheric driving field data set was used as the basic data. Data corresponding to assimilation experiments with other spatial resolutions were derived through spatial interpolation of this atmospheric driving field data.

1.3. Land data assimilation method

Evensen proposed a new solution to square root analysis [25] on the basis of standard EnKF [26]. It does not require observation disturbance in the calculation of analytic field sets and will thus reduce or eliminate sample error brought by observation disturbance. This algorithm also does not require additional assumptions or approximations in calculation of the analytic field structures. For example, it neither requires the assumption that an observation is non-related to the disturbance of the state variable set, nor requires the inversion of the observation error covariance matrix. This algorithm simplifies calculation.

Below are the traditional EnKF equations [25-26]:

(1) State error covariance matrix

Matrix defining the state of the set:

$$A = (\psi_1, \psi_2, \dots, \psi_N) \in R^{n \times N} \quad (2)$$

where ψ_i ($i = 1, \dots, N$) is the member of the sample set, n is dimensionality of the state variable, N is the number of samples in the set.

Matrix defining the disturbance of the set:

$$A' = A - \bar{A} = A(I - 1_N) \quad (3)$$

wherein $\bar{A} = A 1_N$.

State error covariance matrix:

$$P^f = \frac{A' A'^T}{N - 1} \quad (4)$$

Each element in matrix 1_N is $1/N$.

Observation error covariance matrix:

Given an observation $d \in R^m$, observation vectors of N disturbances are defined as follows:

$$d_j = d + \varepsilon_j, j = 1, \dots, N, \quad (5)$$

$$D = (d_1, d_2, \dots, d_N) \in R^{m \times N} \quad (6)$$

$$E = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N) \in R^{m \times N} \quad (7)$$

*Researchers in need of this data set may contact the author at Email: shicx@cma.gov.cn

$$R = \frac{EE^T}{N-1} \quad (8)$$

where m is dimensionality of the observation vectors.

(2) Analysis equation

$$A^a = A + P^f H^T (HP^f H^T + R)^{-1} (D - HA) \quad (9)$$

where H is the operator (H may be non-linear, in which case $HA = H(A)$).

Based on the update of the covariance matrix of the traditional Kalman filter analysis equation, a square root algorithm was used to calculate the update of disturbance in the state variables of the set:

$$P^a = P^f - P^f H^T (HP^f H^T + R)^{-1} HP^f, \quad (10)$$

where H is the observation operator, P and R are the state error covariance matrix and observation error covariance matrix, respectively. The superscripts a , f and T represent analysis field, forecast field, and matrix transposition, respectively. Suppose A and A' are the state matrix and disturbance matrix of the set, respectively. The state variable in this paper is soil volumetric moisture. Therefore, $A = (\theta_1, \theta_2, \dots, \theta_N) \in \mathbb{R}^{n \times N}$. In addition, the observation vector $d \in \mathbb{R}^m$ is set as the satellite-inverted soil moisture. E is the observation disturbance set. Matrices are defined as $S = HA'$ and $C = SS^T + (N-1)R$. The algorithm was performed through the following steps:

- 1) Calculate Matrix C and decompose the characteristic values of C , i.e., $ZA Z^T = C$;
- 2) Update the mean value of the state variable set $\bar{\theta}^a = \bar{\theta}^f + A'S^T Z A^1 Z^T (d - H\bar{\theta}^f)$;
- 3) Calculate Matrix $X_2 = A^{\frac{1}{2}} Z^T S$;
- 4) Perform SVD decomposition $U_2 \Sigma_2 V_2^T = X_2$;
- 5) Solve the disturbance of the analytic set of state variables $A^{a'} = A' V_2 \sqrt{I - \Sigma_2^T \Sigma_2} \Theta$ (Θ is any orthogonal matrix) and then add the mean value of the state variable set $\bar{\theta}^a$ calculated with (2) to derive the analysis field A^a of the state variables.

During the establishment of CLSMDAS, we first used a simple soil water module to establish a land near-surface soil moisture assimilation model based on the EnKF method, constructed a set of ideal near-surface observation data, and then performed an ideal simulation verification test to prove the accuracy and feasibility of the EnKF assimilation module because only near-surface observation was available [22]. On the basis of this analysis, an EnKF assimilation module was coupled with the CLM3.0 model to establish CLSMDAS.

1.4. AMSR-E soil moisture output

The advanced microwave scanning radiometer (AMSR-E) carried by EOS/Aqua is the world's first sensor capable of providing soil moisture service outputs at the global scale. These outputs have been applied widely in hydrological, meteorological, and climatic studies. The working frequencies of AMSR-E are 6.925 GHz, 10.65 GHz, 18.7 GHz, 23.8 GHz, 36.5 GHz, and 89 GHz.

The AMSR-E land inversion algorithm is based on the radiation transfer model. Three major geophysical parameters are obtained through inversion: soil moisture m_e , vegetation moisture content w_e , and surface temperature T_e . According to the relationship between the observed luminance temperatures and the geophysical variables related to the atmosphere, the model equation can be simplified to:

$$T_{Bi} = \Phi_i(x) \quad (11)$$

where $x = \{x_j\}$, x_j are geophysical variables i.e., soil moisture, vegetation moisture content, and surface temperature; T_{Bi} is the luminance temperature observed through channel i ; $\Phi_i(x)$ are standards for the functional

relationship between parameters and the luminance temperature.

AMSR-E luminance temperatures (T_{bs}) were first subjected to projection treatment. Projected Luminance T_{bs} is classified to make ensure it matches pixel points of the inversion conditions. It is matched with auxiliary data. The soil moisture inversion process comprises: 1) Quality control of input data; 2) project re-sampling; 3) surface classification; 4) eliminating data that do not meet inversion requirements; 5) inversion; and 6) obtaining soil moisture [27].

We downloaded and processed AMSR-E daily output data between 2004 and 2007 and used regular nationwide soil moisture observation data, detailed soil moisture observation data from Inner Mongolia and Henan Province, and AMSR-E inverted soil moisture data for a comparative analysis.

Comparative analysis results showed that: 1) Both spatial and temporal variation in AMSR-E inverted soil moisture data was small; 2) The difference between AMSR-E inverted soil moisture and ground-based observed soil moisture was related significantly to level of soil moisture. The inverted soil moisture is more accurate in arid and semi-arid regions than in wet regions. This coincides with the theoretical analysis results; and 3) The difference between AMSR-E inverted soil moisture and ground-level data observed at Station 117 in Inner Mongolia was generally smaller than the differences at Station 115 in Henan. This is because observation points of Station 117 in Inner Mongolia were located mostly in grasslands and observation points in Henan were located mostly in farmlands. When there is high vegetation coverage, the soil moisture inversion capacity of microwaves is reduced. See [27] for details.

2. China Land Soil Moisture Data Assimilation experiment and results verification

Single-point ecological station observations capture key atmospheric driving observation data, and surface and soil data. After CLSMDAS was established, first we used single-point observation data acquired from ecological stations for the land soil moisture data assimilation experiment and analyzed the performance of this assimilation system. After the single-point soil moisture assimilation experiment, we prepared Chinese atmospheric driving field data with temporal resolution of 1 hour and spatial resolution of $0.25^\circ \times 0.25^\circ$ and AMSR-E inverted soil moisture data. We then performed the China soil moisture assimilation experiment. In this section, the processes of the single-point and regional soil moisture assimilation experiments are described and the results analyzed.

2.1. Single-point land soil moisture data assimilation experiment

Shouxian National Climatic Observation Station was located by Huaihe River in the north of Anhui Province at $116^\circ 47'E$, $32^\circ 33'N$. It has an average elevation of 23.5 m ASL. It is 25,000 m² in extent and belongs to Huanghuai Agricultural Ecological Observation Area. Data from Shouxian Observation Station were acquired between April and June 2004. The temporal resolution of the observation data is 30 min. The atmospheric driving data include: atmospheric temperature, relative humidity, wind speed, atmospheric pressure and precipitation at 2 m altitude; surface short-wave radiation and 10 cm soil moisture data observed at ground level. The CLSMDAS established in this paper was used to carry out a soil moisture data assimilation sensitivity experiment. In CLSMDAS, the state error covariance matrix formed a state variable set by randomly disturbing the state variables. Then Eq. (4) was used to calculate the state error covariance matrix. The observation error covariance matrix was calculated with Eq. (8). The following four soil moisture assimilation experiments were designed (Table 1).

Experiment 1: Alternate the model background error and observation error to assess effects of on assimilation results. The model background error was set at 0.03. In other words $\theta_i = \theta_b * (1 + \gamma)$ was used to generate a sample set. γ was a random value within $[-0.03, 0.03]$. Figure 4a shows time-variation curves of modeled, observed, and assimilated soil moisture when the observation error was set at 0.01 and 0.03, respectively (empirical value is used in this paper). The assimilated soil moisture is closer to the observation data when the observation error is set at 0.01.

Experiment 2: Change the size of the EnKF sample set and assess the effects of samples in the set on assimilation results. Fig. 4b shows time-variation curves of modeled, observed, and assimilated soil moisture data when the number of samples in the set was set at 10, 20, and 50. The more samples the set contained, the closer the assimilation result became to observed data. Differences between the assimilation results of 10, 20, and 50 samples were not very significant.

Experiment 3: Assess the effects of soil moisture depth on the assimilation results. Because the first layer of ground-observed soil moisture is the mean value of a 10 cm-deep soil layer which covers the first four soil moisture layers in CLM3.0, we designed two experiments. The observed soil moisture of the first layer was assimilated with the model first-layer soil moisture. The first-layer observed soil moisture (10 cm) was then divided into four layers and then assimilated with soil moisture data of the first four layers of the model for comparison. Fig. 4c shows

time-variation curves of the modeled, observed, and assimilated soil moisture. There were major differences between the results of both assimilation schemes. The more the observed soil moisture information, the closer to the observed results the assimilated soil moisture was.

Experiment 4: Assess the effects of soil moisture frequency on assimilation results. The frequency of the observation data was set at once an hour, once a day or every 3 days. Fig. 4d shows time-variation curves of the modeled, observed, and assimilated soil moisture data. The higher the frequency of the observation data, the closer the assimilation data became to the observed data. The lower the observation frequency, the closer the assimilation result became to the modeled result. Fig. 4e and Fig. 4f are the time-variation curves of temperature and precipitation, respectively, used during the above-mentioned experiments.

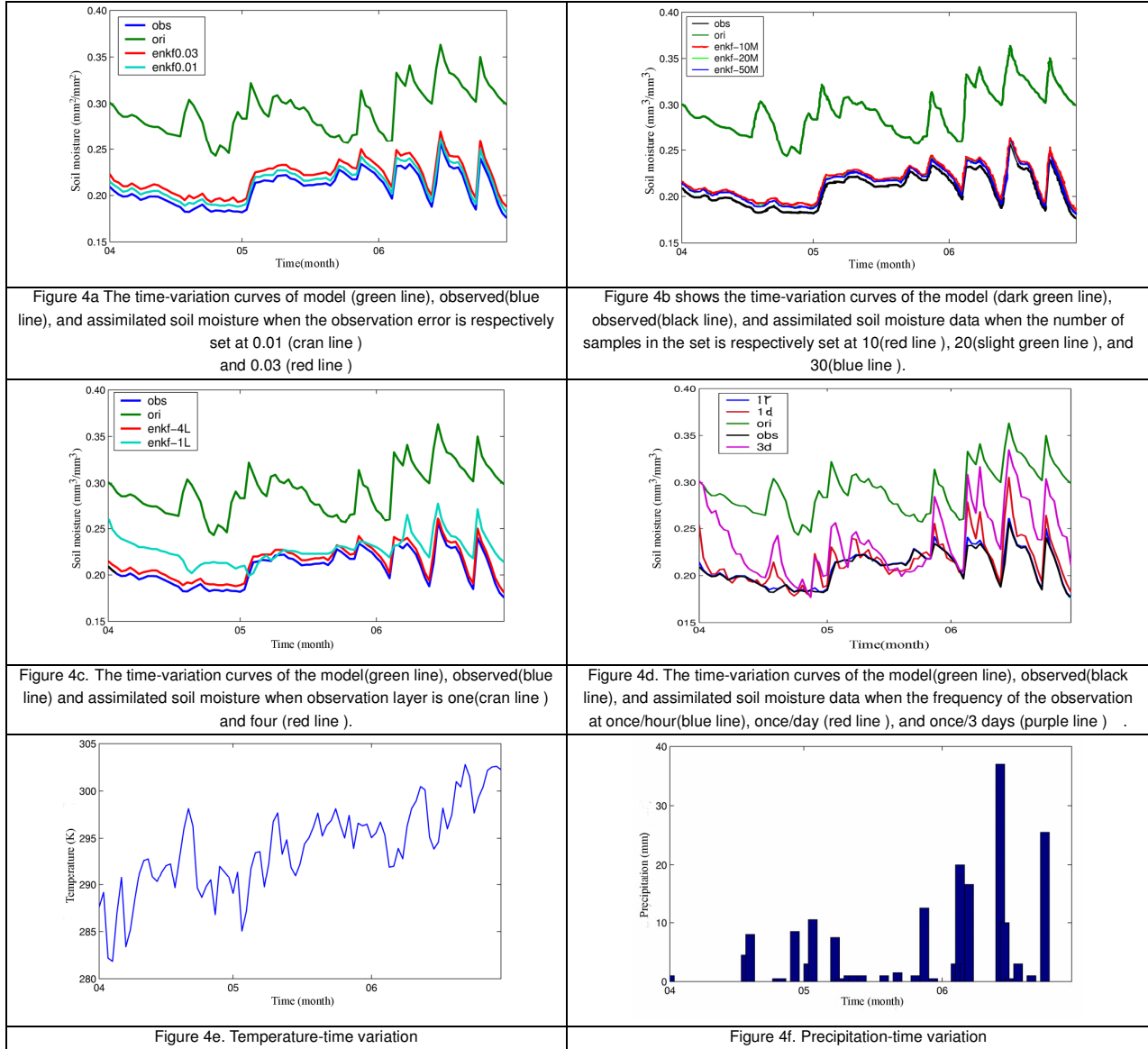


Figure 4 Results of assimilation experiment at Shouxian between April and June, 2004.

Table 1 Four Soil Moisture Assimilation Experiment Schemes

Experiment	Background error (mm^3/mm^3)	Observation error (mm^3/mm^3)	Number of samples	Observation frequency (h)	Number of assimilation layers
1	0.03	0.01/0.03	30	0.5	1
2	0.03	0.01	10/20/50	0.5	1
3	0.03	0.01	30	0.5	1/4
4	0.03	0.01	30	0.5/24/72	1

2.2. China Land Soil Moisture Data Assimilation Experiment

CLSMDAS was used to perform a number of experiments. The spatial extent of the assimilation experiment was 15–55°N, 75–135°E. CLM3.0 was used as the land model. FY2C satellite precipitation estimate data and FY2C ground-incident solar radiation data described in 1.2 were used as the precipitation and radiation in the atmospheric driving data. Temperature, moisture, atmospheric pressure, and wind speed data were derived through interpolation of NCEP reanalysis data. The observation data were AMSR-E inverted soil moisture data downloaded from NASA's official website (<http://nsidc.org/data/amsre/>). EnKF assimilation was used as the assimilation method. The spatial resolution of soil moisture data was $0.25^\circ \times 0.25^\circ$ after assimilation. There were 10 layers in the vertical direction. The classification of soil moisture layers was consistent with that of the CLM3.0 model. The temporal integral step of the model was 30 min. The experiment in 2.1 showed that the sum of the soil moisture values of the topmost four layers in CLSMDAS was consistent in the physical significance of the 10 cm soil moisture data observed at ground level. Therefore, the output soil moisture data of the topmost four layers were combined and then processed into grid-point daily mean and monthly mean soil moisture data for further analysis.

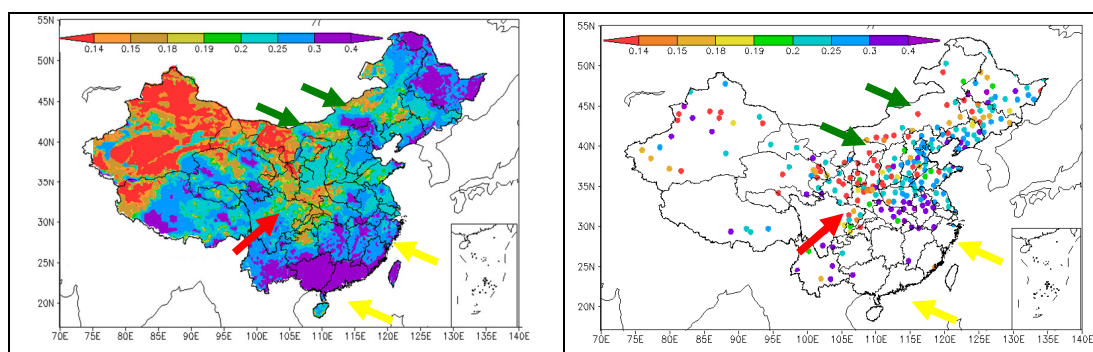


Figure 5 (a) Assimilated Chinese soil moisture distribution. (b) Observation station soil moisture.

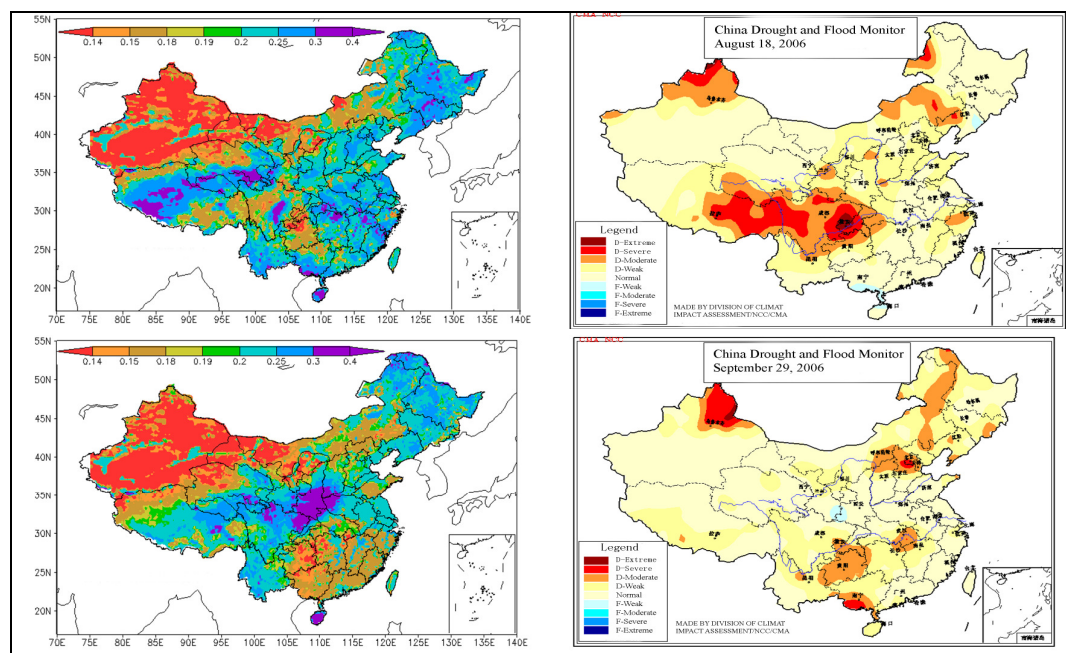


Figure 6 A comparison between Chinese assimilated soil moisture distribution (top left: August 2006, bottom left: September 2006) and Chinese drought and flood monitoring map (right top: August 2006, Right bottom: September 2006)

Analysis of the assimilation experiment between June and September 2006 shows that the CLSMDAS soil moisture distribution is consistent with soil moisture observed on the ground. Due to insufficient available ground-level observation soil moisture data, the large error in observed soil moisture, and the inherent high variability of the spatial distribution of soil moisture, it is difficult to verify the assimilated soil moisture data. The Drought and Flood Climate

Bulletin is an authoritative output issued regularly by the National Climate Center. The climatic drought and flood distribution maps of this are drawn based on comprehensive analysis of station observation data observed. In this paper, the distribution map of ground soil moisture observation stations and the climatic drought and flood distribution map published by the National Climate Center were used to analyze the CLSMDAS soil moisture data (Fig. 5 and Fig. 6).

Fig. 5a shows the CLSMDAS soil moisture on July 18, 2006 and Fig. 5b shows Chinese 10-cm deep soil moisture observations on July 18, 2006 (Fig. 5b). According to the CLSMDAS soil moisture distribution map, soil moisture is relatively high in most southern regions. In Fig. 5b, there are few observation stations in the south, but all of these have high soil moisture (yellow arrows). In the low soil moisture regions in eastern and central of Inner Mongolia, the observed and assimilated soil moisture were consistent (green arrows). In the lower-soil-moisture regions in central China, the observed and assimilated soil moisture data were also quite consistent (red arrows).

Fig. 6 compares the Chinese assimilated soil moisture distribution and a Chinese drought and flood monitoring map published by the National Climate Center [28]. According to the China Drought and Flood Climate Bulletin published by the National Climate Center, the most serious summer drought since 1949 took place in Sichuan Province and Chongqing in August 2006. National climatic drought monitoring results from August 18 showed that the west of Chongqing and the east of Sichuan still remained in severe to extremely severe drought. The great majority of Sichuan, eastern Tibet, southwest Hubei, northwest Hunan, northern Guizhou, northern Xinjiang and southern Gansu and eastern Inner Mongolia remained in medium to severe drought. The east of Jiangnan, the great majority of northern China, and the east of northwest China had mild to medium droughts (<http://climat.cma.gov.cn/>). The CLSMDAS soil moisture distribution on August 18, 2006 showed significantly lower soil moisture values in drought regions of Chongqing and Sichuan than in surrounding regions. According to the China Drought and Flood Climate Bulletin issued by the National Climate Center, in September 2006 droughts persisted or developed in northern China, southern Chongqing, northern Guizhou and northern Xinjiang, and eastern Inner Mongolia. Different degrees of drought occurred in eastern Hubei and southern Guangxi. In the CLSMDAS monthly mean soil moisture map of September 28, 2006, the low value centers of soil moisture are consistent with drought regions reported in the China Drought and Flood Climate Bulletin. It should be noted that the comparison between the Chinese assimilated soil moisture distribution map of China and the national climatic drought and flood distribution map published by National Climate Center is only qualitative because these two maps do not represent the same phenomenon. The climatic drought and flood distribution map contains relative information integrating many conditions and climatic states. Although soil moisture is a physical quantity that most directly reflects droughts and floods, it does not appropriately reflect the drought and flood characteristics of a certain region unless compared with climate data. For example, although soil moisture is very low in the great majority of southern Xinjiang, it is not defined as an arid region in the climatic drought and flood distribution map. Soil moisture is relatively high in normal years in Sichuan and Chongqing. Therefore, when it decreases significantly, it will be considered a sign of severe drought.

3. Summary and discussion

The establishment of the CLSMDAS, especially the application of high temporal and spatial resolution precipitation data acquired from the Chinese geostationary satellite FY2C, and ground-incident solar radiation data inverted with FY2C visible-light data in CLSMDAS, characterized the spatial and temporal distribution characteristics of atmospheric forcing variables that drive the operation of the land model, improved the simulation precision of the land model, and improved the soil moisture assimilation results. The assimilated high-quality soil moisture grid-point data serve as important basic information for monitoring climate changes, including droughts.

One set of satellite-inverted soil moisture data was assimilated. The inversion error introduced in the soil moisture inversion was large and thus reduced the precision of the soil moisture assimilation results. In future research we will use the surface microwave radiation transfer model to conduct direct assimilation of satellite microwave channel radiation luminance temperature data to improve the precision of soil moisture assimilation. With regard to processing atmospheric driving data, we are currently using multi-source data fusion to interface the numerical model with surface and sounding regular observation data, and data recorded automatically at weather stations. We can then use this system to obtain surface atmospheric temperatures, atmospheric pressure, humidity, and wind speed grid-point data with high precision and high spatial and temporal distribution. By combination with precipitation and ground-incident solar radiation data obtained through geostationary meteorological satellite inversion, we can construct a more reasonable atmospheric driving data set and further improve the precision of land model simulation and assimilation.

References

1. Houser P R. Remote-sensing soil moisture using four-dimensional data assimilation. Doctor Dissertation. Arizona: University of Arizona, 1996
2. Houser P R, Shuttleworth W J, Famiglietti J S, et al. Integration of soil moisture remote sensing and hydrologic modeling using data

- assimilation. *Water Resour Res*, 1998, 34(12): 3405-4320
3. Evensen G. Sequential data assimilation with a non-linear geostrophic model using Monte Carlo methods to forecast error statistics. *J Geophys Res*, 1994, 99(C5): 10143-10162
4. Li Xin, Toshio Koike, Cheng Guodong. A land data assimilation algorithm based on simulated annealing. *Advance in Earth Sciences*, 2003, 18(4): 632-636
5. Huang Chunlin, Li Xin. A summary of researches of land data assimilation systems. *Remote Sensing Technology and Application*, 2004, 19(5): 424-430
6. Yang K, Takahiro Watanabe, Toshio Koike, et al. Auto-calibration system developed to assimilate AMSR-E data into a land surface model for estimating soil moisture and the surface energy budget. *J Meteorol Soc Japan*, 2007, 85A: 229-242
7. Tian X J, Xie Z H, Dai A G. A land surface soil moisture data as-simulation system based on the dual-UKF method and the Community Land Model. *J Geophys Res*, 2008, 113: D14127
8. Tian Xiangjun, Xie Zhenghui. A soil moisture assimilation scheme with the consideration of variability of sub-cells and freeze thawing process of soil. *Science in China, Volume D*: 2008, 38(6): 741-749
9. Tian X J, Xie Z H. A land surface soil moisture data assimilation framework in consideration of the model subgrid-scale heterogeneity and soil water thawing and freezing. *Sci China Ser D-Earth Sci*, 2008, 51(7): 992-1000
10. Tian Xiangjun, Xie Zhenghui, and Dai Aiguo, An Ensemble-based Explicit Four-Dimensional Variational Assimilation Method, *J Geophys Res*, 2008, 113, D21124,
11. Tian X Z, Xie Z H, and Dai A G, et al., A dual-pass variational data assimilation framework for estimating soil moisture profiles from AMSR-E microwave brightness temperature, *J Geophys Res*, 2009, 114, D16102
12. Tian X J, Xie Z H. Effects of sample density on the assimilation performance of an explicit four-dimensional variational data assimilation method. *Sci China Ser D-Earth Sci*, 2009, 52(11): 1849-1856
13. Tian X J, Xie Z H. An explicit four-dimensional variational data assimilation method based on the proper orthogonal decomposition: theoretics and evaluation, *Science in China (D)*, 2009, 52(2), 279-286.
14. Tian X J, Xie Z H, Dai A G. An ensemble conditional nonlinear optimal perturbation approach: formulation and applications to parameter calibration, *Water Resour. Res*, 2010, in press.
15. Zhang Shenglei, Xie Zhenghui, Tian Xiangjun, et. al. A soil moisture assimilation method based on soil water model and station data. *Advance in Earth Sciences*, 2006, 21(12): 1350-1362
16. Zhang Shenglei, Xie Zhenghui, Shi Chunxiang, et. al. Application of collective Kalman filter in soil moisture assimilation. *Atmospheric Sciences*, 2008, 32(6): 1419-1430
17. Jia Binghao, Xie Zhenghui, Tian Xiangjun et. al. A soil moisture assimilation scheme based on microwave luminance temperature and collective Kalman filter. *Science in China, Volume D: Geoscience*, 2009
18. Oleson K W, Dai Y J, Bonan G, et al. Technical description of the Community Land Model (CLM), NCAR/TN-461+STR. 2004
19. Lu Qifeng. Influence of Land Surface Boundary Conditions (Such as LAI) on the Regional Climate Simulation and Preparation of Land Surface Remotely-sensed Parameters for Regional Climate Model (CWRP+CLM). Doctoral Dissertation. Nanjing: Nanjing University of Information Science and Technology, 2006
20. Qian T T, Dai A G, Trenberth K, et al. Simulation of global land surface conditions from 1948 to 2004, part I: Forcing data and evaluations. *J Hydrometeorol*, 2006, 7: 953-975
21. Shi Chunxiang, Xie Zhenghui. A Time Downscaling Scheme of Precipitation by Using Geostationary Meteorological Satellite Data. *Advance in Earth Sciences*, 2008, 27 (4) : 15-22
22. Shi Chunxiang. A study of satellite soil moisture data assimilation based on EnKF algorithm. Doctoral Dissertation. Beijing: Graduate School of Chinese Academy of Science, 2008
23. Stamnes K, Tsay S C, Wiscombe W, et al. Numerically stable algorithm for discrete ordinate method radiative transfer in multiple scattering and emitting layered media. *Applied Optics*, 1988, 27(2): 502~2 509
24. Stuhlmann R and Paschke E, An improvement of the IGMK model to derive total and diffuse solar radiation at the surface from satellite data, *J Appl Meteor*, 1990, 29(7): 596-603
25. Evensen G. The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 2003, 53: 343-367
26. Evensen G. Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dynamics*, 2004, 54: 539-560
27. Njoku E G, Jackson T J, Lakshmi V, et al. Soil moisture retrieval from AMSR-E. *IEEE Trans Geosci Remote Sensing*, 2003, 41(2): 215—229
28. Zhu Changhan, Zhang Qiang. A brief introduction of the drought and flood climate monitoring service system of China. *Meteorological Science and Technology*, 1996, 2: 33-3

General introduction on payloads, ground segment and data application of Fengyun 3A

Peng ZHANG (✉), Jun YANG, Chaohua DONG, Naimeng LU, Zhongdong YANG, Jinmin SHI

National Satellite Meteorological Center, China Meteorological Administration, Beijing 100081, China

© Higher Education Press and Springer-Verlag 2009

Abstract Fengyun 3 series are the second-generation polar-orbiting meteorological satellites of China. The first satellite of Fengyun 3 series, FY-3A, is a research and development satellite with 11 payloads onboard. FY-3A was launched successfully at 11 a.m. on May 27, 2008. Since the launch, FY-3A data have been applied to the services on the flood season and the Beijing 2008 Olympic Games. In this paper, the platform, payloads, and ground segment designs are introduced. Some typical images during the on-orbit commission test are rendered. Improvements of FY-3A on Earth observations are summarized at the end by comparing them with FY-1D, the last satellite of Fengyun 1 series.

Keywords Fengyun 3, payloads, ground segment, data application

1 Historical review of Chinese meteorological satellites

Chinese meteorological satellite activities started in 1969. Since then, two parallel works have been advocated. They are receiving, processing, and utilizing foreign satellites, and programming, developing, and applying Chinese meteorological satellites (Fang et al., 2004).

Chinese meteorological satellites contain two systems (Li, 2001; Meng, 2004). They are the polar orbit series and the geostationary orbit series. Each satellite is named with an Arabic numeral and an alphabet. The Arabic numerals represent the satellite series, the odd being the polar orbit satellites and the even representing the geostationary satellites. The alphabet represents the sequence number within the series. So far, five polar orbit satellites and five geostationary satellites have been launched successfully (seen in Table. 1).

Received January 24, 2009; accepted March 5, 2009

E-mail: zhangp@cma.gov.cn

Table 1 Launched Chinese meteorological satellites

satellite name	satellite type	launch time	satellite function
FY-1A	polar	Sept. 7, 1988	R & D
FY-1B	polar	Setp. 3, 1990	R & D
FY-1C	polar	May 10, 1999	Operation
FY-1D	polar	May 15, 2002	Operation
FY-3A	polar	May 27, 2008	R & D
FY-2A	geostationary	June 10, 1997	R & D
FY-2B	geostationary	June 25, 2000	R & D
FY-2C	geostationary	Oct. 18, 2004	Operation
FY-2D	geostationary	Dec. 8, 2006	Operation
FY-2E	geostationary	Dec. 23, 2008	Backup

2 Specification of the FY-3A platform

Fengyun 3 series are the second-generation polar-orbiting meteorological satellites of China. To meet new and higher requirements in modern meteorological services, especially in numerical weather predictions, these series are designed to perform global, three-dimensional, quantitative, and multi-spectral observations under all weather conditions (i.e., cloud-free and cloudy conditions) with multiple sensors onboard (Fan, 2000; Zhang, 2001).

FY-3A is the first satellite of Fengyun 3 series. It was successfully launched on May 27, 2008 from the Taiyuan launch center. Compared with the single payload of Fengyun 1 series, the number of instruments onboard the satellite has increased to 11. Therefore, FY-3A turns over a new chapter in the history of Chinese meteorological satellites and satellite meteorology (Yang, 2008).

FY-3A is a research and development satellite. The designed lifetime of FY-3A is three years. FY-3A is located at 831 km altitude in the sun-synchronous, near-polar orbit. The attitude control adopted the three-axis stabilization techniques. The major technical specifications of the FY-3A platform are listed in Table 2 (National Satellite

Table 2 The major technical specifications of the FY-3A platform

parameters	technical specifications	
launch mass	2298.5 kg	
dimensions	ground configuration	4380 mm×2000 mm×2000 mm (X.Y.Z)
	in-orbit configuration	4440 mm×10000 mm×3790 mm (X.Y.Z)
orbit	sun-synchronous, near-polar circular	
nominal altitude	831 km	
Inclination	98.81°	
Period	101.49 min.	
Eccentricity	0.00013	
number of revolutions/day	14.17	
orbital interception	2,827.6 km (at equator)	
repeat cycle	about 5 days	
local time at descending node	10:05 (a.m)	
local time drift at descending node	less than 15 min. in 2 years	
payload instruments numbers	11	
data transmission	real time	L band; QPSK modulation; 4.2 Mbps bit rate
		X band; QPSK modulation; 18.7 Mbps bit rate
	delayed transmission	X band; QPSK modulation; 93 Mbps bit rate
attitude	control	3-axis stabilized
	control accuracy	≤0.3° (X,Y,Z)
	measurement accuracy	≤0.05° (X,Y,Z)
	stability	≤0.004°/s (X,Y,Z)
designed life time	3 years	

Meteorological Center, 2004; National Satellite Meteorological Center, 2008a).

3 Introduction of the FY-3A payloads

There are 11 payload instruments mounted on FY-3A. They are the visible and infrared radiometer (VIRR), infrared atmospheric sounder (IRAS), microwave temperature sounder (MWTS), microwave humidity sounder (MWHS), medium resolution spectral imager (MERSI), microwave radiation imager (MWRI), solar backscatter ultraviolet sounder (SBUS), total ozone unit (TOU), earth radiation measurement (ERM), solar irradiance monitor (SIM), and space environment monitor (SEM). Among them, the IRAS, MWTS, and MWHS make up the vertical atmospheric sounding system (VASS). The VIRR is the only instrument among them inherited from the formal FY-1 series platform. Other instruments are all first time in orbit.

The VIRR is a 10-channel VIS/IR radiometer for

multi-purpose imagery with 1.1 km resolution at nadir. The swath of the VIRR is 2800 km. The MERSI is a 20-channel VIS/IR radiometer. There are 19 channels in VIS/NIR/SWIR bands and one in TIR band at 10.0–12.5 μm. Spatial resolution at nadir is twofold: 250 m (for four VIS/NIR channels and one TIR channel) and 1 km (for all other channels). Swath width is similar with VIRR at 2800 km. The MWRI is a 10-channel conical-scanning microwave radiometer at five frequencies. All frequencies are in double polarization. Spatial resolution is 9.5×15 km at 90 GHz and 30×50 km at 19 GHz. The swath of the MWRI is 1400 km. The IRAS is a 26-channel IR radiometer for temperature and humidity sounding. Spatial resolution is 17 km and swath width is 2250 km. The MWTS is a four-channel microwave radiometer for nearly-all-weather temperature sounding with the spatial resolution of 70 km at 54 GHz. The MWTS performs on cross-track scanning mode with swath of 2200 km. The MWHS is a five-channel microwave radiometer at four frequencies (one frequency in double polarization) for nearly-all-weather humidity sounding. The spatial resolution is 15 km at 183 GHz band. The swath width is 2700 km with cross-track scanning. The TOU and SBUS make up a suite of two UV spectro-radiometers. The TOU measures total ozone amount with six channels in the 308–360 nm range, with spatial resolution of 50 km with 3000 km swath. The SBUS measures ozone profile with channels in the range of 252–340 nm. Spatial resolution is 200 km at nadir viewing without side scanning. The ERM is a two-broadband channel radiometer for earth-reflected solar flux and earth-emitted thermal flux over total (0.2–50 mm) and short (0.2–4.3 mm) waveband. The ERM has two working modes. One is cross-track scanning mode with 28 km spatial resolution at 2° narrow field of view (NFOV); the swath width is 2300 km. The other is nadir viewing mode with 120° wide field of view (WFOV). The SIM is a three-channel radiometer over 0.2–50 mm wave band for the total incident solar flux. It views the sun near the north polar area. The SEM is the only *in situ* instrument to measure charged particles in solar wind.

From the view point of data application, the payload instruments on the FY-3A can be analogous to those sensors that are well-used in the world. In fact, the VIRR is the expanded AVHRR instrument, and the MERSI is the MODIS-similar sensor. Both of these optical imagers can provide surface characteristics (including cloud surface, land surface, and ocean surface) and aerosol information. The VASS set is made up of ATOVS-similar instruments for atmospheric sounding. The MWRI is the AMSR-similar instrument except for the low frequency at 6.9 GHz. The SBUS and TOU are the SBUV-similar and TOMS-similar sensors, respectively, which can provide ozone profile and total ozone amount separately. The ERM is the CERES-similar instrument inherited from ERBE. The detailed specifications of these 11 payload instruments and their data applications are listed in Table 3 (National

Table 3 The specification of FY-3A 11 payloads

Name	Specification	Purpose
VIRR	spectral range	0.43–12.5 μm
	number of channels	10
	scan range	$\pm 55.4^\circ$
	spatial resolution at nadir	1.1 km
	Cal accuracy in VNIR	5%–10%
	Cal Accuracy in IR	1 K (270 K)
IRAS	quantization	10 bits
	spectral range	0.69–15.0 μm
	number of channels	26
	scan range	$\pm 49.5^\circ$
	spatial resolution at nadir	17 km
	Cal accuracy in VNIR	5%–9%
MWTS	Cal accuracy in IR	1 K (270 K)
	quantization	13 bits
	spectral range	50–57 GHz
	number of channels	4
	scan range	$\pm 48.3^\circ$
	spatial resolution at nadir	50–75 km
MWHS	Cal accuracy	1.2 k
	sensitivity (NE Δ N)	0.4–0.55 k
	quantization	13 bits
	spectral range	150–183 GHz
	number of channels	5
	scan range	$\pm 53.35^\circ$
MERSI	spatial resolution at nadir	15 km
	Cal accuracy	1.5 k
	sensitivity (NE Δ N)	1.1–1.2 k
	quantization	14 bits
	spectral range	0.40–12.5 μm
	number of channels	20
MWRI	scan range	$\pm 55.4^\circ$
	spatial resolution at nadir	0.25–1 km
	Cal accuracy in VNIR	5%–10%
	Cal accuracy in IR	1 K (270 K)
	quantization	12 bits
	spectral range	10–89 GHz
SBUS	number of channels	10
	scan range	$\pm 55.4^\circ$
	spatial resolution at nadir	15–85 km
	Cal accuracy	1–2.8 K
	quantization	12 bits
	spectral range	0.16–0.4 μm
TOU	number of channels	12
	spatial resolution at nadir	200 km
	quantization	16 bits
	stray light	10^{-6}
	Cal accuracy	3% (160–250 nm) 2% (250–400 nm)
	Cal accuracy in diffuse reflection board	3%
TOU	spectral range	0.3–0.36 μm
	number of channels	6
	scan range	$\pm 54^\circ$
	spatial resolution at nadir	50 km
	quantization	10^{-3}
	stray light	12 bits
TOU	Cal accuracy	2%

(Continued)

Name	Specification	Purpose
ERM	spectral total band	0.2–50 μm
	range solar band	0.2–3.8 μm
	number of narrow FOV	2
	channels broad FOV	2
	scan range	$\pm 50^\circ$ (narrow)
	sensitivity (NE Δ N)	$0.4 \text{ Wm}^{-2} \cdot \text{sr}^{-1}$
	Cal total band	0.8%
	accuracy solar band	1%
SIM	stability within 2 years	< 1%
	spectral range	0.2–50 μm
	sensitivity (NE Δ N)	0.2 Wm^{-2}
	Cal accuracy	0.5%
	quantization	16 bits
SEM	stability within 2 years	< 0.02%
	heavy ions, high energy proton, high energy electron, radiation dose, satellite surface charging monitoring, single event upset	space environment

Satellite Meteorological Center, 2004; National Satellite Meteorological Center, 2008a).

4 FY-3A ground segment design

The Chinese meteorological satellite engineering system comprises five parts: satellite segment, launch pad segment, launch vehicle segment, measurement and control segment, and ground segment (Li, 2008). The ground segment responds to data receiving, data pre-processing, and data processing. It is the key to sustaining and promoting satellite data application in the meteorological services (Xu et al., 2006).

The FY-3A ground segment includes the data processing

center, operation and control center, ground receiving center, and data archiving center. They are composed of 10 technical systems and one airborne-based field experiment (National Satellite Meteorological Center, 2006). The technical systems are the Data Acquisition System (DAS), Computer and Network System (CNS), Operation Control System (OCS), Data Pre-Processing System (DPPS), Products Generation System (PGS), Quality Control System (QCS), Utilization Demonstration System (UDS), Archive and Service System (ARSS), Monitoring and Analysis System (MAS), and Simulation and Technical Supporting System (STSS). The purpose of the airborne-based field experiment is to test the engineering model of the payload instruments. The FY-3A ground segment framework is shown in Fig. 1.

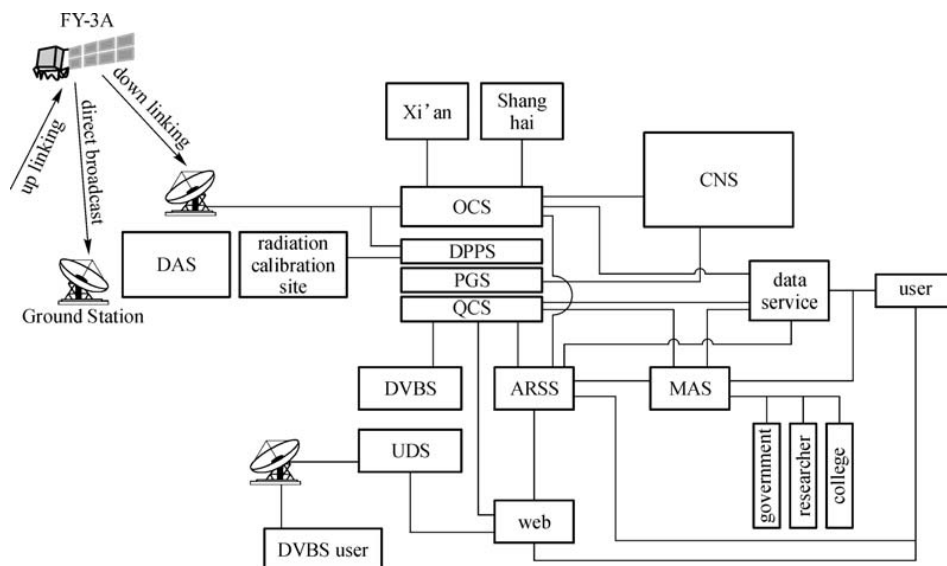


Fig. 1 FY-3A ground segment framework

The FY-3A ground segment has five ground stations to receive the satellite direct-broadcasting data. Four of them are inside China and the last one is located near the northern polar region. Such distribution, especially the ground station in the polar region, guarantees that global data can be received and collected within about three hours. The ground stations receive all of the L-band HRPT data, X-band MPT data, and DPT data. Table 4 shows the information of current ground stations built for FY-3A.

Table 4 Ground station distribution

station name	longitude	latitude
Beijing Station	116° 16' 36" E	40° 03' 06" N
Guangzhou Station	113° 20' 20" E	23° 09' 52" N
Urumchi Station	87° 34' 08" E	43° 52' 17" N
Jiamusi Station	130° 22' 48" E	46° 45' 20" N
Kiruna Station	21° 02' E	67° 32' N

5 Data application and demonstration

There are five systems in the FY-3A ground segment related directly with data application. They are the DPPS, PGS, ARSS, MAS, and UDS. Thereinto, the DPPS generates level 1 products with geolocation and calibration information, and the PGS produces level 2 products to provide geophysical and geochemical information through retrieval algorithms. The ARSS responds to data archiving and data distribution. The MAS provides users a special toolkit to analyze satellite data. The UDS responds to promote the utilization and demonstration of FY-3A data into meteorological services. Figure 2 shows the relationship among the DPPS, PGS, ARSS, MAS, and UDS. Table 2 lists the level 2 products generated from the PGS currently covering the atmosphere, land, ocean, cryosphere, radiation, and space environment (National Satellite Meteorological Center, 2008b).

Since the launch, FY-3A data have been applied to global weather system monitoring, typhoon monitoring,

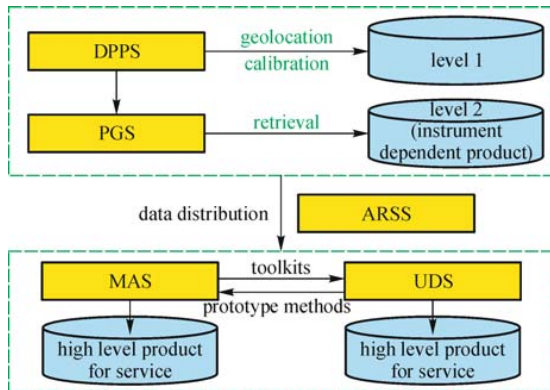


Fig. 2 FY-3A ground segment related with data application

Table 5 Level 2 products generated from PGS

discipline	parameters	FY-3A sensors
atmosphere	cloud/fog properties	MERSI/VIRR/IRAS/
	total water vapor	MWRI/MWHS
	precipitation	VIRR/MERSI/MWRI
	aerosol properties	MWHS/MWRI
	atmospheric temperature and humidity	MERSI/VIRR
	total ozone and ozone profile	IRAS/MWTS/MWHS TOU/SBUS/IRAS
land	land cover	VIRR/MERSI
	surface temperature	VIRR
	vegetation dynamics	VIRR/MERSI
	fire and flood monitoring	VIRR/MERSI
	surface wetness	MWRI
ocean	sea surface temperature	VIRR/MWRI
	sea surface color	MERSI
cryosphere	sea ice	MERSI/VIRR/MWRI
	snow cover	VIRR/MWHS/MERSI/ MWRI
radiation	earth's radiation and solar irradiance	ERM/SIM
space environment	high energy particles radiation dose	SEM

sea and inland water body monitoring, fire monitoring, aerosol and air quality monitoring, the monitoring of hot-island effect in cities, global sea shelf monitoring, global ozone monitoring, etc. It is noted that FY-3A is still on its on-orbit commission test phase. However, FY-3A has served the Beijing 2008 Olympic Games and the flood season in 2008 at the same time.

Figure 3 is the global mosaic image in MERSI channels 3, 2, 1 with 250 m resolution acquired on the same day. The intertropical convergence zone (ITCZ), tropical depression, and subtropical high pressure can be seen clearly. As indicated in Fig. 4, when typhoon Fung-wong was monitored by MERSI on 27 July, 2008, the typhoon hole inside and the spiral cloud outside can be distinguished. Figure 5 shows the ozone hole monitored from

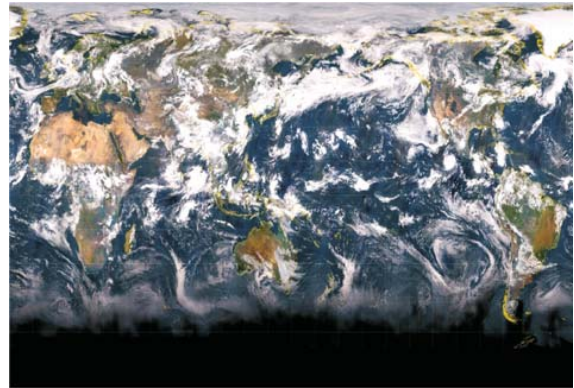


Fig. 3 Global mosaic image in MERSI channel 3, 2, 1 with 250 m resolution



Fig. 4 Typhoon Fung-wong monitored by MERSI on 27 July, 2008

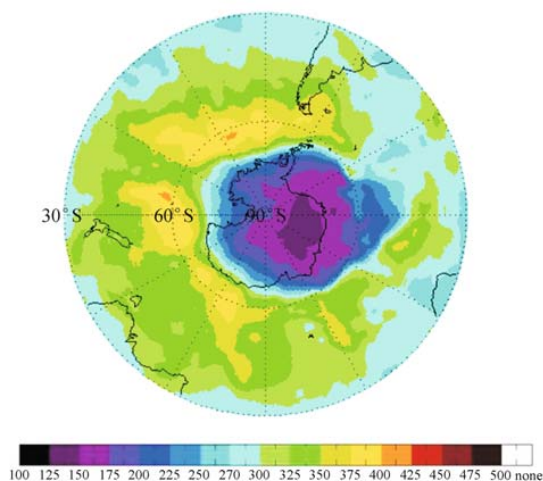


Fig. 5 Total ozone in DU monitored from TOU on Nov. 1, 2008 (Courtesy of Weihe Wang, NSMC)

TOU on November 1, 2008. The region on the cold color side corresponds to the ozone hole in the image.

6 Conclusions

In comparison with FY-1 series, the principal improvements in FY-3A include: 1) atmospheric sounding capacity, 2) microwave imaging capacity, 3) optical imaging with spatial resolution from 1 km to 250 m, 4) atmospheric composition detecting capacity, 5) radiation budget measuring capacity, and 6) global data acquisition from within one day to within two to three hours.

The number of instruments on board the satellite has increased to 11. FY-3A has evolved from single imaging to

comprehensive earth environment observations, from optical to microwave remote sensing, with resolution having been increased from kilometer to hectometre-category, and receivable both in China and up to the polar region. Therefore, FY-3A turns over a new chapter in the history of the Chinese meteorological satellites and satellite meteorology.

The FY-3A satellite provides global air temperature, humidity profiles, and meteorological parameters such as cloud and surface radiation required in producing weather forecasts, especially in making medium numerical forecasting. The FY-3A satellite monitors large-scale meteorological disasters, weather-induced secondary natural hazards and environment changes, and provides geophysical parameters for scientific research in climate change and its variability, climate diagnosis, and predictions. The FY-3A satellite renders global and regional meteorological information for aviation, ocean navigation, agriculture, forestry, marine activities, hydrology, and many other economic sectors.

Acknowledgements This research was supported by the National Key Project of Scientific and Technical Supporting Programs Funded by Ministry of Science & Technology of China (No. 2008BAC40B01). The authors appreciate the contributions from the colleagues who worked for the FY-3A engineering programme in the National Satellite Meteorological Center.

References

- Fan T X (2000). The roles of Fengyun 3 meteorological satellites. Technologic Note of Satellite Meteorology, National Satellite Meteorological Center, 162 (in Chinese)
- Fang Z Y, Xu J M, Zhao F S (2004). Progress and prospect of China meteorological satellite and satellite meteorology. ACTA Meteorological Sinica, 62(5): 550–560 (in Chinese with English abstract)
- Li Q (2001). Development of Chinese geostationary meteorological satellite. Spacecraft recovery and remote sensing, 22(1): 13–19 (in Chinese with English abstract)
- Li Q, Dong Y H (2008). Achievement and forecast of meteorological satellite technology in china. Aerospace Shanghai, 1: 1–10 (in Chinese with English abstract)
- Meng Z H (2004). The polar orbit meteorological satellite in China. Engineering Science, 6(10): 1–5 (in Chinese with English abstract)
- National Satellite Meteorological Center (2004). User requirement of Fengyun 3 series (in Chinese)
- National Satellite Meteorological Center (2006). Initiative design on the first engineering stage of Fengyun 3 series meteorological satellite application system (in Chinese)
- National Satellite Meteorological Center (2008a). FY-3: a new era of Chinese meteorological satellite. FY-3A Introduction Brochure, 1–27 (in Chinese)
- National Satellite Meteorological Center (2008b). Brief Catechism of FY-3A Polar Orbital Meteorological Satellite, 1–64 (in Chinese)
- Xu J M, Niu Y S, Dong C H, Zhang W J, Yang J (2006). Ground segments for FY meteorological satellites. Engineering Science, 8 (11): 13–18 (in Chinese with English abstract)
- Yang J (2008). Development and applications of China's Fengyun (FY) meteorological satellite. Spacecraft engineering, 17(3): 23–28 (in Chinese with English abstract)
- Zhang W J (2001). Status and development of FY series of meteorological satellites. Aerospace Shanghai, 2: 8–12 (in Chinese with English abstract)

NASA Satellite and Model Land Data Services: Data Access Tutorial

Suhung Shen^{1,2}, Gregory Leptoukh¹, Hongliang Fang³

¹ NASA Goddard Earth Sciences Data and Information Services Center, Greenbelt, MD, USA, Suhung.shen@nasa.gov

² George Mason University, Fairfax, VA, USA

³ Institute of Geographic Sciences and Natural Resources Research (IGSNRR), Chinese Academy of Sciences, Beijing, China

Abstract: The NASA Earth Observing System (EOS) program was started in early 1990s, focusing on observation of measurements for climate studies, such as land surface temperature, precipitation, greenhouse gases, aerosols, vegetation index, and soil moisture, etc. Data acquired from many satellites are archived and distributed at different NASA data centers. This document provides brief information about the land data products from NASA satellite measurement missions as well as land surface models. The data search and order methods, data tools, as well as sample applications are documented.

1. Introduction of NASA land products

1.1 NASA satellite observation missions on land surface

The land surface satellite observations begun with the launch of the first in a series of Landsat satellites developed by National Aeronautics and Space Administration (NASA) and operated by U.S. Geological Survey (USGS) (<http://landsat.gsfc.nasa.gov/>). Landsat 1 was launched on July 23, 1972, known as the Earth Resources Technology Satellite (ERTS). Landsat 1 operated until January 1978, outliving its design life by five years. Six Landsat satellites were launched following Landsat 1. Landsat 5 and Landsat 7 are in operation. About 38 years data were collected from Landsat. The next generation of land observation satellite, Landsat Data Continuity Mission (LDCM), is scheduled to launch in December 2012.

The NASA Earth Observing System (EOS) program started to acquire data in early 1990s, focusing on climate science areas: radiation, clouds, water vapor, and precipitation; the oceans; greenhouse gases; land-surface hydrology and ecosystem processes; glaciers, sea ice, and ice sheets; ozone and stratospheric chemistry; and natural and anthropogenic aerosols. Detailed information about the NASA EOS program is documented in the “Earth Science Reference Handbook” by Parkinson et al (2006) and can be found at the web site <http://eospsso.gsfc.nasa.gov/>. More than twenty satellites were launched since 1991. Land surface products are obtained from a number of instruments, such as ETM on Landsat 7 (1999), MODIS on Terra (2000) and Aqua (2002), ASTER on Terra (2000), and SeaWiFS on Obview-2 (1997). Other measurements for land surface study in EOS program include precipitation from TRMM (1997), soil moisture and snow/ice water from AMSR-E (Aqua, 2002), etc.

1.2 NASA Satellite land products, processing levels, resolutions, and data format

Most NASA satellite land measurements are summarized in the NASA Land Measurement Portal (<http://landportal.gsfc.nasa.gov/>). The portal lists land products in four categories: Surface radiation Budget, Vegetation

Parameters, Land-Cover/land Use Changes, and Hydrosphere. The portal provides product metadata, including instrument name, spatial and temporal resolution, data archiving location, and data access policy. Example products are:

- Land Surface temperature/Emissivity from MODIS at 1km, 6km, or 0.05 degree for daily, 8-day, and monthly
- Vegetation index from MODIS at 250m, 500m, 1km, or 0.05 degree for 16-day, and monthly
- Land cover types from MODIS at 250m or 500m for annual; from Landsat 1-7 at 30m for decadal
- Thermal Anomalies/Fire from MODIS at 500m or 1km for daily, 8-day, and monthly
- Snow cover from MODIS at 500m and 0.05 degree for daily, and 8-day
- Evapotranspiration from MODIS at 0.05 degree for 8-day and annual

The standard NASA satellite products have four processing levels as listed in Table 1.

Table 1: Satellite data processing Levels

Levels	Description	Data type
Level-0	Source data, the raw radiance counts at full resolution	swath
Level-1A	The raw radiance counts with time and geolocation referred	Swath
Level-1B	Calibrated radiance, or processed to sensor units, such as brightness temperature at full resolution	swath
Level-2	Derived geophysical variables (e.g, land surface temperature, vegetation index, land cover type) after performing atmospheric correction at the same resolution and location of Level-1 data.	Swath
Level -3	Geophysical variables mapped on uniform space-time grid scales at reduce resolution	Grid
Level -4	Geophysical variables derived from lower level data (i.e., variables are not measured by the instrument, but instead are derived from these measurements by an analysis model)	Grid

The typical data formats of NASA EOS satellite data are HDF (Hierarchical Data Format, <http://www.hdfgroup.org/>), HDF-EOS (Hierarchical Data Format - Earth Observing System, <http://hdfeos.org/>, <http://www.hdfgroup.org/hdfeos.html>), or NetCDF (Network Common data Form, <http://www.unidata.ucar.edu/software/netcdf/>). All these data formats are self-describing, enable an application to interpret the data structure and contents without any outside information, and are platform transparent. Commercial and free tools are available to support above data formats.

1.3 NASA Land surface model data: GLDAS

The Global Land Data Assimilation System (GLDAS) is a land surface data assimilation system developed by the Hydrological Science Branch at NASA Goddard. The system ingests satellite and ground-based observational data products, using advanced land surface modeling and data assimilation techniques, for generating optimal fields of land surface states (e.g., soil moisture and surface temperature) and fluxes (e.g., evaporation and sensible heat flux) parameters (Rodell et al., 2004). The current GLDAS includes data from four land surface models: Mosaic, Noah, the Community Land Model (CLM), and the Variable Infiltration Capacity (VIC). More details about GLDAS models can be found at GLDAS Readme (<ftp://agdisc.gsfc.nasa.gov/data/s4pa/GLDAS/README.GLDAS.pdf>) and NASA Land Data Assimilation System website (<http://ldas.gsfc.nasa.gov/>).

The temporal resolution for the GLDAS products is 3-hourly. Monthly products are also generated through temporal averaging of the 3-hourly products. Data is available from January 1 1979 to present. The spatial resolutions are 0.25 and

1 degree. The data are archived and distributed by the NASA Goddard Earth Science Data and Information Services Center (GES DISC) (<http://disc.sci.gsfc.nasa.gov/>) through ftp and Mirado (<http://mirador.gsfc.nasa.gov/>). The GLDAS data are also provided to GrADS Data Server (GDS) users via <http://agdisc.gsfc.nasa.gov/dods/>. GDS allows a user to access the data, perform subsetting and analysis operations without first downloading them. More advanced online visualization and analysis tool (Giovanni) (<http://disc.sci.gsfc.nasa.gov/giovanni/>, Berrick, et al. 2009) is also available to GLDAS data, which provides a simple and intuitive way to visualize, analyze, and access vast amounts of data without having to download the data.

2. Order NASA Earth Science data products

2.1 NASA Earth science data centers

NASA Earth sciences data are archived and distributed by twelve Distributed Active Archive Centers (DAACs) at different locations. Each center serves a specific Earth system science discipline and provides users with data products, services, and data-handling tools unique to the center's specialty (<http://nasadaacs.eos.nasa.gov/about.html>). The data product can be found through the centralized searching and ordering system such as WIST/ECHO or from a specific data center where data are archived, such as LP DAAC, GES DAAC, NSIDS, and OBPG, etc.. A brief summary of these DAACs are listed in the section 2.4. The following sections give examples to search and order data from centralized services and from data center GES DISC.

2.2 Find and order data from the centralized system

2.2.1 Order data from WIST

<http://wist.echo.nasa.gov>

Wherehouse Inventory Search Tool (WIST) is a primary access point of more than two thousands NASA EOS and other Earth science data sets. WIST system requires a user to register. You may browse data products without login. Here is an example of steps to order MODIS monthly land surface temperature:

- a) Type “MODIS” in Text Search, click go. It will display all MODIS products in the WIST database; or select “MODIS/Aqua” under Category “Land”, which will list all land products of MODIS/Aqua
- b) Select one or more data set, say “MODIS/Aqua Land Surface Temperature/Emissivity Monthly L3 Global 0.05 Deg CGM V005”
- c) Select search area by click, hold, and drag on the map, or enter four corners. The default area is global. In this case, the default option is selected.
- d) Give temporal range as “2009-01-01” to “2009-12-31”
- e) Use default options for “Additional Options”
- f) Click “Start Search”
- g) Select item (granules) and click on “Add selections to cart”
- h) Click on “Accept-continue to Shopping Cart”
- i) Specify ordering options
- j) Fill the ordering form

- k) Review and submit order
- l) Will receive an e-mail in a minutes that contains data downloading instruction

2.2.2 Find data from ECHO

<https://www.echo.nasa.gov/>

NASA EOS Clearinghouse (ECHO) is a warehouse database to archive EOS product metadata from a variety of science disciplines and domains, including Climate Variability and Change, Carbon Cycle and Ecosystems, Earth Surface and Interior, Atmospheric Composition, Weather, and Water and Energy Cycle. All products in WIST system are in ECHO database. Products in ECHO can be navigated by topics, instrument, campaign, data center, etc. On the search or navigation result page, product short description, temporal coverage, and the link to ordering page are provided. ECHO has been working with other organizations to provide their Earth science metadata alongside NASA's for users to search and access.

2.2.3 Find data from GCMD

<http://gcmd.gsfc.nasa.gov>

The Global Change Master Directory (GCMD) is a comprehensive source of satellite and in situ Earth science data, with broad coverage of the atmosphere, hydrosphere, oceans, solid Earth, and biosphere. The system provides Key-word search and navigation to find a data product. For example, on page <http://gcmd.gsfc.nasa.gov>:

- Click on LAND SURFACE → LAND TEMPERATURE → LAND SURFACE, it lists 253 items, refine by text “MODIS”, it reduces items to 63. (Note, the searched items subject to change since the system is updated actively).
- Click “MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V005”, it displays a page with metadata information of this product, including product description, geographic coverage, spatial resolution, instrument/platform, data source link, etc.

2.3 Find data from data centers and services:

EOS Land products are archive and distributed at the following DAACS:

- LP DAAC (<https://lpdaac.usgs.gov/>)

LP DAAC processes, archives, and distributes land data and products derived from the EOS sensors. LP DAAC holds land data from MODIS and ASTER on Terra, and MODIS on Aqua. MODIS data are Level 2 (swath) and Level 3 (grid) in HDF-EOS format. The resolutions range of the Level 3 products vary from 250 m to 6000 m. The ASTER data are Level 1 (swath), Level 2 (swath), and Level 3 (grid) in HDF-EOS or GeoTIFF format. The Level 3 data with resolution 1 km or higher are stored in 10x10 degree tiled files.

- GES DISC (<http://disc.gsfc.nasa.gov/>)

GES DISC archives and distributes satellite data of multiple disciplines, including hydrology, atmospheric composition and dynamic, as well as land surface process. GES DISC archives and distributes data from land and atmospheric assimilation model. The existing satellite land products are global 1x1 degree resolution from MODIS. Higher resolution (1 km) land data from MODIS will be available for entire Asia region, which can be visualized through the online system Giovanni soon. NASA land surface model data are in GES DISC.

- NSIDC (<http://nsidc.org/daac/>)

The National Snow and Ice Data Center (NSIDC) is part of the Cooperative Institute for Research in Environmental Sciences at the University of Colorado at Boulder. NSIDC provides cryosphere-related data and information, including

snow, ice, glaciers, frozen ground, and climate interactions from field and satellite observation. Example satellite land data are soil moisture, and snow water equivalent from AMSR-E.

- OBPG (<http://oceancolor.gsfc.nasa.gov/>):

The Ocean Biology Processing Group (OBPG) at NASA GSFC process and archive ocean biology data from sensors, such as SeaWiFS, MODIS, OCTS, and MERIS etc. Land vegetation index (NDVI) at 4km and 9km resolution from SeaWiFS are produced and distributed by OBPG as additional products.

Landsat data can be obtained from:

- USGS GioVis, The Global Visualization Viewer (<http://glovis.usgs.gov/>): A web-based tool to search and order data by clicking on an interactive map or by entering the geographic coordinates (latitude and longitude) of the site you are searching for.
- USGS earth Explorer (<http://edcsns17.cr.usgs.gov/EarthExplorer/>): A Web-based tool that allows you to custom tailor your search parameters for Landsat data.

Land data are also available from:

- GLCF - Global Land Cover Facility at University of Maryland
<http://glcf.umiaccs.umd.edu/index.shtml>. This site provides satellite imagery of Land cover and other land products from ASTER, IKONOS, Landsat, MODIS, QickBird, Orbview, and SRTM, as well as products from AVHRR, GOES, Landsat, and MODIS, etc.
- LEDAPS – Landsat Ecosystem Disturbance Adaptive Processing System
<http://ledaps.nascom.nasa.gov/>. This site is a NASA-funded project to map North American forest disturbance since 1975 from the Landsat and ASTER satellite data.
- Land Measurement Portal
<http://landportal.gsfc.nasa.gov/>. This portal provides comprehensive metadata of land measurement from NASA EOS, NOAA and other Agencies and Institutions.

2.4 Order data from GES DISC

The NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) is one of twelve NASA Earth Science data archive and distribution centers. GES DISC holds data and information of Precipitation, Atmospheric Chemistry and Dynamics, and land surface process from satellite observation and atmospheric and land assimilation models. The center provides several methods to access data sets, including Mirado, OPeNDAP, GDS, OGC Web Map Service, and Giovanni. All data from GES DISC are open to the public. No registration is required to access the data.

Mirado (<http://mirador.gsfc.nasa.gov/>) is a major data search and order system at GES DISC. Mirado has a very simple and easy to use interface that can perform Google-like keyword search. Features such as data file hit estimator, an interactive shopping cart, and spatial and parameter subsetting (subsetting feature is available for some products) are available in Mirado. In addition, the system can search data through navigation of projects and science area by applying semantic web technology. A list of product in Mirado is available by clicking “Data Holding” from the Mirado page. The following is an example for ordering soil moisture from GLDAS land surface model:

- From page: <http://mirador.gsfc.nasa.gov/>
- **Using Keyword Search:**
 - a) Keyword: soil moisture

Location [(minLat, minLon), (maxLat, maxLon)]: (15,55), (55,135) over China

Time Span: 2000-01-01, 2000-01-31

It returns a list of all products with soil moisture. The list contains product information, file numbers, parameters, spatial & temporal resolution.

If the list is too long, and you would like to narrow down it, go back to “keyword” search page, and add more keyword. For example, add model name in Keyword as “GLDAS CLM soil moisture”

b) select the product, say “GLDAS CLM Land Surface Model L4 Monthly 1.0 x 1.0 degree (GLDAS_CLM10_M)”, then click on “Add selected files to cart”

c) select service by clicking available services, say “convert to NetCDF”

d) click “Continue to shopping cart”

e) click “check out” (you may modify the cart at this time)

f) Follows download instruction to download data. There are multiple download methods. You may select one that works best for you.

- **Using Projects Navigation:**

There are two types of projects, one is satellite mission, such as AIRS, OMI, TRMM, the other is research project, such as A-Train, NEESPI, GLDAS.

In this case, navigate in the following path to get the above same data:

Projects → GLDAS → Common Land Model (CLM) → GLDAS_CLM10_M.001 → 2000

Then do the steps c) to f) in Keyword.

- **Using Science Area Navigation:**

This is for navigation parameters starting with science areas, for example:

Science Areas → Water and Energy Cycles → Soil → Average Layer Soil Moisture → GLDAS_CLM10_M.001 → 2000. Then do the steps c) to f) in Keyword.

Some data are available through OPeNDAP and GDS, which are software framework that allows to access remote data. Data in OPeNDAP can be accessed directly through many tools, such as IDV, Panoply, Ferret, and GrADS, etc. GrADS Data Service (GDS) provide subsetting and analysis services across the internet for any GrADS-readable dataset. Sample data sets available through GDS are land surface model data (GLDAS), and atmospheric reanalysis data (MERRA). The following are examples to access GLDAS data in GDS service by using GrDAS.

- **Perform global averaging analysis for soil moisture content**

```
$ grads
```

```
ga-> sdfopen http://agdisc.gsfc.nasa.gov/dods/\_expr\_{GLDAS\_MOS10SUBP\_3H}\_{aave\(soilm3.1,lon=-180,lon=180,lat=-60,lat=90\)}{0:360,-60:90,1:1,00z22jan1980:21z22jan1980}
```

```
ga-> d
```

```
Result value = 443.626 mm
```

- **Sample GrADS script to access GLDAS data via GDS**

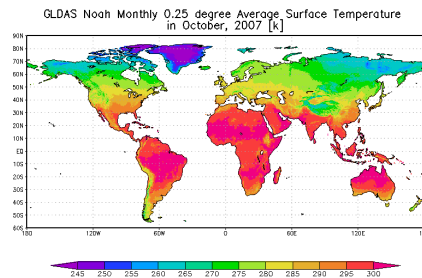
Save the following script into a file named: map_gldas.gs

```
'reinit', 'sdfopen http://agdisc.gsfc.nasa.gov/dods/GLDAS\_NOAH025\_M', 'set lon -180 180', 'set lat -60 90', 'set gxout grfill', 'set grads off', 'set t 92', 'd avgsurft', 'set rbcsls', 'run cbarn', 'draw title GLDAS Noah Monthly 0.25
```

degree Average Surface Temperature \ in October, 2007 [k]', 'printim
GLDAS_NOAH025_M_avgsurft.A200710.001.gif white'

Run script and generate the following map:

\$grads, ga->map_gldas



Open Geospatial Consortium (OGC) Web Map Service (WMS) is an interface that allows the use of data and enables clients to build customized maps with data coming from a different network. Currently, data in OGC WMS at GES DISC are these from AIRS, TRMM, and OMI.

Many Level 3 and Level 2 satellite and model data are available through Giovanni, which is an online visualization and analysis tool. More will be described in the section 3.1.

3. Online Visualization Services of NASA data

A number of online visualization services are available for the NASA data. Five popularly used services will be described here.

3.1 Giovanni

<http://disc.sci.gsfc.nasa.gov/giovanni/>

Giovanni (GES-DISC (Goddard Earth Sciences Data and Information Services Center) Interactive Online Visualization AND aNalysis Infrastructure) is a Web-based application developed by GES DISC, which can visualize and analysis data online by a few clicks without downloading the data. The Giovanni database holds more than a thousand of geophysical parameters. Visualization types include maps, time series, cross-section, profile, scatter plots, and correlation maps, etc. The Giovanni system generates images in formats as png and KMZ for Google Earth. The processed data can be downloaded in ASCII, hdf, and NetCDF formats. To use Giovanni, only internet and a browser are needed, no need to install data processing software or downloading data. The current Giovanni system consists of more than thirty instances (interfaces) for different collections of data. Sample Giovanni instances are:

Land surface model GLDAS in Giovanni:

http://gdata1.sci.gsfc.nasa.gov/daac-bin/G3/gui.cgi?instance_id=GLDAS10_M

Data collection for MAIRS:

http://gdata1.sci.gsfc.nasa.gov/daac-bin/G3/gui.cgi?instance_id=mairs_monthly

Note: It is requested that the GES DISC be specifically and clearly acknowledged if Giovanni (or data downloaded from Giovanni) is used for data analyses and visualizations in publications, posters, oral presentations, reports, Web pages, and other types of scientific media. Please check Giovanni page to find sample acknowledge statements.

Satellite Land Measurements in Giovanni include:

- Precipitation:
TRMM (1997.12 – present), GPCP (NASA Global precipitation climate project, 1979.01 – present)
- Wind:
QuikSCAT (1999.06 – present)
- Land surface temperature:
MODIS/Terra (2000.03 – present), MODIS/Aqua (2002.07 – present)
- Vegetation index:
MODIS/Terra (2000.03 – present), MODIS/Aqua (2002.07 – present)
- Soil moisture:
AMSR-E/Aqua (2002.07 – present)
- Active fire:
MODIS/Terra (2000.03 – present), MODIS/Aqua (2002.07 – present)
- Snow/ice (Northern Hemisphere):
NOAA/NESDIS (2000.01 – present)

Assimilation Model Data in Giovanni:

- GLDAS (Global Land Data Assimilation System), 1979-present, 1x1 degree, monthly,
Contains land surface state (e.g., soil moisture and surface temperature) and flux (e.g., evaporation and sensible heat flux) products simulated by four land surface models (CLM, Mosaic, Noah and VIC). Please read GLDAS Readme for details: <ftp://agdisc.gsfc.nasa.gov/data/s4pa/GLDAS/README.GLDAS.pdf>
- MERRA (MODERN ERA RETROSPECTIVE-ANALYSIS FOR RESEARCH AND APPLICATIONS), 1979-present, 2/3x1/2 degree, monthly, Contains land surface diagnostic parameters like transpiration, vegetation greenness fraction, leave area index, top soil layer wetness, surface temperature of saturated (unsaturated, wilted) zone, etc. Please read MERRA file specific document for details: http://gmao.gsfc.nasa.gov/research/merra/file_specifications.php

A hands on sample cases for using Giovanni is available at:

http://disc.sci.gsfc.nasa.gov/mdisc/documentation/demo_cases.doc

3.2 MODIS Rapid Response System

<http://rapidfire.sci.gsfc.nasa.gov/>

MODIS Rapid Response System provides high quality real-time and past-time true-color and false-color imagery to support monitoring and study fire, dust storm, and other natural hazards. Images are at resolutions of 1 km, 500 m or 250 m and in formats of gif, KMZ for GoogleEarth, and Worldfile for GIS. Links are provided to data used to generate MODIS images.

3.3 NASA Earth Observations (NEO)

<http://neo.sci.gsfc.nasa.gov/Search.html>

This tool is designed for outreach and education purpose. A user can explore remote sensing products from an easy use Web interface by generating images on-the-fly. The generated image can be mapped onto Google Earth.

3.4 NASA Earth observatory

<http://earthobservatory.nasa.gov/>

Earth Observatory is another outreach and education site of NASA. In addition to pre-generated high quality, a story/article is written about the image. Images and stories about natural hazards (fires, dust storms, floods, and droughts, Volcanoes and earthquakes) are documented. You may sign up to receive daily or weekly electronic news letters to keep track of recent natural hazards stories. Global maps and articles of precipitation, surface temperature anomaly, snow cover, vegetation, etc. are also available.

3.5 NASA Visible Earth

<http://visibleearth.nasa.gov>

This is a huge archive of high quality images and animations from many instruments of NASA, NOAA, DOD, and non-US agencies. Images can be searched by key word, satellite, sensor, country, collection, and GCMD topics, etc, and be downloaded freely.

4. Support Research Projects and sample use of data and service

This section introduce a number projects at GES DISC to support researches on land cover and land use changes and climate variations by providing data and services. Sample plots from satellite observation and model data by using Giovanni are presented.

4.1 NASA data to support research projects: NEESPI and MAIRS

The NASA NEESPI and MAIRS data centers at GES DISC are funded by the NASA land cover and land use change program to support international programs, Northern Eurasia Earth science Partnership Initiative (NEESPI) (<http://neespi.org/>) and Monsoon Asia Integrated Regional Study (MAIRS) (<http://www.mairs-essp.org/>), by providing satellite remote sensing and model data and information. The NEESPI and MAIRS program focus on studies of land processing and climate variations at Northern Eurasia and Monsoon Asian regions, respectively.

The NEESPI data support project (<http://disc.sci.gsfc.nasa.gov/neespi/>) was started in 2006 (Leptoukh, et al, 2007). Infrastructure of an automated data management system was established to support the data end-to-end, including data transfer, data format convert, and data ingest into archive and distribute system. Tools are provided to read data and analysis data. Products collected and processed are land cover types, land surface temperature, soil moisture, vegetation index, active fire, and snow/ice. The spatial resolution is 1x1 degree. Customized Giovanni interfaces for monthly and daily data have been created to allow exploring easily data from multiple instruments of multiple disciplines (<http://disc.sci.gsfc.nasa.gov/neespi/additional/visualization.shtml>)

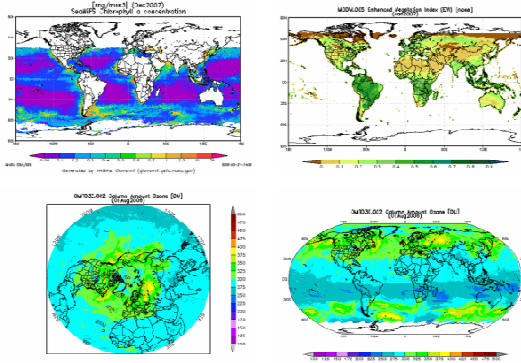
Following successful support of the NEESPI program with NASA satellite remote sensing data, the MAIRS data support project (<http://disc.sci.gsfc.nasa.gov/mairs/>) was started in 2009. It uses and leverages the established data management and service infrastructure for NEESPI. A customized Giovanni system has been created for MAIRS that contains satellite observations from multiple sensors and model output from the NASA Global Land Data Assimilation System (GLDAS), and from the NASA atmospheric reanalysis project, MERRA. Higher resolution satellite data, such as 5km and 1km land surface temperature, and vegetation index, etc. are under processing to support MAIRS regional studies. For data that are not archived at the GES DISC, a product metadata portal is under development to serve as a gateway for providing product level information and data access links, which include both satellite, model products and ground-based

measurements information collected from MAIRS scientists. Due to the large overlap of geographic coverage and many similar scientific interests of NEESPI and MAIRS, these data and tools will serve both projects.

4.2 Sample plots by using Giovanni

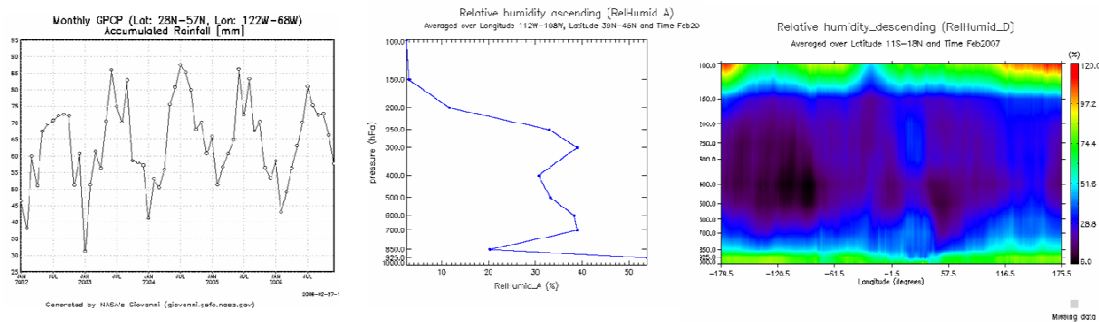
4.2.1 Basic visualization features

- Display maps with customized color, projection:



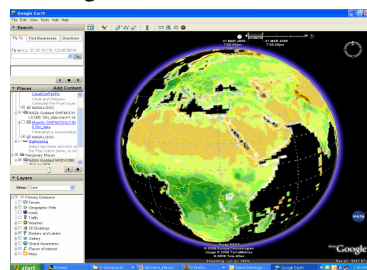
The color palette of images can be adjusted to get the best result. Maps can be viewed in Equidistant Cylindrical, North/South polar stereographic, and Robinson projection. The sample images are Chlorophyll a concentration from SeaWiFS (upper left), vegetation index from MODIS-Terra (upper right), ozone from OMI (lower left and right).

- Display cross-section image, profile, and time series:



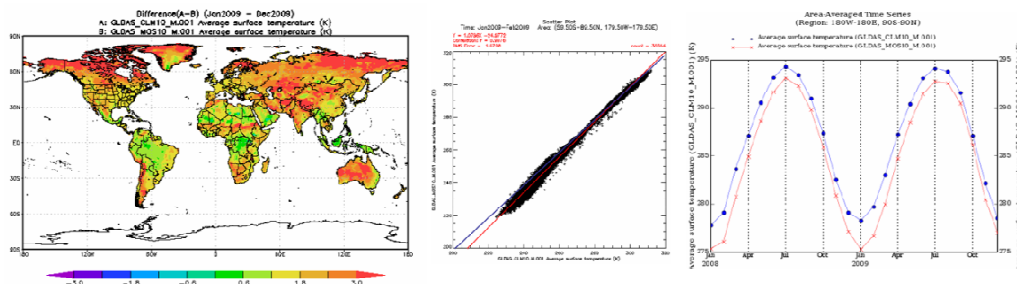
Giovanni can generate time series for a selected area (left, precipitation from GPCP). For 3-dimensional data, a vertical profile (center, relative humidity from AIRS) or vertical cross-section map (right, relative humidity from AIRS) can be generated.

- Display image with GoogleEarth:



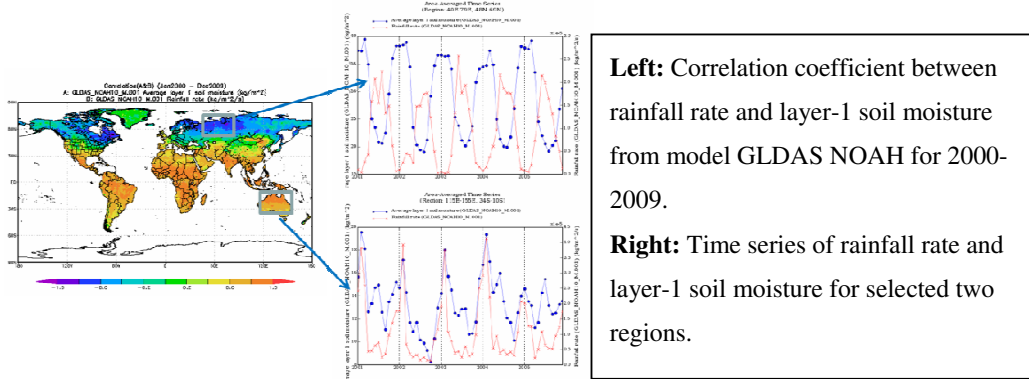
Giovanni images can be saved as KMZ file which can be displayed

- Comparison functions: difference, scatter plot, times series overlay:



Left Panel: The difference of annual averaged surface temperature from two GLDAS models, CLM and MOS of year 2009; **Middle Panel:** scatter plot of surface temperature of models CLM and MOD for Jan& Feb 2009; **Right Panel:** time series of global averaged surface temperature of models CLM and MOS from Jan 2008 to Dec 2009.

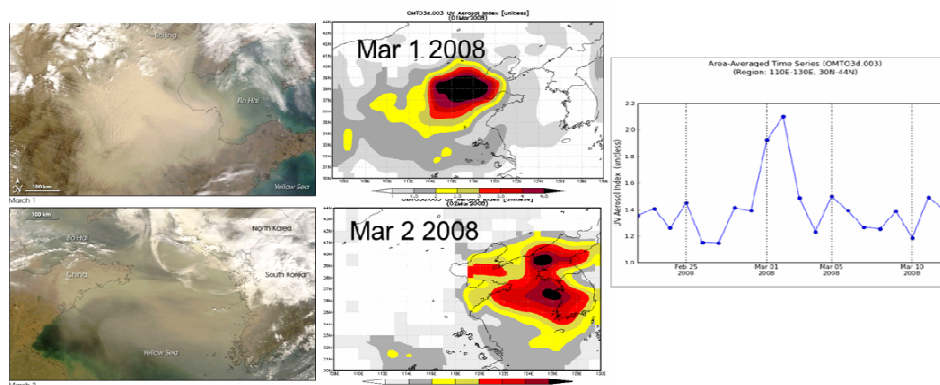
- Study relationship between two parameters: correlation map, time series overlay:



In addition to above plotting features, the Giovanni system can create animations; calculate zonal mean, histogram, trend of a time series; as well as perform anomaly analysis for some products.

4.2.2 Monitoring natural hazard events

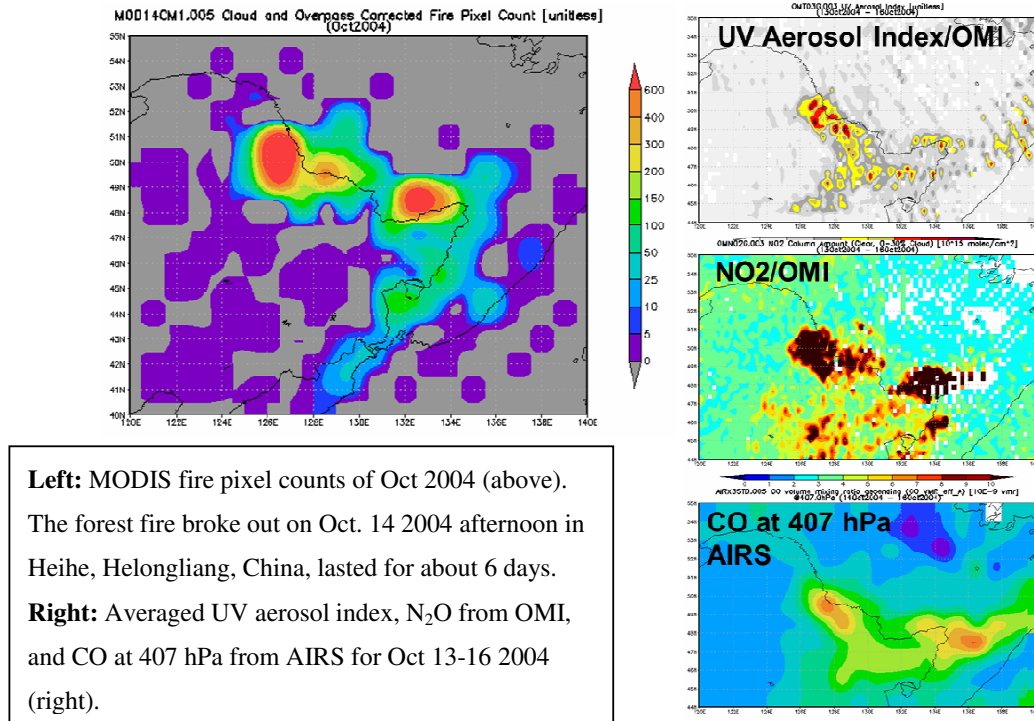
- Dust storm over East China, March 1-2 2008:



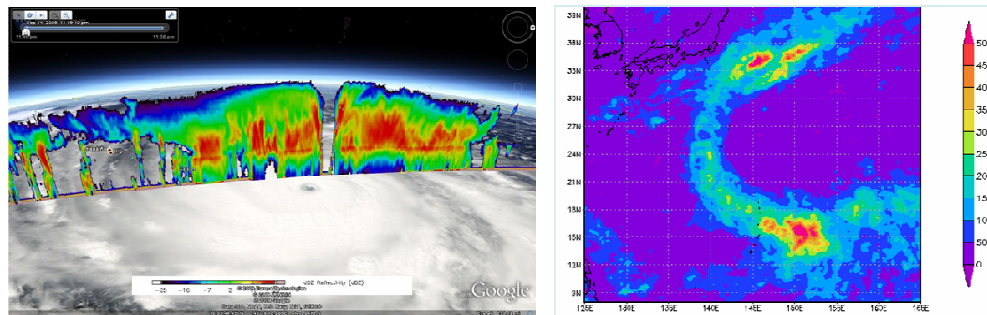
Left Panel: True Color images from MODIS Terra on Mar 1 03:20Z and Mar 2 02:25Z 2008 (from NASA MODIS Rapid Response System); Middle panel: UV Aerosol index from OMI of above two days; Right panel: time series of daily UV

aerosol index from Feb 22 to Mar 12 over (110°E-130°E, 30°N-44°N), indicating a significant aerosol increase on Mar 1-2.

- Forest fire in northeast China, October 14-19 2004:



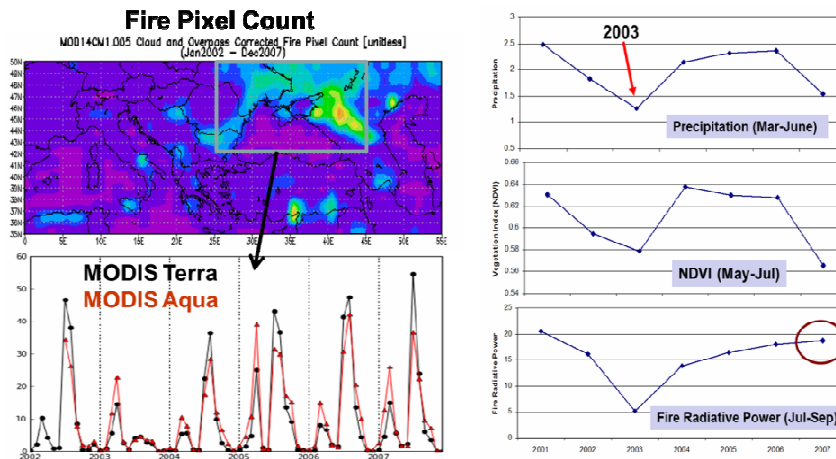
- Typhoon Choi-Wan (彩云) over the Western Pacific, Sep 11-20 2009:



Typhoon Choi-Wan reached Category 5 strength on the Saffir-Simpson hurricane/typhoon scale, formed on Sep 11 2009 over the Western Pacific near 15°N, 152°E. It moved to northwest and locked about 410 miles southeast of Tokyo, Japan on Sep 19 and moving further into the open North Pacific Ocean. Upper right image shows the accumulated rainfall of Sep 11-19 2009, which indicates the Typhoon track. The maximum accumulated rainfall along the track exceeded 500 mm. The upper left image is the vertical cross-section of the cloud reflectivity from CloudSat overlay on MODIS true color images on Sep 15 2009 displayed with GoogleEarth.

4.2.3 Studying seasonal and interannual variations

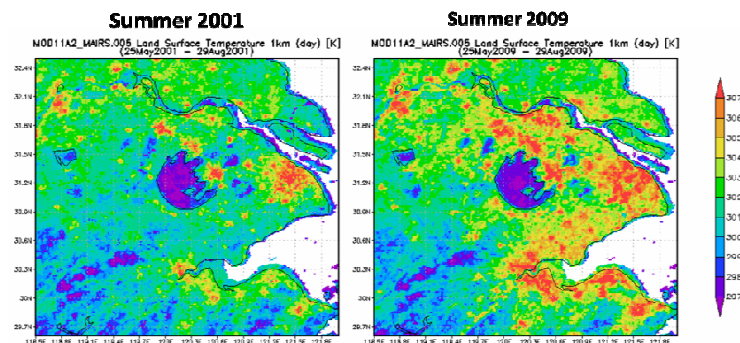
- Variation of active fire in temperate Europe. This is from the work of Shen et al. (2009):



Satellite observations show that in temperate Europe fires or hot spots occur mainly in croplands (upper left image, average of MODIS fire pixel count from 2002.01 to 2007.12). Time series of fire pixel count over the boxed area shows seasonal cycle of maximum in summer (Jul-Aug) and a second peak in spring (Mar-Apr) (curve in lower left panel), both may be associated with agricultural burning of crop residue for inexpensive and quick stubble removal while adding nutrients to the soil and killing weeds and pests at the same time. The value of 2003 summer is very low. It is found that the precipitation in spring to early summer is anomalous low which caused anomalous low crops in this region.

- The urban heat effect shown clearly in the 1km land surface temperature data:

NASA GES DISC MAIRS project are processing 1-km MODIS land surface temperature (LST) for MAIRS region. Above images are day-time LST from MODIS-Terra of summer season (Jun-Aug) for 2001 (left) and 2009 (right) over the Yangtze River Delta region, indicating that day-time LST is higher significantly than the rural areas. The difference between 2001 and 2009 indicates the fast urbanization in the Yangtze River Delta region.



4.3 Sample research results by using NASA satellite data

NASA Earth system science data and service publishes featured research articles annually, named “Sensing Our Planet”, to illustrate the use of Earth observing remote sensing data from NASA earth science data centers. The articles are available online at <http://nasadaacs.eos.nasa.gov/articles/index.html>. Hard copies can be ordered by sending email to nasadaacs@eos.nasa.gov

Sample research results by using GLDAS and other data are available at: <http://disc.sci.gsfc.nasa.gov/hydrology/additional/science-focus>. For example,

- “Monitoring water storage with GLDAS and GRACE”, based on the work of Rodell, et al. (2006): Scientists use GLDAS to help interpret the valuable and unique but low resolution hydrological data provided by GRACE.
- “Diurnal cycle of summertime precipitation from NLDAS data products”, based on works of several research papers. The diurnal (day-night) cycle of rainfall in the summer over the United States is examined in this study utilizing NLDAS-2 data sets.

Relevant Links:

NASA EOS program: <http://eospsso.gsfc.nasa.gov/>

US Landsat program: <http://landsat.gsfc.nasa.gov/>

NASA MAIRS data support project: <http://disc.gsfc.nasa.gov/mairs/>

Web-based visualization tool, Giovanni: <http://daac.gsfc.nasa.gov/giovanni>

NASA Earth System Science Data and Services: <http://nasadaacs.eos.nasa.gov/>

NASA EOS data order primary entry point, WIST: <http://wist.echo.nasa.gov>

NASA EOS data and documents searching system, ECHO: <http://www.echo.nasa.gov/>

Earth Science data, service, and information searching engine, GCMD: <http://gcmd.gsfc.nasa.gov>

NASA Land Measurement Portal: <http://landportal.gsfc.nasa.gov/>

References:

- Berrick, S.W., Leptoukh, G., Farley, J.D., and Rui, H. (2009). Giovanni: A Web service workflow-based data visualization and analysis system. *IEEE Trans. Geosci. Remote Sens.*, 47(1), 106-113
- Leptoukh, G., Csizsar, I., Romanov, P., Shen S., Loboda T., & Gerasimov, I. (2007). NASA NEESPI data center for satellite remote sensing data and services. *Global and Planetary Change, Environ. Res. Lett.*, 2. 045009, doi:10.1088/1748-9326/2/4/045009.
- Parkinson, Claire, Alan Ward, Michael D. King, 2006, Earth Science Reference Handbook – A guide to NASA’s Earth Science Program and Earth Observing Satellite Missions, Sterling Spangler, 291 p.
- Rodell, M., P.R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J.K. Entin, J.P. Walker, D. Lohmann, and D. Toll, The Global Land Data Assimilation System, *Bull. Amer. Meteor. Soc.*, 85(3), 381-394, 2004
- Shen, S., Leptoukh, G., Loboda, T., Csizsar, I., Romanov, P., and Gerasimov, I., The NASA NEESPI Data Portal to Support Studies of Climate and Environmental Changes in Non-boreal Europe". *In : Regional Aspects of Climate-Terrestrial-Hydrologic Interactions in Non-boreal Eastern Europe*, Groisman, Pavel Ya.; Ivanov, Sergiy (Eds.), Springer, 2009, 255 p., pp 9-16.



A parameterized multiple-scattering model for microwave emission from dry snow

Lingmei Jiang^{a,b,*}, Jiancheng Shi^c, Saibun Tjuatja^d, Jeff Dozier^e, Kunshan Chen^f, Lixin Zhang^{a,b}

^a State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and the Institute of Remote Sensing Applications of Chinese Academy of Sciences, School of Geography, Beijing Normal University, Beijing, 100875 China

^b Beijing Key Laboratory for Remote Sensing of Environment and Digital Cities, 100875 China

^c Institute for Computational Earth System Science, University of California, Santa Barbara, CA 93106-3060, USA

^d Wave Scattering Research Center, The University of Texas at Arlington, Arlington, TX 76019-0016, USA

^e Donald Bren School of Environmental Science and Management, University of California, Santa Barbara, CA 93106-5131, USA

^f Center for Space and Remote Sensing Research, National Central University, 32054 Chung-Li, Taiwan

Received 15 August 2006; received in revised form 20 January 2007; accepted 17 February 2007

Abstract

Snow water equivalent (SWE) is one of the key parameters for many applications in climatology, hydrology, and water resource planning and management. Satellite-based passive microwave sensors have provided global, long-term observations that are sensitive to SWE. However, the complexity of the snowpack makes modeling the microwave emission and inversion of a model to retrieve SWE difficult, with the consequence that retrievals are sometimes incorrect. Here we develop a parameterized dry snow emission model for analyzing passive microwave data, including those from the Advanced Microwave Scanning Radiometer-Earth Observing System (AMSR-E) at 10.65 GHz, 18.7 GHz, and 36.5 GHz for SWE estimation. We first evaluate a multiple-scattering microwave emission model that consists of a single snow layer over a rough surface by comparing model calculations with data from two field measurements, from the Cold Land Process Experiment (CLPX) in 2003 and from Switzerland in 1995. This model uses the matrix doubling approach to include incoherent multiple-scattering in the snow, and the model combines the Dense Media Radiative Transfer Model (DMRT) for snow volume scattering and emission with the Advanced Integral Equation Model (AIEM) for the randomly rough snow/ground interface to calculate dry snow emission signals. The combined model agrees well with experimental measurements. With this confirmation, we develop a parameterized emission model, much faster computationally, using a database that the more physical multiple-scattering model generates. For a wide range of snow and soil properties, this parameterized model's results are within 0.013 of those from the multiple-scattering model. This simplified model can be applied to the simulation of the microwave emission signal and to developing algorithms for SWE retrieval.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Radiative transfer; Dry snow; Passive microwave remote sensing; Parameterization

1. Introduction

Characteristics of spatial and temporal distributed snow properties play important roles in global energy and water cycles. Snow cover significantly influences the Earth's surface radiative balance and acts as the frozen storage term in the water

balance. Snow water equivalent (SWE) is important for hydrological applications and water resource management. In situ snow cover and SWE data, however, are available only at point measurements in a few areas that are poorly distributed globally (Robinson et al., 1993). Satellite passive microwave imagery has been used as a source of snow cover information because of all-weather imaging capabilities, rapid scene revisit time, and the ability to derive quantitative estimates of SWE (Derksen et al., 2000). Currently, there has been a growing use of microwave radiometry satellite observation in weather and climate prediction model (Marshall et al., 2005). It required

* Corresponding author. School of Geography, Beijing Normal University, Beijing, 100875 PR China. Tel.: +86 10 58809966; fax: +86 10 58805274.

E-mail address: jlingmei@hotmail.com (L. Jiang).

high accuracy and fast simulations of the satellite observation coupled with forecast modeling. Therefore, the accuracy and efficiency of forecast model is basically linked to both the accuracy and computation time of the radiative transfer model.

Passive microwave remote sensing of snow parameters, such as snow extent, snow water equivalent, and wet/dry state, have been investigated by many researchers using various microwave sensors (Goodison & Walker, 1994; Foster et al., 1997; Derksen et al., 2000; Pulliainen & Hallikainen, 2001; Kelly & Chang, 2003; Roy et al., 2004; Tedesco et al., 2004; Derksen et al., 2005a,b; Macelloni et al., 2005; Pulliainen, 2006). Over a broad range of frequencies, 3–90 GHz, microwave brightness temperature is sensitive to snow crystal characteristics, snow density, and water equivalent (Wiesmann & Mätzler, 1999; Pulliainen et al., 1999; Tsang & Kong, 2001; Macelloni et al., 2001), but it also depends on the physical temperature and properties of the underlying soil. At the lower frequencies, emission from dry snow is mainly affected by underlying soil dielectric and roughness properties; at higher frequencies, emission is sensitive to snow water equivalent and snow particle size since the volume scattering by snow particles becomes important (Mätzler, 1996). Because dry snow emits considerably less microwave radiation than soil, the brightness temperature of snow is inversely related to the snow water equivalent. When snow starts to melt, emission will significantly increase because water droplets absorb and re-emit rather than scatter microwave radiation (Foster et al., 2005).

In recent years, theoretical modeling of microwave emission from snow has advanced significantly and has provided a better understanding of snowpack scattering and emission processes (Mätzler & Wiesmann, 1999; Wiesmann & Mätzler, 1999; Pulliainen et al., 1999; Tsang & Kong, 2001). Snow is a dense medium owing to the high volume fraction of ice grains (10% to 50%), and there are interactions between the emitted microwave signal with the snow volume and surfaces. Snow volume scattering include both coherent (dense medium effect) and incoherent multiple-scattering. Several microwave snow emission models have been reported including the MEMLS model — a multilayer and multiple-scattering radiative transfer model (Wiesmann & Mätzler, 1999; Mätzler & Wiesmann, 1999) and the HUT model (Pulliainen et al., 1999). The dense medium radiative transfer (DMRT) model have been developed for modeling microwave signals of snow cover with either Rayleigh or Mie spherical scattering phase matrices with the quasi-crystalline approximations (Chuah et al., 1996; Tsang et al., 2000; Tsang & Kong, 2001). The DMRT takes into account the coherent wave interactions by the pair distribution function of the particle positions (Percus–Yevick equation) and is suitable for snow application (Tsang & Kong, 2001). The DMRT model predictions are in good agreement with numerical solutions of Maxwell's equations based on three-dimensional simulations (NMM3D), with laboratory controlled measurements (Chen et al., 2003a), and with field measurements for a variety of snow depths, grain sizes and densities (Tsang et al., 2000; Macelloni et al., 2001; Jiang et al., 2004; Tedesco et al., 2006).

Furthermore, theoretical modeling of surface emission and scattering has also significantly improved. The Integral

Equation Model (IEM) has demonstrated applicability to a much wider range of surface roughness conditions compared to conventional models. Recently, Chen et al. (2003b) extended the original IEM and developed the Advanced Integral Equation Model (AIEM), by removing some weak assumptions in the original IEM model development. Comparisons of AIEM with NMM3D-simulated data (Chen et al., 2003b) and field experimental data over the frequency range from 6 to 37 GHz (Shi et al., 2005) showed significantly better agreement than the original IEM model over a wide range of surface dielectric, roughness, and sensor frequencies. These efforts have established a fundamentally improved understanding of the effects of snow physical parameters and underlying surface dielectric and roughness properties on the microwave measurements of snow-covered terrain, making it possible to characterize microwave emission more accurately.

Vector radiative transfer theory (VRT), which is based on energy transport of partially polarized electromagnetic waves inside a medium, has been used for studying snow's effects on microwave signatures. A snow-layer emission model based on VRT accounts for incoherent multiple-scattering effects within the layer and the incoherent interactions between the volume and the layer surfaces (Fung, 1994). The VRT equations for a snow layer can be solved numerically using the eigen-analysis technique (Tsang et al., 2000; Tsang & Kong, 2001) or the matrix doubling (MD) method (Ulaby et al., 1986; Tjuatja et al., 1993). In terms of computation, matrix doubling is a more efficient method for layers that are optically thick, as is usual with snow.

By combining the recent advancements of theoretical modeling developments in each scattering and emission components of snow layer and surface with the technique of the multi-scattering radiative transfer solution, it is possible to develop a new multiple-scattering snow emission model that implements the most recent achievements in theoretical model developments to improve our understanding of the effects of the properties of the snow and of the underlying soil. However, a multiple-scattering snow emission model is complex, in general, and makes its direct application for analyses of microwave radiometer data or inferring snow parameters computationally difficult. Therefore, the field needs a simple and accurate snow emission model that can correctly represent the characteristics and relationships of the emission signals, so that a model can be used to drive algorithm development for SWE retrieval by passive microwave remote sensing and to provide a fast dry snow data assimilation model for applications in climate prediction and land surface process modeling affords. The major purpose of this study is to present such a model for terrestrial dry snow cover.

The manuscript is organized as following: In Section 2, we describe a dry snow multiple-scattering microwave emission model that implemented the recent achievements in theoretical model developments for each component. In Section 3, we compare this model with results from two field experiments. In Section 4, we demonstrate the development of a parameterized snow microwave emission model using the simulated data from our dry snow multiple-scattering microwave emission model. Section 5 presents our conclusions.

2. A slab multiple-scattering microwave emission model for dry snow

We have implemented the recent advancements in theoretical model development for both volume and surface emission models. The matrix doubling approach is used to include multiple-scattering and combines the Dense Media Radiative Transfer Model (DMRT) (Tsang, 1992) for snow volume scattering and emission, the Advanced Integral Equation Model (AIEM) for soil emission (Chen et al., 2003b), and the interactions of microwave signals between snow and soil with the surface bistatic scattering model from the AIEM model (Chen et al., 2003b) to calculate dry snow emission. Our multiple-scattering model is a combined DMRT-AIEM-MD microwave emission model for terrestrial dry snow cover with the consideration of one snow layer over a rough soil or rock surfaces.

In VRT formulation, the volume scattering phase matrix characterizes the coupling of intensities in any direction inside the layer caused by scattering and is usually derived using the electromagnetic wave formulation. The volume scattering phase matrix of an optically thin (infinitesimal) layer is determined by the properties of a collection of scatterers. A dry snow layer is a heterogeneous medium composed of ice particles with different sizes and microstructures. The shape and orientation of the snow crystal have little effect on the snow microwave emission (Foster et al., 1999, 2000), so the ice particles can be effectively modeled as spheres. The high volume fraction of ice grains causes the volume scattering to include both coherent scattering (the dense medium effect) and incoherent multiple-scattering. The multiple incoherent scattering is accounted for in the VRT formulation. Several volume scattering models for dense media account for the coherent interactions (Chuah et al., 1996; Tsang et al., 2000; Tsang & Kong, 2001). This study utilizes Tsang and Kong's (2001) volume scattering phase matrix for spatially-correlated spherical scatterers based on the dense medium radiative transfer (DMRT) model with the quasi-crystalline approximation.

Emission of electromagnetic waves from dry snow combines surface and volume scattering. To account for the surface scattering and surface-volume interactions, the VRT equations are subject to the boundary conditions at the air–snow and snow–ground surfaces. In the matrix doubling formulation, these boundary conditions are enforced by including the surface scattering phase matrices (Ulaby et al., 1986; Tjuatja et al., 1993; Fung, 1994). At the air–snow interface, the dielectric contrast between dry snow and air is commonly small, and for emission we neglect surface roughness. Therefore, we consider the air–snow interface as smooth and the snow–ground surface as rough.

To determine the total emission from a snow layer above ground, consider the geometry shown in Fig. 1. The total emission source within a layer can be separated into three components: the total upwelling emission u_u , the total downwelling emission u_d , and the ground emission into the layer u_g . The ground emission u_g is determined by the temperature and emissivity of the soil calculated by AIEM, that depends on the underground soil moisture and roughness at snow–ground

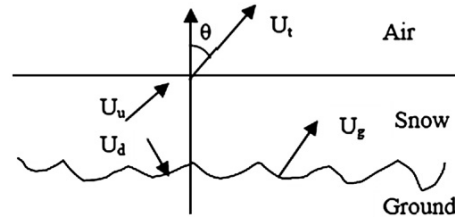


Fig. 1. Geometry of a snow-layer emission problem.

interface. The upwelling and downwelling emissions, u_u and u_d , are functions of the snow-layer temperature profile and its volume scattering properties. For a snow layer with optical thickness τ_0 , u_u and u_d are:

$$u_u(\tau_0) = \int_0^{\tau_0} T^0(\tau) \left\{ I - \frac{1}{4} S^0(\tau_0 - \tau) S^0(\tau) \right\}^{-1} \{ I + S^0(\tau_0 - \tau) \} (1 - a) \frac{K}{\lambda^2} U^{-1} T(\tau) d\tau \quad (1)$$

$$u_d(\tau_0) = \int_0^{\tau_0} T^0(\tau_0 - \tau) \left\{ I - \frac{1}{4} S^0(\tau) S^0(\tau_0 - \tau) \right\}^{-1} \{ I + S^0(\tau) \} (1 - a) \frac{K}{\lambda^2} U^{-1} T(\tau) d\tau \quad (2)$$

where $S^0(\delta)$ and $T^0(\delta)$ are the zeroth-order Fourier components in the backward and forward scattering phase matrices of the snow layer, they are calculated by the spherical Mie scattering phase matrices with the quasi-crystalline approximation (QCA) and the pair distribution function of the particle positions and coherent wave interactions using Percus–Yevick equation (Tsang & Kong, 2001). I is the identity matrix, U is the directional cosine matrix for the polar angles, K is the Boltzmann constant, T is the temperature profile of the layer, a is the albedo, and λ is wavelength.

The total emission from the snow layer, u_T , is

$$u_T = L_u u_u + L_d u_d + L_g u_g \quad (3)$$

where the L 's are multiple-scattering operators that account for all surface and surface-volume multiple-scattering. The volume scattering phase matrices S^0 and T^0 and the multiple-scattering operators L 's are solved using the matrix doubling method with the snow/ground boundary conditions computed using AIEM model (Chen et al., 2003a,b) for the surface reflectivity matrix. In this way, the interactions between the scattering and emission from snowpack and the ground rough surface can be taken into account.

3. Comparisons of the model with experimental data

To validate the DMRT-AIEM-MD emission model, we compared the model simulations with two field experimental data sets that were obtained from the ground radiometer measurements over dry snow covers. The first one is from the passive microwave experiments of snowpacks in the Alps (Wiesmann et al., 1996) measured on Dec. 22, 1995 at Weissfluhjoch (46°49', 83°N, 9°48', 62°E) in Davos, Switzerland. The other field experimental data set with both the ground radiometer and snow pit

Table 1
Snow profile data on December 22, 1995 at Weissfluhjoch

Height (m)	Grain Shape	Temperature of snow (°C)	Snow density (kg/m ³)
0.40–0.60	++	−2.2	109.0
0.25–0.40	/ \	—	177.0
0.00–0.25	□ □	−0.7	259.0

measurements at Fraser (39.9066°N, 105.8829°W) is from the Cold Land Processes Experiment (CLPX) during Feb. 2003 in northern Colorado, U.S.A.

3.1. Comparison with Weissfluhjoch data

In the Weissfluhjoch experiment on 22 Dec 1995, ground radiometric measurements at 11, 35, and 94 GHz were obtained with a set of portable linearly polarized Dicke radiometers, about 160 cm above the surface. The measurements were obtained with incidence angles from 20° to 70° at 5° intervals. Snow properties — collected nearly simultaneously with the radiometric measurements — included snow depth, grain shapes, temperature, permittivity, density and weather conditions. A snow profile was measured with temperature, grain shape, permittivity, and density in 10 cm steps. Permittivity was measured with open coaxial resonators (Mätzler, 1996), and grain shapes were classified according to the international classification for seasonal snow (Colbeck, 1986). The ground is covered with stones and rocks composed of serpentine. The experiment was carried out during a sunny day, with low sky microwave brightness temperatures — 5.7 K, 11.7 K, and 29.0 K at 11, 35 and 94 GHz. The total emissivity (e) of the snowpack observed by the radiometer is

$$e = \frac{T_b - T_{\text{sky}}}{T_s - T_{\text{sky}}} \quad (4)$$

where T_b , T_s are the observed radiation and snow physical temperature and T_{sky} is the brightness temperature of the downwelling atmospheric radiation.

Table 1 shows the snow profile data (Wiesmann et al., 1996). In a winter snowpack of 60 cm depth, the top layer consists of 20 cm of new snow above a thin crust. Below the crust, the bottom layer consisted of coarse grains. Snow temperature was

271.2 K. The vertical profiles of snow grain shape and density were also available as shown in Table 1. Due to the fact that the DMRT-AIEM-MD model is a one-layer snow model, snow input parameters for the model were a density of 220 kg m^{−3} and a grain radius of 0.4 mm that were determined by fitting the high frequency measurements at 94 GHz because the measurements at 94 GHz cannot “see” the soil surface 60 cm below. There were no observations of the ground surface roughness and soil moisture during the field experiment. The ground surface dielectric and roughness parameters are treated as fitting parameters with the minimum mean square difference between the modeled and the measured brightness temperatures at 11 GHz, where the subsurface emission signals are the dominant emission source. Results are 1 cm of surface rms height and 20 cm of surface correlation length. The ground surface soil dielectric constant was computed using 10% volumetric soil moisture (Dobson et al., 1985).

Fig. 2 shows the observations and the model calculations for vertical and horizontal polarization. At all three frequencies, the DMRT-AIEM-MD model predictions are able to match the measured snow emissivity data reasonably well in terms of magnitudes in both polarizations and in difference between polarizations. Snow emissivity decreases as both v and h polarizations as frequency increases. Similarly, the polarization difference between v and h polarizations also decreases, especially at incidence angle less than 55°. Table 2 summarizes the root mean square error (RMSE) between the model predictions and the measured snow emission data at each frequency. It shows very good agreement in v polarization at all three frequencies with maximum RMSE less than 0.015. However, the results for h polarization are not as good, as with maximum RMSE 0.039. The overall RMSE for all three frequencies is 0.013 for v and 0.033 for h polarization. The reason for the discrepancy between the model and measurements is mainly due to underground signal fitting is poor in h polarization with RMSE 0.038 at X-band, which can result in a poor calculation at Ka-band as shown in Table 2. The other reason might be because the DMRT-AIEM-MD emission model is a slab emission model with the average snow properties from the vertical snow profile, which were taken with a slight adjustment from the high frequency measurements as our model inputs for all frequencies. At lower frequencies, more emission would come from the lower part of the snowpack. The

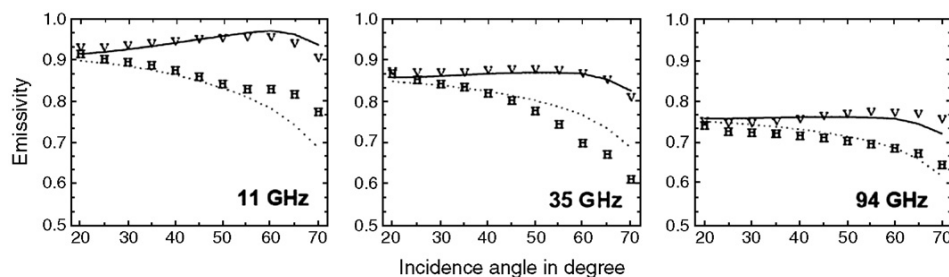


Fig. 2. Emissivity versus incidence angle at 11 GHz, 35 GHz, and 94 GHz. Solid lines show model calculations for vertical polarization, dotted lines for horizontal. Letters represent the corresponding field measurements.

Table 2
RMSE of the comparison of the DMRT-AIEM-MD model with the experiment data on December 22, 1995 at Weissfluhjoch

Polarization	Frequency			Overall
	11 GHz	35 GHz	94 GHz	
ν -pol	0.015	0.009	0.015	0.013
h -pol	0.038	0.039	0.015	0.033

underground emission signal might play an important role. At high frequency, however, the underground emission signal might not be significant. The emission signal might mainly come from the upper part of the snowpack due to its limited penetration capability. The simulated emissivity at 94 GHz is very sensitive to snow density, so only a slight change in snow density would result in a significant change in the simulated emissivity. It indicates that the microwave measurements at different frequencies may look at the different snow properties at the different depth depending on its penetration capability. The average snow parameters from the vertical profile measurements may not be the optimal snow parameters for microwave modeling. In addition, the ice lens or a thin crust layer inside the snowpack can considerably affect the spectral behavior of emissivity at the horizontal polarization but not at the vertical polarization (Wiesmann et al., 1998).

3.2. Model comparisons with CLPX'03 data

The Cold Land Process Experiment in 2003 included measurements with the ground-based passive microwave radiometer (GBMR) from the University of Tokyo at three frequencies — 18.7 GHz, 36.5 GHz, and 89 GHz — at both vertical and horizontal polarizations (Graf et al., 2003). The measurement site had an area about 100 m × 100 m at Fraser (39.9066°N, 105.8829°W) in the Rocky Mountains of Colorado. The study area has flat topography and is surrounded by a uniform pine forest and a discontinuous pine forest. Data used in this study were the measurements from dry snow conditions during February 2003. There were two experiment setups, one with a constant incidence angle at 55° and measurements averaged over all azimuths, the other with incidence angles from 30° to 70° but with a fixed azimuth. Each experimental setup was performed at every other day.

During the experiment, there were two light snowfalls on February 20 and 24, and a larger snowfall on 22 February, which added about 20 cm of new fresh snow. Snow pit data were collected around noon from two different snow pit

locations within 10 m of the area observed by the radiometer. Faceted crystals were the most plentiful type of ice crystal for this site except on the last day (25 February), when rounded particles and wet agglomerates appeared. This may cause some discrepancy between the real and modeled shapes of ice particles.

The model input parameters were the mean snow depth and density obtained from the ground measurements (Hardy et al., 2003). Snow grain size measurements include the information on minimum, average, and maximum determined by the product of two-dimension optical grain length with a range from about 0.2 mm to 7 mm for each grain size group: small, moderate, and large grain groups. Therefore, the input ice particle size for the model was a fitting parameter selected within the grain size measurement range, which actually was smaller than the observed mean particle size from all measurements. The volumetric soil moisture of 24% measured at 1.5 cm near the site was used to calculate soil dielectric constant (Dobson et al., 1985). It varied little during the entire experiment. The soil roughness properties were the fitting parameters. Since there were no sky brightness temperature observation during this experiment, we calculated T_{sky} for clear-sky conditions (Ulaby et al., 1986; Westwater et al., 1990) at 18.7 GHz, 36.5 GHz and 89 GHz at 18 K, 30 K, and 80 K, respectively. All model input parameters used for simulation are shown in Table 3.

Fig. 3 shows the comparisons between the DMRT-AIEM-MD model emissivities and the measurements. The top plot shows variability with observations angles, and the bottom plot shows a time-series. They show that the predictions of the DMRT-AIEM-MD model compare well with the measurements at the different incidence angles and at all three frequencies. As Table 4, shows the RMSEs are 0.024, 0.019 and 0.017 for ν polarization and 0.023, 0.022 and 0.009 for h polarization at 18.7 GHz, 36.5 GHz, and 89 GHz. Some discrepancies between the measurements and the model calculations exist and can be also explained by the possible reasons as described in analyses of Weissfluhjoch data and model comparison in Section 3.1. For CLPX experimental data, our model performance is better at 36.5 GHz and 89 GHz than at 18.7 GHz. Observed emissivity at 18.7 GHz for both ν and h polarizations increases from 65° to 70° incidence, but at 36.5 GHz and 89 GHz emissivity does not increase. One possible reason might be that the radiometer observed different footprints at difference incidence angles, because the instrument was installed at a fixed height above ground. There may have been different soil properties at the different incidence angles, but these would not affect the signal at the higher frequencies.

Table 3
Averaged snow parameters at Fraser

Date	Feb. 19	Feb. 20	Feb. 21	Feb. 22	Feb. 23	Feb. 24	Feb. 25
Depth (m)	0.80	0.83	0.78	1.02	0.99	1.08	0.99
Density (kg/m ³)	227	228	204	202	206	212	214
Radius (mm)	0.38	0.39	0.36	0.35	0.34	0.33	0.33
T_{snow} (°C)	−2.0	−2.0	−1.0	−0.0	−3.0	−3.0	−2.0
T_{ground} (°C)	−0.0	−0.0	−0.0	−0.0	−0.0	−1.0	−0.0

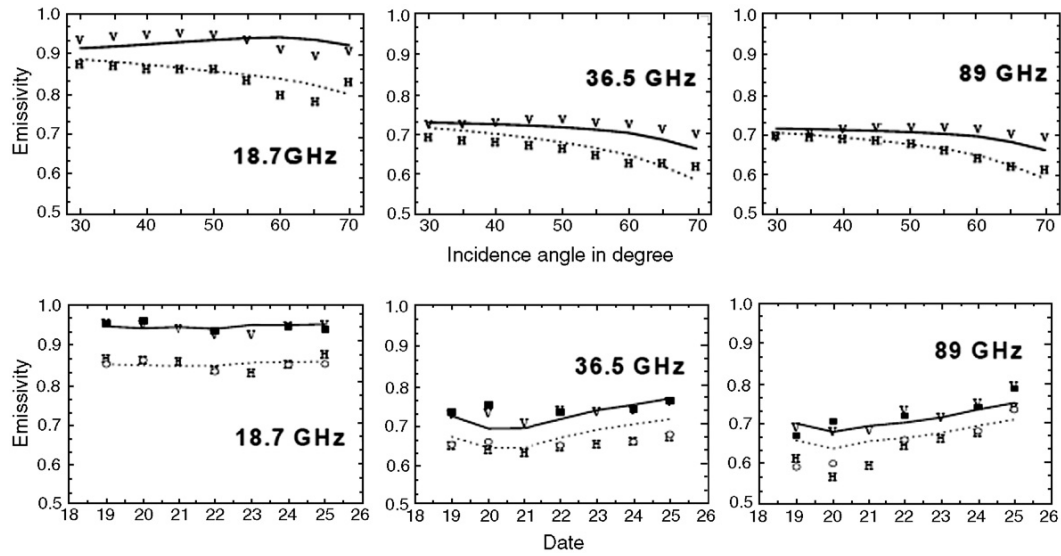


Fig. 3. Comparison of the model with observation at different angles (bottom) on 22 Feb 2003 (top) and with time-series observation at 55° incidence angle from 19–25 Feb 2003 (bottom). Solid and dotted lines show model calculations for v and h polarization. Letters show the measurements.

In the time-series observations, RMSEs are 0.011, 0.023 and 0.024 for v polarization and 0.013, 0.031 and 0.04 for h polarization at 18.7 GHz, 36.5 GHz, and 89 GHz, respectively (Table 4). The significant jump in the 89 GHz observations on 25 Feb is probably caused by melting snow in the surface layer.

Recognizing there are also uncertainties in the field data, we believe that the model and observations agree reasonably well.

4. Development of a parameterized model

At low frequencies, the commonly used ω – τ model is derived from the analytical solution of the 0th-order radiative transfer equations. Without considering snow and ground temperature, total emissivity E_p^t is

$$E_p^t = \left(E_p^v + E_p^v \cdot L_p \cdot R_p^c + L_p \cdot E_p^s \right) \cdot \Psi_p \quad (5)$$

The superscripts t , v , and s represent the emissivity components for total, volume, and surface. The subscript p represents the polarization status v or h . Ψ is the power transmittivity at the air–snow interface. $L_p = \exp(-\tau/\cos(\theta_r))$ is the attenuation factor. θ_r is the refractive angle in the snow layer and τ is snow optical thickness. $R_p^c = 1 - E_p^s$ is the ground surface effective reflectivity. Eq. (5) is commonly considered as a three-component model. The first term is the direct snow emission component $E_p^v = (1 - \omega) \cdot (1 - L_p)$ where ω is the snow volume scattering albedo. The second term in Eq. (5) is the snow–ground interaction term and represents that the downward snow emission signal is reflected back through the snow layer again. The last term in Eq. (5) represents the underground emission signal after passing through snowpack.

The 0th-order and 1st-order radiative transfer models cannot predict emission very well when the snow volume scattering albedo and optical thickness are large (Ulaby et al., 1986; Fung,

1994). They commonly underestimate snow emission signals. On the other hand, the multiple-scattering snow emission model is very complex, without an analytic solution, and computationally intensive. It is unrealistic to apply the multiple-scattering model directly in analyses of satellite measurements and for SWE algorithm development. Therefore, it is necessary to develop a simple but accurate dry snow emission model that can be used for both fast forward simulation and development of SWE inversion models. To avoid the weakness of some empirical approaches, we use a database generated by the multiple-scattering model that covers most possible snow and soil surface conditions. For specifications of a simple model, we have adopted the following constraints:

- The model must be accurate at each frequency and polarization that is used in the remote sensing application.
- The model must use the snow's physical characteristics, as indicated by the simple radiative transfer model.

Although the low scattering-order models such as 0th-order or 1st-order radiative transfer models that do not take multiple-scattering into account, they provide basic descriptions of the characteristics of microwave emission signal in responding to snow and soil properties. They are valid when snow scattering

Table 4
RMSE of the comparisons of the DMRT-AIEM-MD model with the experiment data at Fraser

Polarization	Frequency			
	18 GHz	36 GHz	89 GHz	Overall
v -pol (time-series)	0.011	0.023	0.024	0.0205
h -pol (time-series)	0.013	0.031	0.040	0.030
v -pol (angular)	0.024	0.019	0.017	0.020
h -pol (angular)	0.023	0.022	0.009	0.019

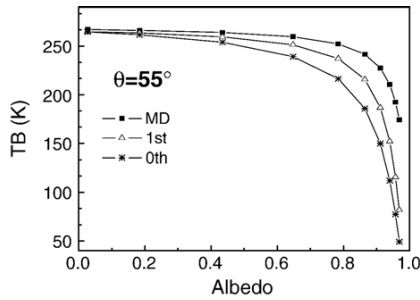


Fig. 4. Emission comparison for different scattering-order solutions at 55° incidence angle at vertical polarization (MD — matrix doubling method, 1st — first-order solutions, 0th — the zeroth-order solutions).

albedo and optical thickness are not too large. For large snow albedo and optical thickness, the low scattering-order models underestimate the emission signals, as Fig. 4 shows. The simulated emissivities by the different scattering-order models clearly show the same trend, i.e., the emission signal decreases as the scattering albedo increases. Although the 0th-order radiative transfer model underestimates the brightness temperature, it is possible to select this model as the basic form and add corrections to include multiple-scattering for our parameterized dry snow emission model.

To start, we simulated dry snow emission for both v and h polarizations at an incident angle of 55° using our DMRT-AIEM-MD model, using AMSR-E frequencies 10.65 GHz, 18.7 GHz, and 36.5 GHz. The simulations cover a wide range of snow and soil dielectric and roughness values as summarized in Table 5. The commonly used Gaussian correlation function for the soil surface was used in the simulation since it is a better approximation for high frequency microwave measurements than the exponential correlation function. With the different combinations of each snow and ground parameter, there are 153,600 DMRT-AIEM-MD model simulated emissivities at each frequency and polarization.

The 0th-order radiative transfer model (4) can be rearranged to a two-component model as

$$E_p^t = \left(E_p^v \cdot (1 + L_p) + L_p \cdot (1 - E_p^v) \cdot E_p^s \right) \cdot \Psi_p \\ = \left(\text{Intercept} + \text{slope} \cdot E_p^s \right) \cdot \Psi_p \quad (6)$$

The total emission and soil emission signals can be described as a linear function with the slope and intercept that depend only

Table 5
Snow and ground parameters for the database simulation

Input parameters	Minimum	Maximum	Step	Units
Snow density	150	450	100	Kg m ⁻³
Grain radius	0.2	1.6	0.2	Mm
Snow depth	0.1	2.0	0.1	m
Ground rms height	0.5	3.0	0.5	cm
Ground surface rms slope	0.05	0.25	0.05	—
Soil moisture	5	40	5	%

Table 6
Coefficients in Eq. (8) for calculating Cf_p^v

Frequency (GHz)	a	b	c	d	e
10.7	2.098563	−0.8514113	0.07340312	0.4980226	−1.874751
18.7	2.112606	−0.8829258	0.108098	0.5742384	−1.279307
36.5	2.120885	−0.7705214	0.07772004	0.3342744	−1.280769

on snowpack emission and attenuation properties or that are affected only by snow properties. Based on this description, we first divided the DMRT-AIEM-MD model simulated emissivity by the power transmissivity Ψ_p , then carried out the linear regression analyses between a set of all simulated ground surface emission signals and the corresponding total emission signals for a given snow grain size and density. In this way, the linear regression coefficients represent the snowpack emission and attenuation properties with the multiple-scattering effects for that snow grain size and density. By looping through different combinations of snow grain size and density in the ranges as listed in Table 5, all slopes and intercepts can be determined for all snow grain sizes and densities. Unfortunately, this technique can be only applied to the conditions when the soil emission signals can penetrate the snowpack. Therefore, we limit our analyses to the data with snow optical thickness $\tau \leq 2$ at each frequency in our simulated database.

Through our analyses and comparison with the components of the 0th-order radiative transfer model, we develop our parameterized dry snow emission model that includes multiple-scattering:

$$E_{mp}^t \approx \left(E_p^v \cdot Cf_p^v + L_p \cdot (1 - E_p^v) \cdot Cf_p^{svs} E_p^s \right) \cdot \Psi_p \quad (7)$$

E_{mp}^t is the total emissivity simulated by our multiple-scattering model. The first term $E_p^v \cdot Cf_p^v$ is the intercept determined in the linear regression analyses. E_p^v is direct snow volume emissivity in the 0th-order form as given in Eq. (5). Cf_p^v is the multiple-scattering correction factor that corrects for the difference in the direct volume emission signal between the 0th-order and the multiple-scattering models:

$$Cf_p^v = a + b \cdot \omega + \tau' \cdot (c + d \cdot \omega + e \cdot \omega^2) \quad (8)$$

where $\tau' = \tau / \cos(\theta_t)$ is the optical path length and where the coefficients a , b , c , d and e are determined by the linear regression analysis and are given in Table 6.

The second term in Eq. (7), excluding the underground emissivity E_p^s , $L_p (1 - E_p^v) Cf_p^{svs}$ represents the slopes determined

Table 7
Coefficients in Eq. (9) for calculating Cf_p^{svs}

Frequency (GHz)	A	B	C	D
10.7	−0.0268096	0.43702293	0.89230177	−0.75151270
18.7	−0.0737394	0.52525146	0.71524232	−0.61666468
36.5	−0.1397970	0.6268216	0.5559191	−0.4737233

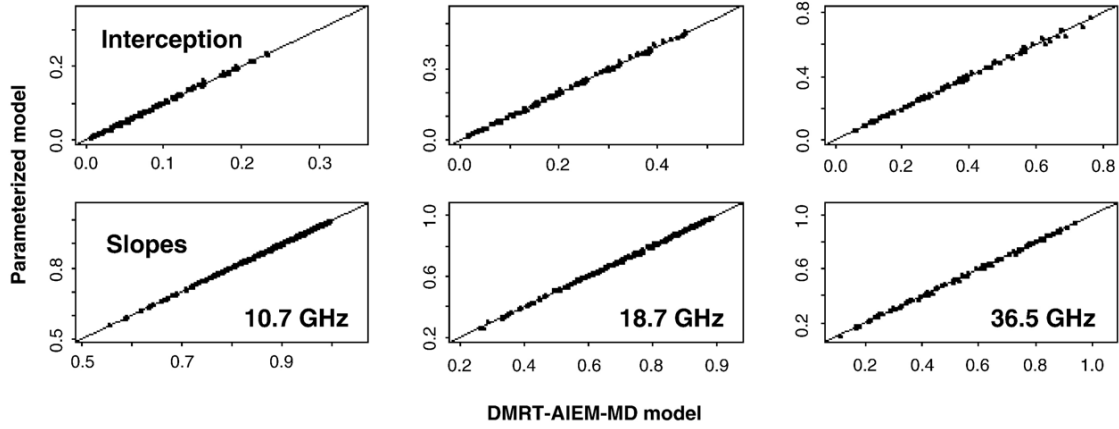


Fig. 5. Comparison of the intercepts (top row) and slopes (bottom row) determined by the DMRT-AIEM-MD model simulated data and these modeled by (7) for frequencies of 10.7, 18.7, and 36.5 GHz from left to right.

by the linear regression analyses. The multiple-scattering correction factor Cf_p^{svs} can be expressed as

$$Cf_p^{svs} = \exp[\tau' \cdot (A + B \cdot \omega) + \tau^2 \cdot (C \cdot \omega + D \cdot \omega^2)] \quad (9)$$

Similarly, the coefficients A , B , C , and D are constants determined by regression analysis, given in Table 7.

Fig. 5 (top row) shows the comparisons between the intercepts that were determined by the linear regression analyses and the model for the first term in Eq. (7): $E_p^v \cdot Cf_p^v$. The differences between these two are extremely small with RMSEs of 0.0017, 0.0033, and 0.007 for 10.65, 18.7, and 36.5 GHz, respectively. Fig. 5 (bottom row) shows the comparisons between the slopes that were determined by the linear regression analyses and the model in the second term of Eq. (7): $L_p (1 - E_p^v) Cf_p^v$. Similarly, the errors are also very small with RMSEs of 0.0015, 0.0039 and 0.0081 for 10.65, 18.7, and 36.5 GHz, respectively.

Fig. 6 (top row) shows the comparisons between the DMRT-AIEM-MD model simulated emissivities and the simple para-

meterized model for v polarization at 10.65, 18.7, and 36.5 GHz. The bottom row in Fig. 6 shows these for h polarization. The differences between these two models are extremely small with RMSEs of 0.0041, 0.0071, and 0.010 for v polarization at the frequencies of 10.65, 18.7, and 36.5 GHz, respectively. These for h polarization are 0.0052, 0.0087, and 0.013 at 10.65, 18.7, and 36.5 GHz. This simple parameterized model approaches the complex multiple-scattering model fairly well. The RMSEs at 10.7 GHz are the smallest among these three frequencies because multiple-scattering at lower frequency is smaller than that at higher frequencies. From these comparisons, the errors resulted in Eqs. (7–9) are not significant. The simple parameterized model provides both a fast way to simulate dry snow emission signals and a future possibility for improving algorithms for remote sensing of snow water equivalent. The newly developed dry snow emission model Eq. (7) is very simple and suitable for the microwave remote sensing applications with the negligible error in comparison to the complex multiple-scattering model simulations.

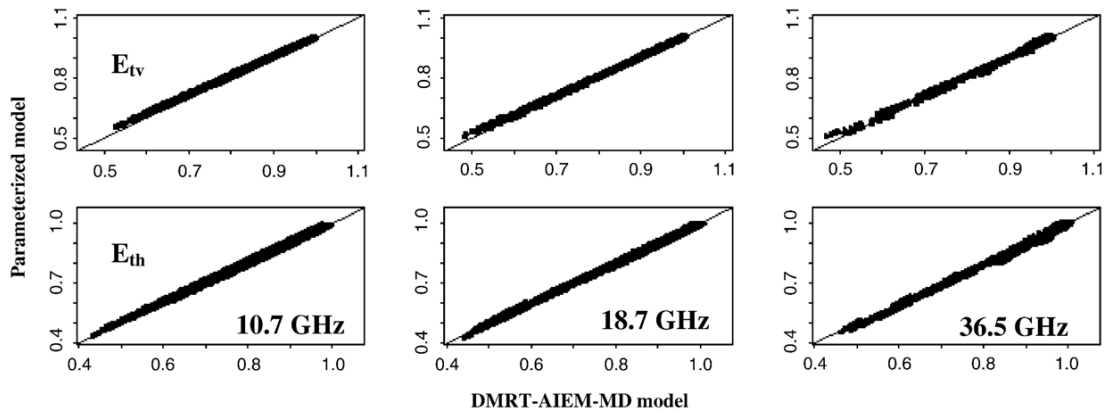


Fig. 6. Comparison of the total emissivities calculated by the DMRT-AIEM-MD model simulated data and these modeled by (7) for v polarization (top row) and h polarization (bottom row) at the frequencies of 10.7, 18.7, and 36.5 GHz from left to right.

5. Conclusions

The recent improvements in both volume and surface scattering models have made it possible to simulate dry snow emission signals for a very wide range of snow and soil conditions. We have integrated these recent developments into the DMRT-AIEM-MD emission model to simulate dry snow microwave emission with multiple-scattering. Comparisons with field data at Weissfluhjoch and CLPX'03 show good agreement for all frequencies and both polarizations.

With confirmed confidence in our multiple-scattering snow emission model, we develop a parameterized multi-frequency-polarization dry snow emission model for fast simulation and analyses of the passive microwave satellite measurements from AMSR-E. This model has a simple form similar to the 0th-order radiative model but adjusted with multiple-scattering correction factors. The differences between the simple parameterized model and the multiple-scattering model DMRT-AIEM-MD are extremely small, with the RMSEs in 10^{-3} range for all three study frequencies at 10.65, 18.7, and 36.5 GHz and both v and h polarizations, except the RMSE is 0.013 at 36.5 GHz and h polarization. This simple model provides a simple, accurate connection between dry snow emission at different frequencies and polarizations. It can be applied to the conditions with the snow optical thickness $\tau \leq 2$ at each of the three studied frequencies and with the dielectric homogeneous half space underground. The simple dry snow emission model developed in this study should be useful in understanding and analyzing the current and past passive microwave satellite measurements of the fully snow-covered pixels. In global SWE monitoring, however, the sub-grid heterogeneity is expected to have the significant impact and has to be taken into account. This issue needs to be further studied.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (90302008), by the NASA NNG04GC52A and by Program for Changjiang Scholars and Innovative Research Team in University (PCSIRT). The two sets of ground field data used in this study were obtained from Snow Tools Signature Database and from CLPX data at National Snow and Ice Data Center (NSIDC). We also thank the anonymous reviewers for their helpful comments on this paper.

References

- Chen, C. T., Tsang, L., Guo, J., Chang, A. T. C., & Ding, K. H. (2003). Frequency dependence of scattering and extinction of dense media based on three-dimensional simulations of Maxwell's equations with applications to snow. *IEEE Transactions on Geoscience and Remote Sensing*, 41(8), 1844–1852.
- Chen, K. S., Wu, T. -D., Tsang, L., Li, Q., Shi, J., & Fung, A. K. (2003). Emission of rough surfaces calculated by the integral equation method with comparison to three-dimensional moment method simulations. *IEEE Transactions on Geoscience and Remote Sensing*, 41(1), 90–101. doi:10.1109/TGRS.2002.807587
- Chuah, H. -T., Tjuatja, S., Fung, A. K., & Bredow, J. W. (1996). A phase matrix for a dense discrete random medium: Evaluation of volume scattering coefficient. *IEEE Transactions on Geoscience and Remote Sensing*, 34(5), 1137–1143. doi:10.1109/36.536529
- Colbeck, S. C. (1986). Classification of seasonal snow cover crystals. *Water Resources Research*, 22(9, suppl.), 59S–70S.
- Derksen, C., LeDrew, E., Walker, A., & Goodison, B. (2000). Influence of sensor overpass time on passive microwave-derived snow cover parameters. *Remote Sensing of Environment*, 71, 297–308.
- Derksen, C., Walker, A., Goodison, B., & Walter, S. (2005). Integrating in situ and multiscale passive microwave data for estimation of subgrid scale snow water equivalent distribution and variability. *IEEE Transactions on Geoscience and Remote Sensing*, 43(5), 960–972.
- Derksen, C., Walker, A., & Goodison, B. (2005). Evaluation of passive microwave snow water equivalent retrievals across the boreal forest/tundra transition of western Canada. *Remote Sensing of Environment*, 96, 315–327.
- Dobson, M. C., Ulaby, F. T., Hallikainen, M. T., & El-Reyes, M. (1985). Microwave dielectric behavior of wet soil, part 2: Dielectric mixing models. *IEEE Transactions on Geoscience and Remote Sensing*, 23, 35–46.
- Foster, J. L., Chang, A. T. C., & Hall, D. K. (1997). Comparison of snow mass estimates from prototype passive microwave snow algorithm, a revised algorithm and a snow depth climatology. *Remote Sensing of Environment*, 62(2), 132–142. doi:10.1016/S0034-4257(97)00085-0
- Foster, J. L., Hall, D. K., Chang, A. T. C., Rango, A., Wergin, W., & Erbe, E. (1999). Effects of snow crystal shape on the scattering of passive microwave radiation. *IEEE Transactions on Geoscience and Remote Sensing*, 37(2), 1165–1168. doi:10.1109/36.752235
- Foster, J. L., Barton, J. S., Chang, A. T. C., & Hall, D. K. (2000). Snow crystal orientation effects on the scattering of passive microwave radiation. *IEEE Transactions on Geoscience and Remote Sensing*, 38(5), 2430–2434.
- Foster, J. L., Sun, C., Walker, J. P., Kelly, R. E. J., Chang, A. T. C., Donga, J., et al. (2005). Quantifying the uncertainty in passive microwave snow water equivalent observations. *Remote Sensing of Environment*, 94(2), 187–203. doi:10.1016/j.rse.2004.09.012
- Fung, A. K. (1994). *Microwave Scattering and Emission Models and their Applications* (pp. 573). Boston: Artech House.
- Goodison, B., & Walker, A. (1994). Canadian development and use of snow cover information from passive microwave satellite data. In B. Choudhury, Y. Kerr, E. Njoku, & P. Pampaloni (Eds.), *Passive Microwave Remote Sensing of Land-Atmosphere Interactions* (pp. 245–262). Utrecht: VSP BV.
- Graf, T., Koike, T., Fujii, H., Brodzik, M., & Armstrong, R. (2003). CLPX-ground: Ground Based Passive Microwave Radiometer (GBMR-7) Data. *National Snow and Ice Data Center*: Boulder, CO.
- Hardy, J., Pomeroy, J., Link, T., Marks, D., Cline, D., Elder, K., et al. (2003). Updated 2003 CLPX-Ground: Snow measurements at the Local Scale Observation Site (LSOS). *National Snow and Ice Data Center*: Boulder, CO.
- Jiang, L., Shi, J., Tjuatja, S., & Chen, K. S. (2004). A comparison of dry snow emission model with field observations. *Proceedings IGARSS 2004, Vol. 6* (pp. 3709–3712). doi:10.1109/IGARSS.2004.1369926
- Kelly, R. E. J., & Chang, A. T. C. (2003). Development of a passive microwave global snow depth retrieval algorithm for Special Sensor Microwave Imager (SSM/I) and Advanced Microwave Scanning Radiometer-EOS (AMSR-E) data. *Radio Science*, 38(4), 8076. doi:10.1029/2002RS002648
- Macelloni, G., Paloscia, S., Pampaloni, P., & Tedesco, M. (2001). Microwave emission from dry snow: a comparison of experimental and model results. *IEEE Transactions on Geoscience and Remote Sensing*, 39(12), 2649–2656. doi:10.1109/36.974999
- Macelloni, G., Paloscia, S., Pampaloni, P., Brogioni, M., Ranzi, R., & Crepaz, A. (2005). Monitoring of melting refreezing cycles of snow with microwave radiometers: The Microwave Alpine Snow Melting Experiment (MASMEX 2002–2003). *IEEE Transactions on Geoscience and Remote Sensing*, 43(11), 2431–2442.
- Mätzler, C. (1996). Microwave permittivity of dry snow. *IEEE Transactions on Geoscience and Remote Sensing*, 34(2), 573–581. doi:10.1109/36.485133
- Mätzler, C., & Wiesmann, A. (1999). Extension of the microwave emission model of layered snowpacks to coarse-grained snow. *Remote Sensing of Environment*, 70, 317–325.
- Marshall, J. L., Weng, F., Lord, S., Riishojgaard, L., Phoebus, P., & Yoe, J. (2005). Recent advances at the joint center for satellite data assimilation.

- Ninth Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS).*
- Pulliainen, J. T., Grandell, J., & Hallikainen, M. T. (1999). HUT snow emission model and its applicability to snow water equivalent retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3), 1378–1390. doi:10.1109/36.763302
- Pulliainen, J., & Hallikainen, M. (2001). Retrieval of regional snow water equivalent from space-borne passive microwave observations. *Remote Sensing of Environment*, 75, 76–85.
- Pulliainen, J. (2006). Mapping of snow water equivalent and snow depth in boreal and sub-arctic zones by assimilating space-borne microwave radiometer data and ground-based observations. *Remote Sensing of Environment*, 101, 257–269.
- Robinson, D. A., Dewey, K. F., & Heim, R. R. (1993). Global snow cover monitoring: an update. *Bulletin of the American Meteorological Society*, 74, 1689–1696.
- Roy, V., Goita, K., Royer, A., Walker, A. E., & Goodison, B. E. (2004). Snow water equivalent retrieval in a Canadian boreal environment from microwave measurements using the HUT snow emission model. *IEEE Transactions on Geoscience and Remote Sensing*, 42(9), 1850–1859. doi:10.1109/TGRS.2004.832245
- Shi, J., Jiang, L., Zhang, L., Chen, K. -S., Wigneron, J. -P., & Chanzy, A. (2005). A parameterized multifrequency-polarization surface emission model. *IEEE Transactions on Geoscience and Remote Sensing*, 43(12), 2831–2841. doi:10.1109/TGRS.2005.857902
- Tedesco, M., Pulliainen, J., Takala, M., Hallikainen, M., & Pampaloni, P. (2004). Artificial neural network-based techniques for the retrieval of SWE and snow depth from SSM/I data. *Remote Sensing of Environment*, 90, 76–85.
- Tedesco, M., Kim, E. J., Cline, D., Graf, T., Koike, T., Armstrong, R., et al. (2006). Comparison of local scale measured and modeled brightness temperatures and snow parameters from the CLPX 2003 by means of a dense medium radiative transfer theory model. *Hydrological Processes*, 20(4), 657–672. doi:10.1002/hyp.6129
- Tjuatja, S., Fung, A. K., & Dawson, M. S. (1993). An analysis of scattering and emission from sea ice. *Remote Sensing Reviews*, 7, 83–106.
- Tsang, L. (1992). Dense media radiative transfer theory for dense discrete random media with particles of multiple sizes and permittivities. *Progress in Electromagnetic Research*, 6(5), 181–225.
- Tsang, L., Kong, J. A., & Ding, K. -H. (2000). Scattering of electromagnetic waves. *Theories and Applications*. New York: Wiley-Interscience 445 pp.
- Tsang, L., & Kong, J. A. (2001). Scattering of electromagnetic waves. *Advanced Topics*. New York: Wiley-Interscience. 432 pp.
- Ulaby, F. T., Moore, R. K., & Fung, A. K. (1986). *Microwave Remote Sensing: Active and Passive, from Theory to Applications*. Boston: Artech House 1120 pp.
- Wiesmann, A., Stozzi, T., & Weise, T. (1996). *Passive microwave signature catalogue of snowcovers at 11, 21, 35, 48 and 94 GHz. IAP Research Report, Vol. 96–8. University of Bern.*
- Wiesmann, A., Mätzler, C., & Weise, T. (1998). Radiometric and structural measurements of snow samples. *Radio Science*, 33(2), 273–289.
- Wiesmann, A., & Mätzler, C. (1999). Microwave emission model of layered snowpacks. *Remote Sensing of Environment*, 70, 307–316.
- Westwater, E. R., Snider, J. B., & Falls, M. J. (1990). Ground-based radiometric observations of atmospheric emission and attenuation at 20.6, 31.65, and 90.0 GHz: A comparison of measurements and theory. *IEEE Transactions on Antennas and Propagation*, 38(10), 1569–1580. doi:10.1109/8.59770

Other Key Reference Papers for Prof. Shi's Lecture:

- [1] Shi, J., L. Jiang, et al. (2006), Physically Based Estimation of Bare-Surface Soil Moisture With the Passive Radiometers, *IEEE Transactions On Geoscience And Remote Sensing*, 44(11): 3145 - 3153
- [2] Shi, J., T. Jackson, et al. (2008), Microwave vegetation indices for short vegetation covers from satellite passive microwave sensor AMSR-E, *Remote Sensing of Environment* 112: 4285 -4300
- [3] Shi, J., L. Jiang, et al. (2005), A Parameterized Multifrequency-Polarization Surface Emission Model, *IEEE Transactions On Geoscience And Remote Sensing*, 43(12): 2831 - 2841

PART II: LAND SURFACE MODELING

This part depicts the development of land surface modeling, proposes some novel skills to characterize the atmosphere-land surface coupling processes, and show application cases in typical regions. Moreover, there is a special focus on the parameter estimation and the uncertainty quantification and analysis for the complex earth system.

Role of land surface processes in monsoon development: East Asia and West Africa

Yongkang Xue,^{1,2} H.-M. H. Juang,³ W.-P. Li,¹ S. Prince,⁴ R. DeFries,^{4,5} Y. Jiao,⁶
and R. Vasic¹

Received 3 March 2003; revised 3 November 2003; accepted 14 November 2003; published 6 February 2004.

[1] Evidence is presented that exchanges of water and energy between the vegetation and the atmosphere play an important role in east Asian and West African monsoon development and are among the most important mechanisms governing the development of the monsoon. The results were obtained by conducting simulations for five months of 1987 using a general circulation model (GCM) coupled with two different land surface parameterizations, with and without explicit vegetation representations, referred to as the GCM/vegetation and the GCM/soil, respectively. The two land surface models produced similar results at the planetary scale but substantial differences at regional scales, especially in the monsoon regions and some of the large continental areas. In the simulation with GCM/soil, the east Asian summer monsoon moisture transport and precipitation were too strong in the premonsoon season, and an important east Asian monsoon feature, the abrupt monsoon northward jump, was unclear. In the GCM/vegetation simulation, the abrupt northward jump and other monsoon evolution processes were simulated, such as the large-scale turning of the low-level airflow during the early monsoon stage in both regions. With improved initial soil moisture and vegetation maps, the intensity and spatial distribution of the summer precipitation were also improved. The two land surface representations produced different longitudinal and latitudinal sensible heat gradients at the surface that, in turn, influenced the low-level temperature and pressure gradients, wind flow (through geostrophic balance), and moisture transport. It is suggested that the great east-west thermal gradient may contribute to the abrupt northward jump and the latitudinal heating gradient may contribute to the clockwise and counterclockwise turning of the low-level wind. The results showed that under unstable atmospheric conditions, not only low-frequency mean forcings from the land surface, such as monthly mean albedo, but also the perturbation processes of vegetation were important to the monsoon evolution, affecting its intensity, the spatial distribution of precipitation, and associated circulation at the continental scale. *INDEX*

TERMS: 1866 Hydrology: Soil moisture; 1833 Hydrology: Hydroclimatology; 3322 Meteorology and Atmospheric Dynamics: Land/atmosphere interactions; 3337 Meteorology and Atmospheric Dynamics: Numerical modeling and data assimilation; *KEYWORDS:* land surface, monsoon, SSiB

Citation: Xue, Y., H.-M. H. Juang, W.-P. Li, S. Prince, R. DeFries, Y. Jiao, and R. Vasic (2004), Role of land surface processes in monsoon development: East Asia and West Africa, *J. Geophys. Res.*, 109, D03105, doi:10.1029/2003JD003556.

¹Department of Geography, University of California, Los Angeles, California, USA.

²Department of Atmospheric Sciences, University of California, Los Angeles, California, USA.

³National Centers for Environmental Prediction, National Oceanic and Atmospheric Administration, Camp Springs, Maryland, USA.

⁴Department of Geography, University of Maryland, College Park, Maryland, USA.

⁵Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland, USA.

⁶Department of Earth and Atmospheric Sciences, University of Quebec at Montreal, Montreal, Canada.

1. Introduction

[2] Monsoons are macroscale phenomena. Differential heating of the land and the ocean, latent heat release into the atmosphere, and planetary rotation are considered to be the factors that determine the strength, duration and spatial distribution of large-scale monsoons [Webster *et al.*, 1998]. Land surface characteristics of the continents have also been suggested to be an important factor in the modulation of the monsoon circulation and surface hydrology [Webster, 1987]. Despite the importance of the monsoon systems in providing water for agriculture in some of the Earth's most populous regions monsoons have not been adequately modeled (see Webster *et al.* [1998] for a comprehensive review) and the role of land surface processes in the systems are still not well understood.

[3] In the case of the Indian monsoon, for example, the role of Eurasian snow cover is not agreed. Although many observational studies [e.g., *Hahn and Shukla*, 1976; *Dey and Bhanu Kumar*, 1982; *Liu and Yanai*, 2002] generally show a negative correlation between Eurasian snow cover and subsequent summer Indian monsoon, modeling studies with general circulation models of the atmosphere (GCM) and regional models [e.g., *Barnett et al.*, 1989; *Yasunari et al.*, 1991; *Vernekar et al.*, 1995; *Douville and Royer*, 1996] indicate that the effects of anomalous snow cover over Eurasia on the Indian monsoon are highly variable (see *Douville and Royer* [1996] for a review). A recent study using observational data has challenged this relationship [*Robock et al.*, 2003].

[4] In addition to snow, a number of studies have explored the roles of other land surface processes and the mechanisms that govern land surface/monsoon interactions in monsoon systems. In an investigation of the relative roles of land surface evaporation and sea surface temperature (SST) on the Asian monsoon [*Lau and Bua*, 1998], it was found that land/atmosphere interactions did not seem to alter basic, planetary-scale features, but local effects over east Asia/Indochina were quite pronounced. In a sensitivity study, *Meehl* [1994] found that stronger Asian summer monsoon were associated with lower surface albedo, greater soil moisture, less snow cover, and greater sea/land contrast. *Douville et al.* [2001] indicated that although African summer rainfall increased with increased soil moisture, there was no response in the Indian subcontinent, which they attributed to the more dynamic and chaotic nature of the Asian monsoon. A GCM simulation of the desertification in Mongolian and Inner Mongolian grassland [*Xue*, 1996] produced negative monsoon rainfall anomalies in northern and southern China and positive rainfall anomalies along the Changjiang (Yangtze) river region, which were generally consistent with observed anomalies. The large reduction in evaporation due to land degradation resulted in less convection and lowered atmospheric heating rates, which was associated with relative subsidence and, in turn, weakened the northward movement of the monsoon flow and lowered the rainfall and evaporation, leading to a positive feedback system.

[5] In this study, we used the National Center for Environmental Prediction's (NCEP) GCM which belongs in the higher hierarchy of numerical models used for climate studies [*Kalnay et al.*, 1990; *Kanamitsu et al.*, 1991]. The Simplified Simple Biosphere model (SSiB) [*Xue et al.*, 1991] was coupled with the NCEP GCM for this study. The simulations with the NCEP GCM/SSiB were compared with those from the NCEP GCM coupled with a land scheme where the biophysical processes were not explicitly parameterized. Using these comparisons, we explored the influence of the soil and vegetation biophysical processes on intra-seasonal monsoon development. This study focuses mainly on the impact of land surface processes on monsoon precipitation. This paper discusses the monsoons in east Asia and West Africa; in another paper we will focus on the Americas.

2. Model Descriptions

[6] The NCEP GCM [*Kalnay et al.* [1990], *Kanamitsu et al.* [1991], <http://sgl62.wwb.noaa.gov:8080/research/>

mrf.html] was used with 28 levels and with T62 horizontal resolution (slightly less than 2 degrees in equatorial and midlatitude areas) for a range of model runs. The effects of using the GCM coupled with a simple two-layer soil model (NCEP GCM/SOIL), as used in the original NCEP GCM, were compared with the GCM coupled with SSiB [*Xue et al.*, 1991] (NCEP GCM/SSiB), a comprehensive soil-vegetation-atmosphere model. The two land parameterization schemes represent land surface processes with two different approaches.

[7] In NCEP GCM/SOIL the ground hydrology was simulated by the soil model, and the distributions of monthly mean vegetation albedo and surface roughness length were separately prescribed on the basis of an existing data set [*Dorman and Sellers*, 1989], which has similar monthly mean values to those used in SSiB, but no explicit biophysical processes are included. Soil temperature and soil volumetric water content were computed in two layers at depths 0.1 and 1.0 m in a fully implicit time integration scheme [*Pan and Mahrt*, 1987]. The lowest atmospheric model layer was the surface layer and the Monin-Obukhov similarity profile relationship was applied to obtain the surface stress and sensible and latent heat fluxes [*Miyakoda and Sirutis*, 1986]. A bulk aerodynamic formula was used to calculate the fluxes once the turbulent exchange coefficients had been obtained. In this approach the land surface properties that regulate land/atmosphere interactions were regarded as separable parameters, which could be independently prescribed as boundary conditions in the GCM for each month.

[8] In NCEP GCM/SSiB the radiative transfer in the canopy was simulated, which produces diurnal variation in surface albedo. There were three soil layers and one vegetation layer. *Deardorff's* [1977] force-restore method was used to predict the surface and the deep soil temperatures. SSiB includes processes such as water interception loss, direct evaporation from bare soil, and canopy transpiration (controlled by photosynthesis), to describe the surface water balance. The aerodynamic resistance controls interactions of heat fluxes between the vegetated surface and the atmosphere. Similarity theory was used to calculate the aerodynamic resistance from the canopy to the reference height. On the basis of the *Paulson* [1970] and *Businger et al.* [1971] equations, a relationship between the Richardson number, vegetation properties, and aerodynamic resistance at the vegetated surface was developed. Many GCMs use *Louis's* [1979] parameterization to calculate the aerodynamic resistance, where the total aerodynamic resistance including both neutral and non-neutral parts is a function of the Richardson number. This implies that only one value of surface roughness length is used for the parameterization. Although this parameterization is simple and easy to use, it does not satisfy the vegetated surface, where the range of values of surface roughness length could be as large as 1 order of magnitude. In SSiB, the resistance of the neutral part is dependent on vegetation and soil properties. In the non-neutral part, a parameterization is related to atmospheric stability conditions and some adjustments based on the vegetation conditions are introduced [*Xue et al.*, 1991, 1996a].

[9] In the NCEP GCM/SSiB model, land surface properties were specified according to vegetation-cover type. A

Table 1. Designations of the Model Runs and the Different Initial and Boundary Conditions Used

Case	Model	Initial Conditions	Land Cover Map
C	NCEP/SOIL	Reanalysis	none
S1	NCEP/SSiB	Reanalysis	NEW SSiB MAP
S2	NCEP/SSiB	Reanalysis and GSWP initial soil moisture	NEW SSiB MAP
S3	NCEP/SSiB	Reanalysis	OLD SSiB MAP

parameter set for each of the vegetation types was used on the basis of a variety of sources [Dorman and Sellers, 1989; Willmott and Klink, 1986; Xue et al., 1996a, 1996b], many of which are invariant with season. Seasonally varying monthly values of some vegetation properties, such as leaf area index (LAI), green leaf fraction, and surface roughness length, were prescribed for most vegetation types or calculated in the model for the crop type [Xue et al., 1996b]. SSiB provided fluxes of momentum, sensible heat and latent heat, radiative skin temperature, visible and near-infrared albedo for both direct and diffuse radiative to the GCM.

3. Experimental Design and Initial and Boundary Conditions

[10] The GCM simulations consisted of five month-long integrations through the boreal monsoon season. Initial conditions were obtained from NCEP/NCAR Global Reanalysis for three dates, 1, 3, and 4 May 1987. The date 2 May was skipped because of errors in the reanalysis data for that day. 1987 was an ENSO year and was 1 of 2 years for which a comprehensive soil moisture data set was available. The results of the three model runs with different initial conditions (1, 3, and 4 May) were averaged. The NCEP GCM/SOIL and the NCEP GCM/SSiB runs are referred to as cases C and S1, respectively (Table 1).

[11] The means of surface albedo for case C and case S1 were very similar during June–July–August (JJA) with the exception of some, mostly dry, areas where SSiB simulated slightly higher values (Figure 1). The 1987 NCEP/NCAR Global Reanalysis [Kalnay et al., 1996; Kistler et al., 1999] (referred to as Reanalysis) was used in both case C and case S1 as the source of initial conditions (atmosphere, soil moisture, and soil temperatures), ocean surface boundary conditions (SST and sea ice), and initial snow depth for all GCM runs, as originally used by NCEP for prediction/forecasting. Comparisons between these two cases indicate the effects of explicit description of biophysical processes in the GCM.

[12] Soil moisture was simulated in both GCM/SSiB and GCM/SOIL during model simulations without nudging. Specified SST and sea ice were updated using the observational data during the simulation. Observational data for verification were from the Climate Prediction Center Merged Analysis of Precipitation (CMAP) [Xie and Arkin, 1997] in which observations from rain gauges were merged with precipitation estimates from satellite system.

[13] The impact of initial soil moisture on the model simulations was studied using soil moisture data from the GEWEX soil wetness project (GSWP) [Dirmeyer et al., 1999]. GSWP is a pilot study intended to produce a soil wetness global data set by using 1987 and 1988 meteorolog-

ical observations and analyses to drive land surface models. SSiB participated in this project and the results produced by SSiB were used for this study. The average of three runs using initial soil moisture from GSWP in the NCEP GCM/SSiB is referred to as case S2 and comparisons between cases S1 and S2 indicate the effects of the different initial soil moisture.

[14] For numerical simulations with the NCEP GCM/SSiB, a global vegetation classification map was used in the coupled surface-atmosphere model to provide land surface conditions required by the SSiB. A 1 km² resolution global land cover map, based on remote sensing [Hansen et al., 2000] (referred to as NEW SSiB MAP) was used in cases S1 and S2. The vegetation map was aggregated to the GCM grid system by grouping the cover types into the 12 SSiB vegetation types [Xue et al., 2001] and selecting the most common type in each T62 cell (Figure 2a). The land cover classes originally used in SSiB were based on the physiognomic classification of Kuchler [1983] and the land use database of Matthews [1984, 1985] (referred to as OLD SSiB MAP, Figure 2b). OLD SSiB MAP was used in case S3.

[15] The most significant differences between NEW SSiB MAP and OLD SSiB MAP were in semi-arid and arid areas (Figure 2). For example, OLD SSiB MAP classified central Asia, including the Tibetan plateau, as desert, which is not appropriate [Shi and Smith, 1992], while in NEW SSiB MAP it was classified as grasslands or shrubs with bare soil. OLD SSiB MAP classified the Sahara desert as bare soil and shrubs with bare soil; NEW SSiB MAP classified it as bare soil only. In addition to central Asia and the Sahara desert, in NEW SSiB MAP, Europe had more cropped area and India's vegetation cover was changed from crops and forests to wooded grassland and small areas of grassland and shrubs. The comparison between cases S1 and S3 allowed comparison of the effects of the land cover maps.

4. Simulation Results

4.1. General Features

[16] The JJA period is the monsoon season for many areas in the Northern Hemisphere. Case C simulated the spatial distribution of JJA precipitation reasonably (Figure 3) with the maximum values in the Inter-Tropical Convergence Zone (ITCZ) and a second peak in the midlatitudes of both hemispheres. The monsoon regions in India, east Asia, Africa, and the Americas were evident. The main deficiencies in the simulation were the rather weak precipitation in the West Pacific (the Mei-Yu or "Plum" rainband of east Asia and southern Japan), too strong precipitation in the east Pacific, and excessively large area of light-precipitation at higher latitudes (Figure 3c).

[17] Case S1 produced very similar spatial distributions of precipitation as case C at the planetary scale (not shown), but there were substantial differences at regional/continental scales (Figure 3d). These were mainly in tropical and subtropical monsoon areas, and in midlatitude and high-latitude continents. For instance, case S1 increased the precipitation in Central America, reduced the precipitation in southern China and India, including oceans nearby, and increased precipitation along the Asian monsoon trough (south and east of the Tibetan Plateau). It also eliminated the excess precipitation in some continental areas and increased the precipitation in West Africa. We next inves-

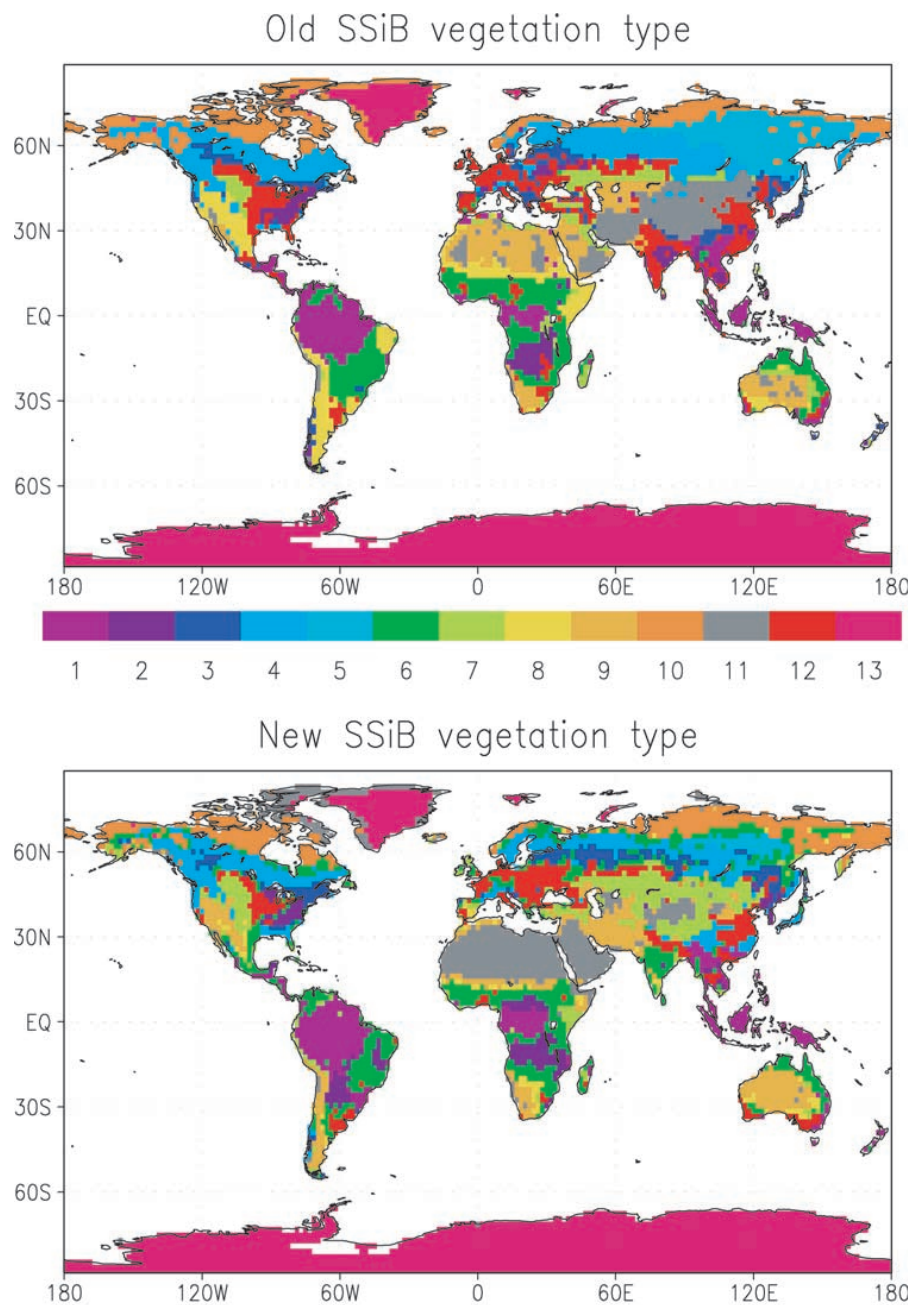


Figure 1. NCEP GCM/SSiB land cover classification map. (top) OLD SSiB MAP; (bottom) NEW SSiB MAP. Type 1, tropical rain forest; type 2, broadleaf deciduous trees; type 3, broadleaf and needleleaf trees; type 4, needleleaf evergreen trees; type 5, needleleaf deciduous trees; type 6, broadleaf trees with ground cover; type 7, grassland; type 8, broadleaf shrubs with ground cover; type 9, broadleaf shrubs with bare soil; type 10, dwarf trees with ground cover; type 11, desert; type 12, crops; type 13, permanent ice.

tigate whether these regional differences were associated with land surface processes.

4.2. East Asian Simulation

[18] The east Asian monsoon covers China, Korea, Japan, Indochina, as well as parts of surrounding countries and

nearby oceans and, together with the Indian monsoon, forms the major part of the Asian Monsoon system [Flohn, 1957; Ding, 1994]. The observed precipitation over east Asia had strong seasonal, interannual, and inter-decadal variations, in particular a dramatic shift from dry condition to wet condition in central eastern China and an opposite

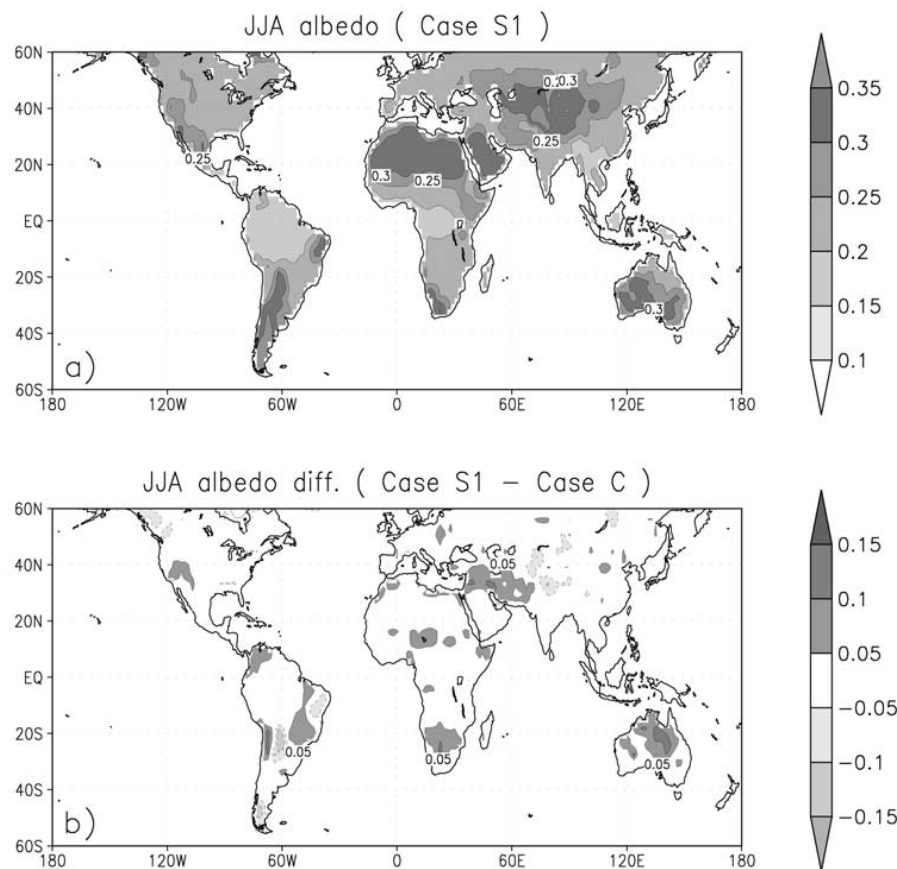


Figure 2. JJA average albedo for (a) case C; (b) case S1–case C.

shift in northern China in the 1970s [Chen *et al.*, 1991; Ding, 1994; Yatsagai and Yasunari, 1994; Yanai and Tomita, 1998; Weng *et al.*, 1999], related to the El Niño-like SST anomalies [e.g., Yang and Lau, 1998; Weng *et al.*, 1999], midlatitude circulation, Indian monsoon [Chen *et al.*, 1991; Yatsagai and Yasunari, 1995], and land surface processes such as snow and land degradation [e.g., Yasunari *et al.*, 1991; Vernekar *et al.*, 1995; Xue, 1996].

4.2.1. Evolution of the East Asian Monsoon

[19] The rainy season in east Asia starts from south of the Changjiang river in April and moves southward to South China in May with light precipitation. This is caused by the confluence of cold air from north and the southwesterly flow from subtropical high and westerly flow from the subtropical region of south Asia, forming a typical subtropical rain belt [Chen *et al.*, 2001]. In mid-May, the tropical monsoon develops in the South China Sea and large amounts of moisture are transported northward into the east Asian continent. Heavy precipitation occurs first in South China and moves northward, indicating a pre-Mei-Yu season. The CMAP precipitation for May 1987 (Figure 4a) indicated a precipitation center in the coastal region of South China with a northeast-southwest precipitation belt. The 4–6 mm d⁻¹ isohyet was located slightly north of the Changjiang river.

[20] Case C simulated the south-north gradient of the spatial distribution of the precipitation in the pre-Mei-Yu

season. The rainfall maximum was centered on the Changjiang river with a south-north precipitation belt (Figure 4b). The 4–6 mm d⁻¹ isohyet extended much farther to the north, almost reaching Inner Mongolia. Compared to case C, the spatial distribution of the precipitation in case S1 was shifted to the south (Figure 4c). The 4–6 mm d⁻¹ isohyet was located to the south of the Yellow River, and the center of the maximum precipitation was located to the south of the Changjiang river with a northeast-southwest precipitation belt. The observed, case C, and case S1 average precipitation over 110°E ~ 120°E and 20°N ~ 40°N was 5.96, 6.51 (±0.5), and 5.74 (±0.4) mm d⁻¹, respectively. The results for case S2 in Figure 4d will be presented in Section 5.

[21] We focused on the results from the three case means to minimize the effects of spurious results and standard deviations are given for all the results presented in the paper (in parentheses). To check the reliability of the results, we also compared the precipitation patterns in Figure 4 with those for each pair of runs in case C and case S1 and found they were very similar. In all three initial conditions, the positions of the 2–4 mm d⁻¹ isohyets in case C were 5 to 10 degree to the north compared to those in case S1. The monthly mean precipitation over 110°E ~ 120°E and 30°N ~ 40°N (north of the Changjiang river) was 6.48, 5.1, and 7.08 mm d⁻¹ for runs in case C and 4.15, 3.25, and 4.44 mm d⁻¹ for runs in case S1, which indicated a

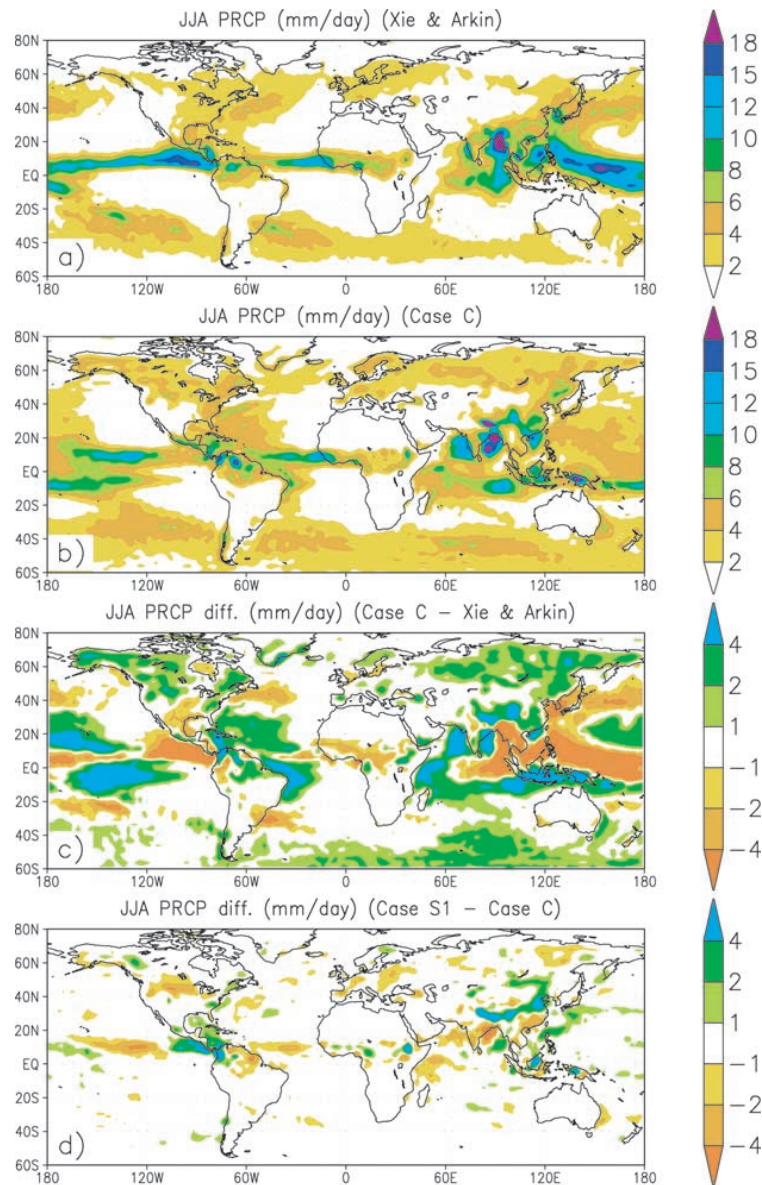


Figure 3. JJA 1987 precipitation for (a) CMAP; (b) case C; (c) case C–CMAP; (d) case S1–case C (mm d^{-1}).

consistently farther north extension of precipitation in case C. The precipitation was 2.98 mm d^{-1} for observation over that area. The highest rainfall in case S1 for this area was lower than the lowest rainfall in case C. This showed that the differences between cases S1 and C were significant.

[22] Both case C and case S1 produced spuriously heavy rainfall in a region centered at 104°E and 33°N . Because the Loess Plateau is to the north of this area and the Sichuan Basin is to the south, there is a steep topographic gradient and the complex regional topography may have contributed to this simulation error. This spurious simulation of precipitation also occurred in the NCEP Reanalysis, and fifth generation Penn State University/NCAR Meso-scale Model (PSU/NCAR MM5) [Grell *et al.*, 1994] with 50 km horizontal resolution (W. Li, personal communication, 2002).

[23] The east Asian premonsoon and monsoon evolution during the rainy season is illustrated by the zonally averaged, 10-day mean precipitation between 105°E and 120°E from May through September (Figure 5). The observed time evolution of the May–September precipitation in 1987 (Figure 5a) was similar to that of Chen *et al.* [2001, Figure 2a] and Lau *et al.* [1988, Figure 7], both of which are the means of observational data, 1961–1995 and 1950–1979, respectively.

[24] Intense precipitation originated around 27°N in early May and by late May had moved southward to about 22°N (Figure 5a, solid arrow). In June the heavy rain moved abruptly northward (dashed arrow) and another precipitation maximum appeared to the north of 30°N . This point marks the start of the Mei-Yu rains. This development is an

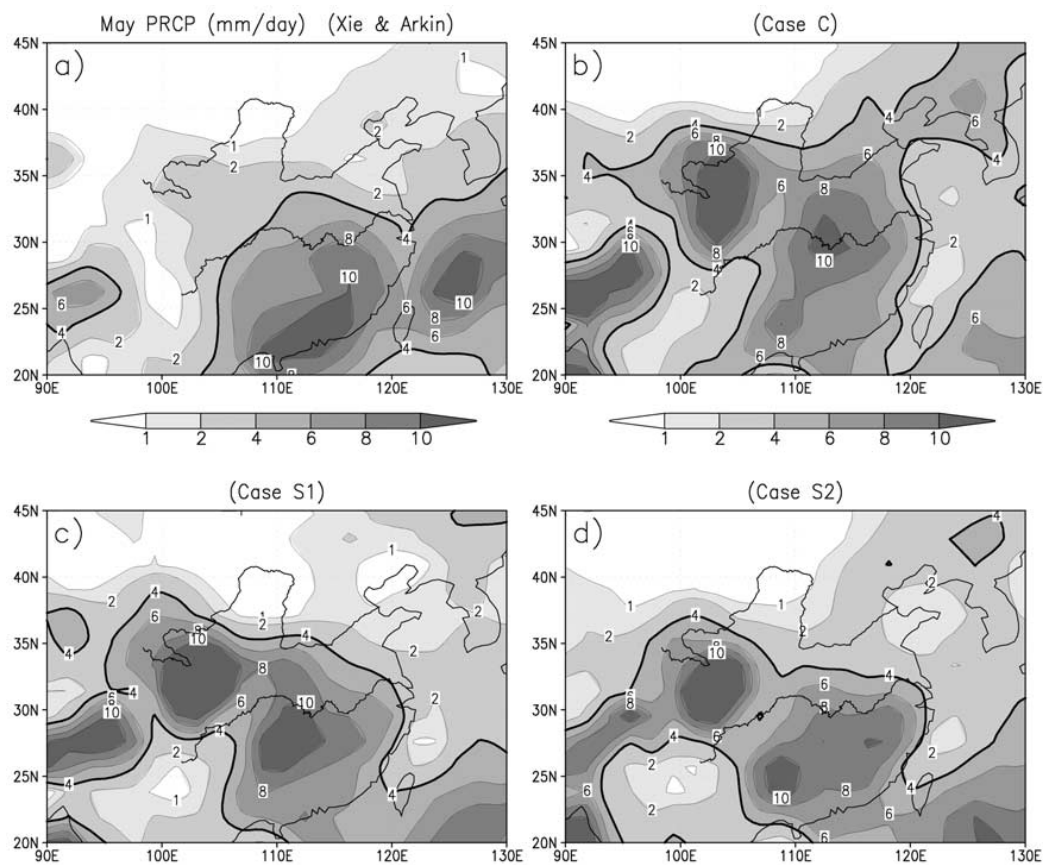


Figure 4. May 1987 precipitation for (a) CMAP; (b) case C; (c) case S1; (d) case S2 (mm d^{-1}).

important signature of the east Asian monsoon, and is referred to as the abrupt monsoon northward jump and has been described in numerous studies [e.g., Lau *et al.*, 1988; Chen *et al.*, 1991; Ding, 1994]. According to climatological data [e.g., Lau *et al.*, 1988], the monsoon rain expands farther into northern China (about 40°N) from July to August and initiates another jump, leading to the start of the monsoon season in northern China. In 1987, this second jump was not clear and there was only a gradual expansion of the rainfall band. There was a relatively dry area between 25°N and 30°N during part of July and August. The monsoon then retreated southward in late August and early September (solid arrow).

[25] There are clear differences between cases C and S1 simulations of the evolution of the monsoon (Figures 5b and 5c). Case C correctly simulated the rainy season, the south-north precipitation gradient, as well as the rainfall peak in July between 20°N and 25°N (Figure 5b). However, the monsoon evolution process was unclear. In addition to the overly extended precipitation in May as discussed above, there was only one persistent wet season during the entire period with the maximum precipitation located around 23°N (coast of South China). This pattern persisted in all three runs in case C. Furthermore, the simulated rain was more intense than that indicated by the observations. Case S1, on the other hand, simulated the features of the monsoon evolution and captured the northward jump. The northward

jump, however, started about 10d earlier and extended over a slightly longer period (Figure 5c, dashed arrow). The rainfall maxima around 22°N and 32°N were simulated, but expanded to northern China too early. The intensity was also stronger than observed (Figure 3d). Overall, however, the evolution processes were simulated, including the dry area between 25°N and 30°N during July and August. All three runs in case S1 were consistent, but with slightly different dates for the start of the abrupt northward movement and their durations. The results for case S2 and case S3 in Figure 5 will be discussed in sections 5 and 6, respectively.

4.2.2. Physical and Dynamic Mechanisms of Land Surface and Atmospheric Effects

[26] The differences between cases C and S1 were caused only by the different parameterizations of land surface processes. These affect the water and energy balances on land surface and then the atmosphere through land/atmosphere interactions. In May, southwest airflow brought moisture to the southern part of China and formed a cyclone-like system according to Reanalysis (Figure 6a). The strong convergence zone at 850 hPa to the south of the Changjiang river was consistent with the observed precipitation. To the north of 35°N , westerly winds and divergence prevailed. The bold lines in the figure show the locations where the zonal wind was zero for a better view of the circulation patterns. Case C produced a convergence

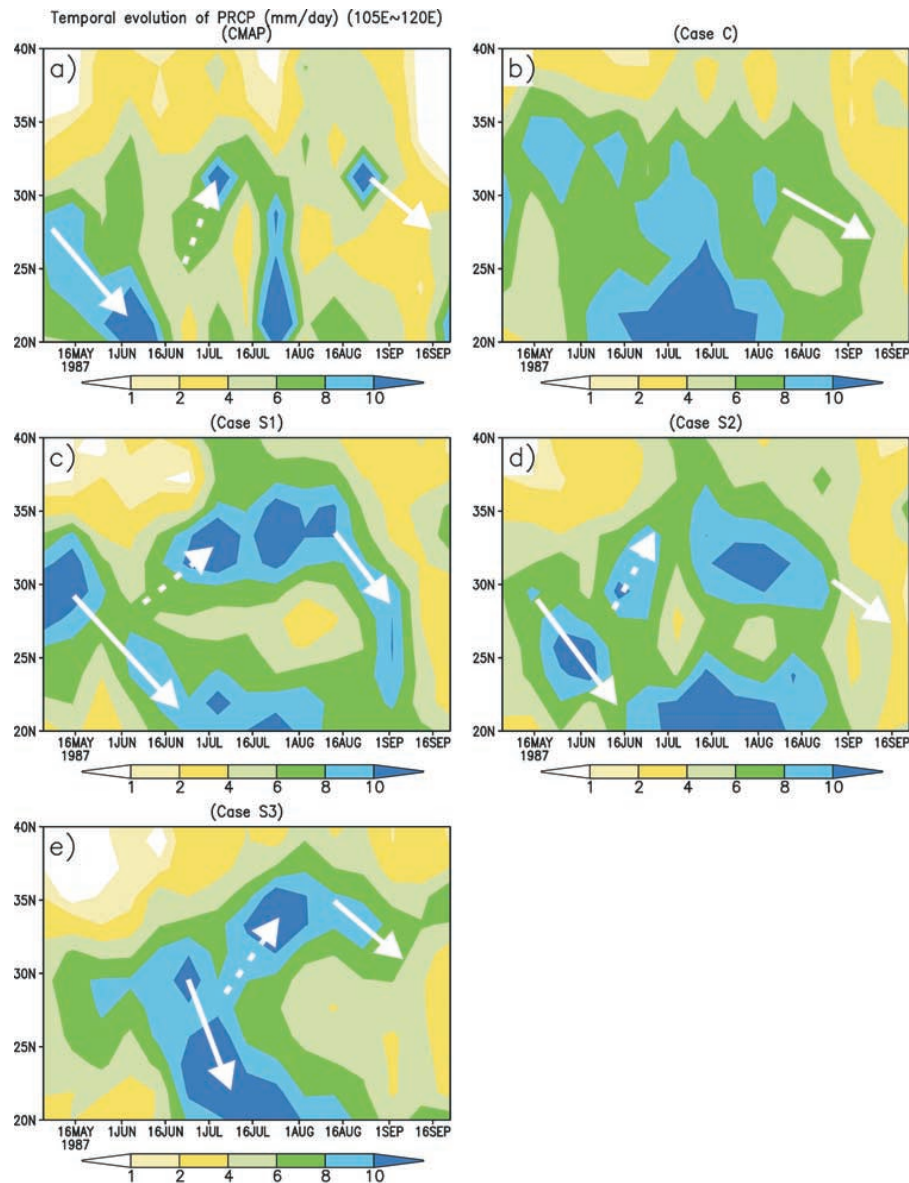


Figure 5. Temporal evolution of the 10-day mean precipitation (mm d^{-1}) averaged over 105° – 120°E from May through September. (a) CMAP; (b) case C; (c) case S1; (d) case S2; (e) case S3.

band to the south and a divergence band to the north in east Asia (Figure 6b), however the southwesterly flow was unrealistically strong, pushing the convergence zone farther to the north. The cyclonic flow was too weak and the easterly wind between 30°N and 35°N almost disappeared. Case S1 produced circulation and associated convergences and divergences that were closer to the Reanalysis (Figure 6c). These differences in wind fields were consistent with those in precipitation (Figure 4). Each run in case C and case S1 gave very similar patterns.

[27] Many studies have investigated the mechanisms responsible for the atmospheric circulation. It has been found that the circulations in summer subtropics seem to be more related to thermal forcing, and the formation

mechanism is more complicated compared with other latitudes [Hoskins, 1987]. In this study, it was found that circulations in May and June had the largest differences between two cases. To understand the causes, we analyzed the differences in surface heating. Table 2 shows the May average upward heating differences between case S1 and case C over the east Asian continent, and indicates the difference in sensible heat flux was dominant in the upward heating components. The May sensible heat flux, 850 hPa geopotential height, and the differences between cases C and S1 (Figure 7) showed that in case C, the east Asian area to the south of 30°N was a heat sink and the area to the north of 30°N was a source. In contrast, in case S1, the entire east Asian continent was a heat source. Therefore

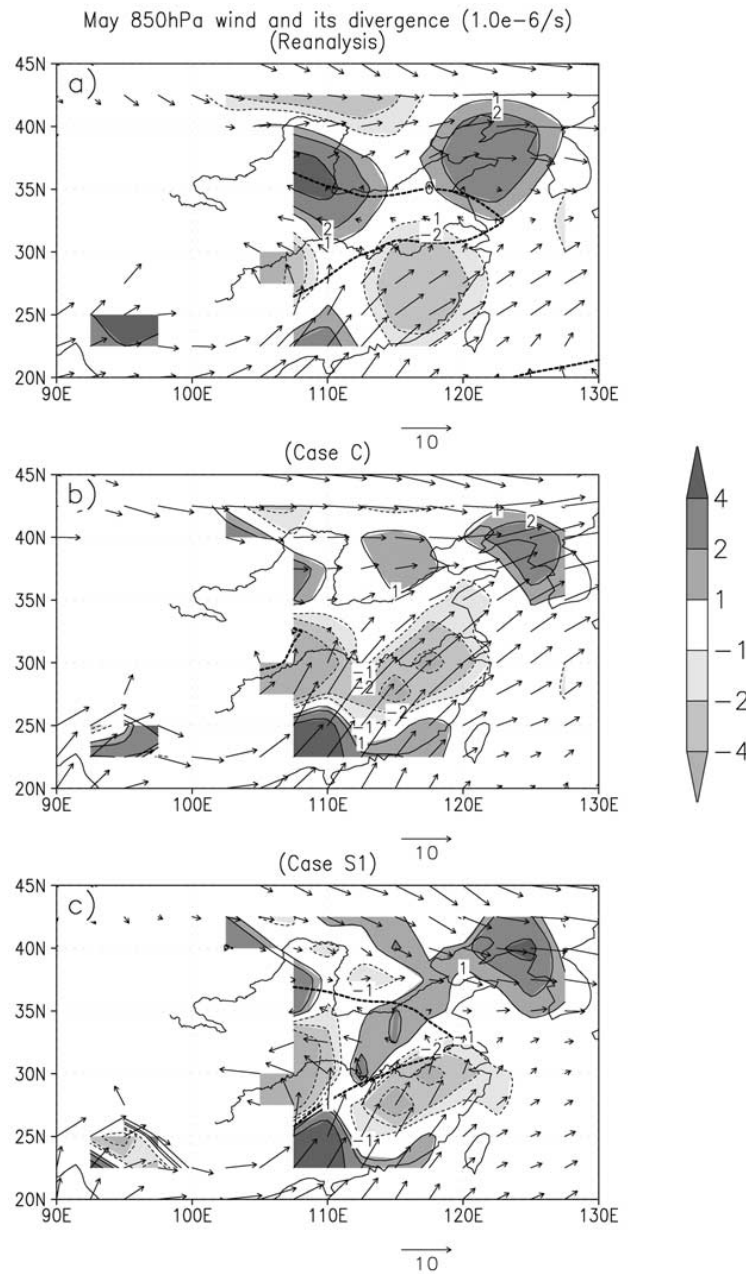


Figure 6. May 1987 wind field (m s^{-1}) and divergence ($10\text{e-}6 \text{ s}^{-1}$) at 850 hPa (a) Reanalysis; (b) case C; (c) case S1. To clarify the circulation patterns, the bold lines show the locations where the zonal wind was zero.

the gradients of geopotential height in southeastern China were stronger in case C than in S1 (Figures 7a and 7b), consistent with the large heating gradient, which produced strong southwesterly in case C (Figure 6b).

[28] The counterclockwise turning of the low-level flow between 25°N and 32°N (Figure 6c) was consistent with the pressure gradient difference and the associated easterly wind anomaly between cases S1 and C in that region. Case S1 produced relative lower pressure to the south and relative higher pressure to the north (Figure 7c). An anom-

alous eastward wind in case S1 would be produced while Coriolis forcing balanced the pressure gradient force based on geostrophic balance. In June, case S1 still produced the counterclockwise turning while case C did not, because of the same cause as in May (not shown). In July and August, while the monsoon was mature, both cases S1 and C produced the turning. In another study [Wu and Liu, 2003] the July Reanalysis from 1980 to 1997 was used to analyze the relationship between circulation and boreal summer subtropical heating, which included vertical distri-

Table 2. Monthly Mean Surface Upward Heating Fluxes and Low-Level Air Temperature^a

	Latent Heat Flux	Sensible Heat Flux	Short Wave Up	Long Wave Up	T (925 hPa)	T (850 hPa)
May, case S1–case C	−15.9	26.9	−9	6.5	0.33	0.87
June, case S1–case C	−9.4	35	−7.7	16.2	2.54	2.08

^aHeating fluxes are given in W m^{-2} , and temperatures are given in $^{\circ}\text{C}$ (25° – 35°N , 110° – 120°E).

butions of long-wave radiative cooling, sensible heating and condensation heating in atmosphere. Condensation heating was identified as the main heating source in the east Asian lowland, which influenced the July monsoon. Our study reveals the role of vegetation processes, espe-

cially the spatial distribution of the surface heating in the processes, in the early stages of monsoon.

[29] Figure 8 shows the differences in JJA mean precipitation, vertically integrated moisture flux and its divergence, and evaporation between cases S1 and C. The

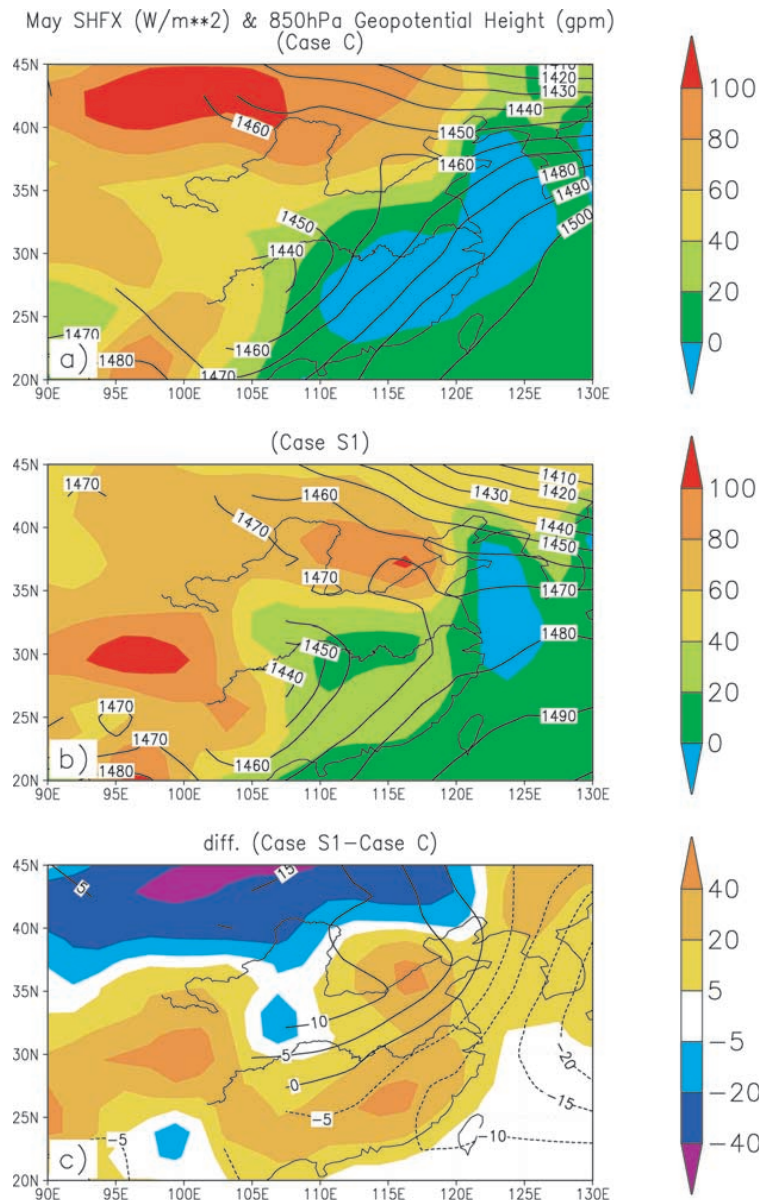


Figure 7. May sensible heat flux (W m^{-2}) and 850 hPa geopotential height (gpm). (a) Case C; (b) case S1; (c) case S1–case C.

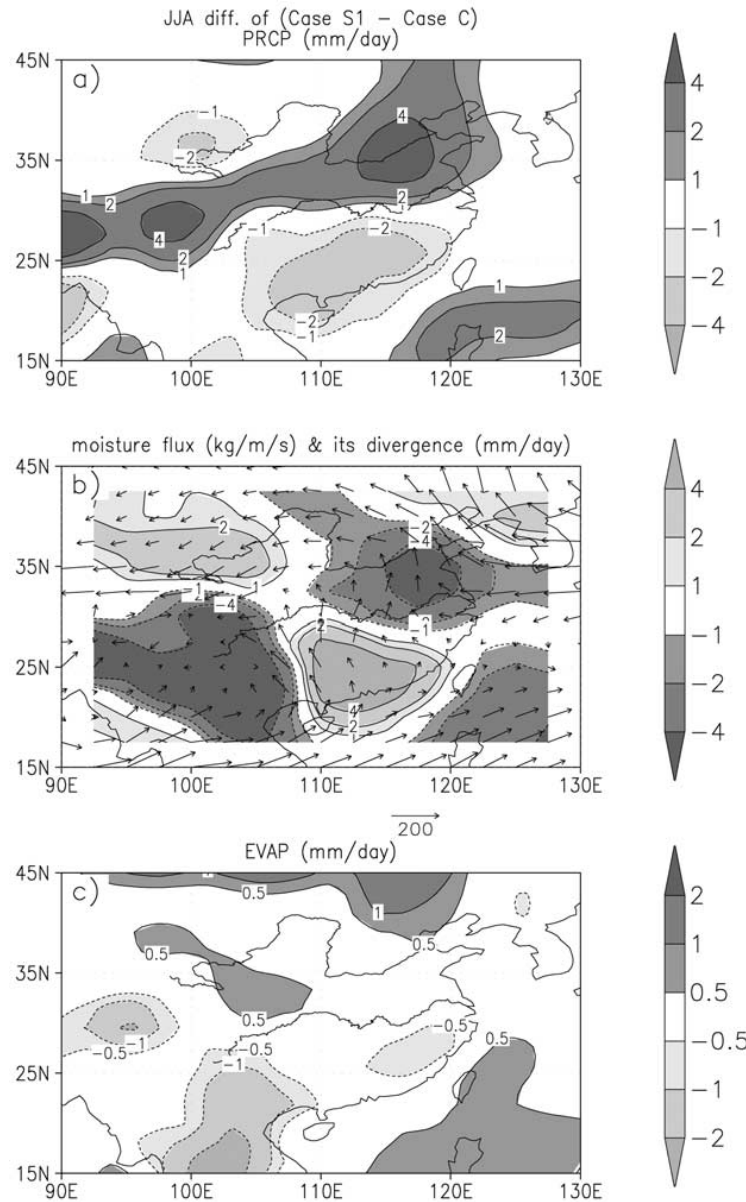


Figure 8. JJA mean differences between case S1 and case C (a) precipitation (mm d^{-1}); (b) Vertically integrated moisture flux ($\text{kg m}^{-1} \text{s}^{-1}$) and divergences (mm d^{-1}); (c) evaporation (mm d^{-1}).

northeast-southwest band with positive precipitation difference between 45°N and 25°N in Figure 8a was consistent with stronger moisture flux convergence (Figure 8b). The lower precipitation to the south of the Changjiang river and west of the Yellow river was located in the divergence areas (Figure 8b). Most changes in evaporation only appeared in the Indochina Peninsula and north of the Yellow River (Figure 8c). This indicates that the main differences in the monsoon precipitation simulation between cases C and S1 were related more closely to moisture-divergence field, the patterns of which were similar to the low-level wind field (not shown), rather than surface evaporation. Therefore it is necessary to examine the influence of land surface processes

on the circulation to understand the evolution of the monsoon rainfall.

[30] The latitudinal and longitudinal means of several variables were evaluated to examine how land surface processes affected the northward jump. It was found that cases C and S1 had large differences in northward low-level moisture transport. Figure 9 shows the 10-day mean of 925 hPa specific humidity zonally averaged over 105°E and 120°E for cases C and S1. In early May, the northward transport of the moisture in case C was stronger and produced relatively wetter condition than case S1, as discussed in section 4.2.1 (Figure 4b). In the late part of June, the moist region in case S1 had a dramatic northward expansion

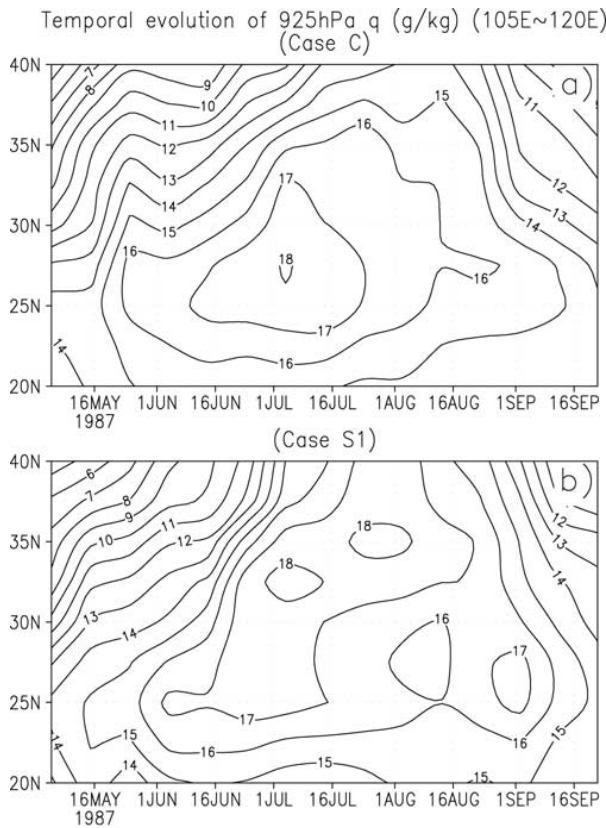


Figure 9. Temporal evolution of the 10-day mean specific humidity (g kg^{-1}) at 925 hPa averaged over $105^\circ\text{--}120^\circ\text{E}$ from May through September. (a) Case C; (b) case S1.

between 25°N and 40°N , which was consistent with the timing of the northward jump of the precipitation band (Figure 5c) and provided the necessary moisture. Case C, on the other hand, did not show such dramatic northward expansion. The maximum humidity (17 g kg^{-1} contour line) was confined to the south of the Changjiang river (at 30°N). Further analysis showed the differences in moisture fields were consistent with the differences in simulated meridional wind. The meridional wind at 925 hPa in case S1 also increased dramatically in the late part of June (Figure 10b), when the northward expansion of wet region occurred (Figure 9b). The 5 m s^{-1} contour line reached around 37°N . Case C, on the other hand, did not have an increase in meridional wind during June (Figure 10a). The 4 m s^{-1} contour line was confined to around 30°N . Since northward transport of the water vapor was the main moisture source of the east Asian summer monsoon (Figure 6), the differences shown in Figures 9 and 10 would have great impact on monsoon development. In the following, we further explore how land surface processes contribute to these differences.

[31] The abrupt northward jump in the east Asian monsoon is an important feature. Thus far there have only been theoretical studies of its cause, based on a quasi-geostrophic vorticity equation for a barotropic dissipative system with thermal forcing to investigate the mechanism of abrupt change of equilibria state [Liu and Tao, 1983; Miao and

Ding, 1985; Wang, 1986]. These studies found that seasonally varying thermal forcing and interactions between thermal forcing and nonlinear motion of atmosphere under certain geographic conditions could produce abrupt changes in atmospheric circulation, but the abrupt change would not occur under weak meridional or zonal thermal gradients.

[32] Enlightened by these studies, we examined the surface heating sources to understand the role of land surface processes in the northward abrupt jump of rainfall band. The low-level temperature in case S1 in June was higher than in case C (Table 2), which would enhance the land-sea temperature gradient since the temperatures over the ocean were the same for the two cases (not shown). Table 2 shows the differences in radiative heating were not large, and case S1 had slightly less latent heat release from the surface. The major differences were in the sensible heat flux as in May. On the basis of the geostrophic balance in midlatitude, the northward wind (shown in Figure 10b) should be produced by the east-west pressure gradient. Among surface upward heating components, only sensible heat fluxes exhibited a clear east-west gradient.

[33] Figure 11 shows the 21–30 June mean sensible heat flux and geopotential height and the differences in 850 mbar. There was no clear gradient of sensible heat flux between the eastern part of the east Asian continent and the Pacific

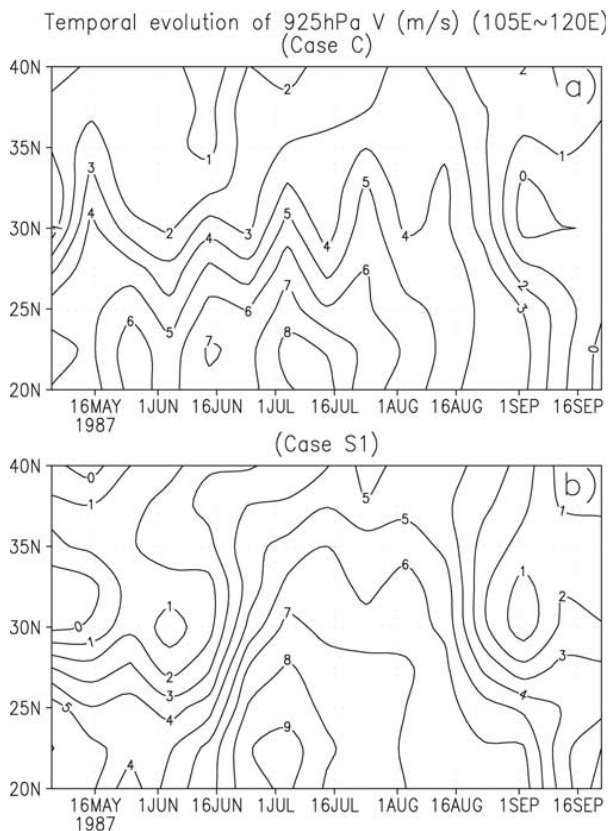


Figure 10. Temporal evolution of the 10-day mean meridional wind (m s^{-1}) at 925 hPa averaged over $105^\circ\text{--}120^\circ\text{E}$ from May through September. (a) Case C; (b) case S1.

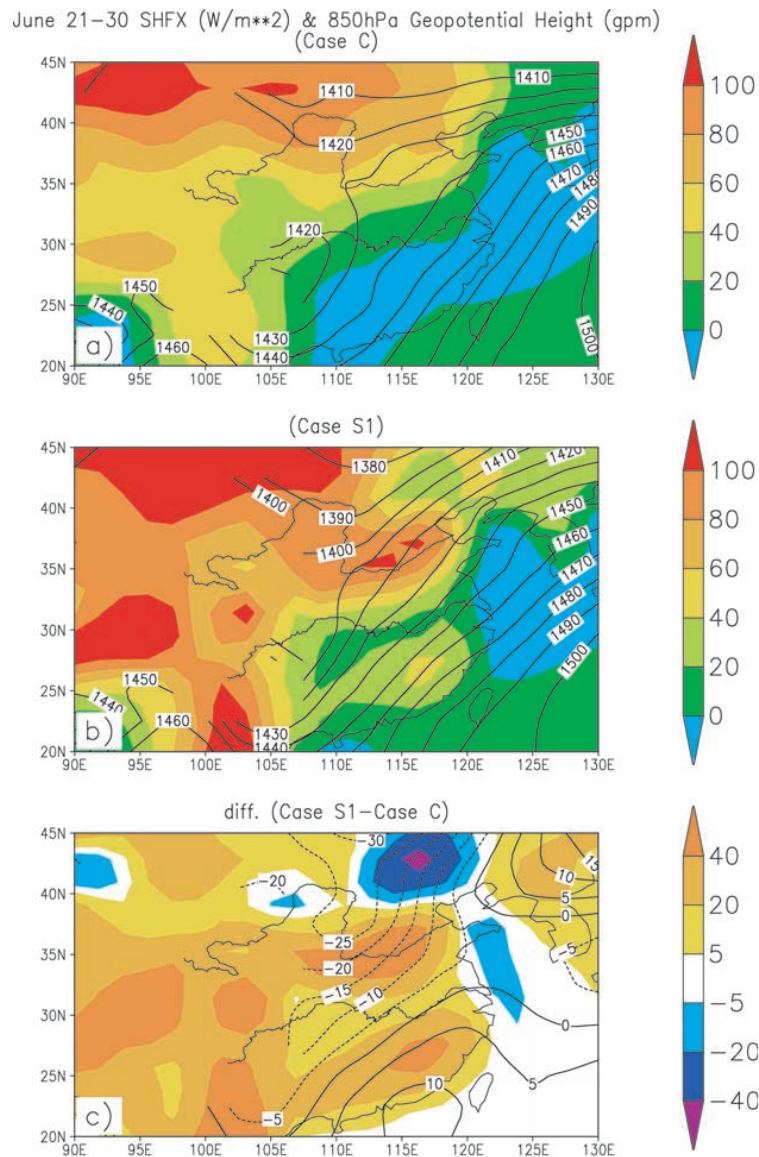


Figure 11. The 21–30 June sensible heat flux (W m^{-2}) and 850 hPa geopotential height (gpm). (a) Case C; (b) case S1; (c) case S1–case C.

Ocean in case C (Figure 11a). In fact, many parts of the land were heat sinks during June. In contrast to case C, S1 produced substantial east-west gradients of sensible heat flux between land and ocean, which in turn produced greater temperature gradients and consequent pressure gradients in the lower atmosphere (Figure 11b). The additional northward meridional wind was evident (Figure 11c). On the basis of these analyses and previous theoretical studies, we suggest that the greater east-west thermal gradient, which produced strong northward transport of moisture and a cyclone condition, may contribute to the abrupt northward jump of the monsoon.

4.3. West African Simulation

[34] The West African monsoon is relatively weak compared to the Asian monsoon [Griffiths, 1972; Nicholson,

1976]. The similarity of the climate in the east-west direction contrasts dramatically with the strong North-South gradient. The relationship between SST and seasonal to interannual rainfall variations in the Sahel region has long been discussed. Several observational and modeling studies have suggested that the Atlantic SST anomalies and global SST anomalies play important roles in producing rainfall anomalies over the Sahel and the adjoining regions [e.g., Lamb, 1978; Hastenrath, 1984; Lamb and Pepler, 1991; Folland et al., 1991; Palmer et al., 1992; Rowell et al., 1995]. Meanwhile, the role of biophysical feedbacks in the Sahel region has also been examined [e.g., Charney et al., 1977; Walker and Rowntree, 1977; Sud and Fennessy, 1982; Laval and Picon, 1986; Kitoh et al., 1988; Wang and Eltahir, 1999]. These studies consistently demonstrated impacts of land surface conditions on the climate of the

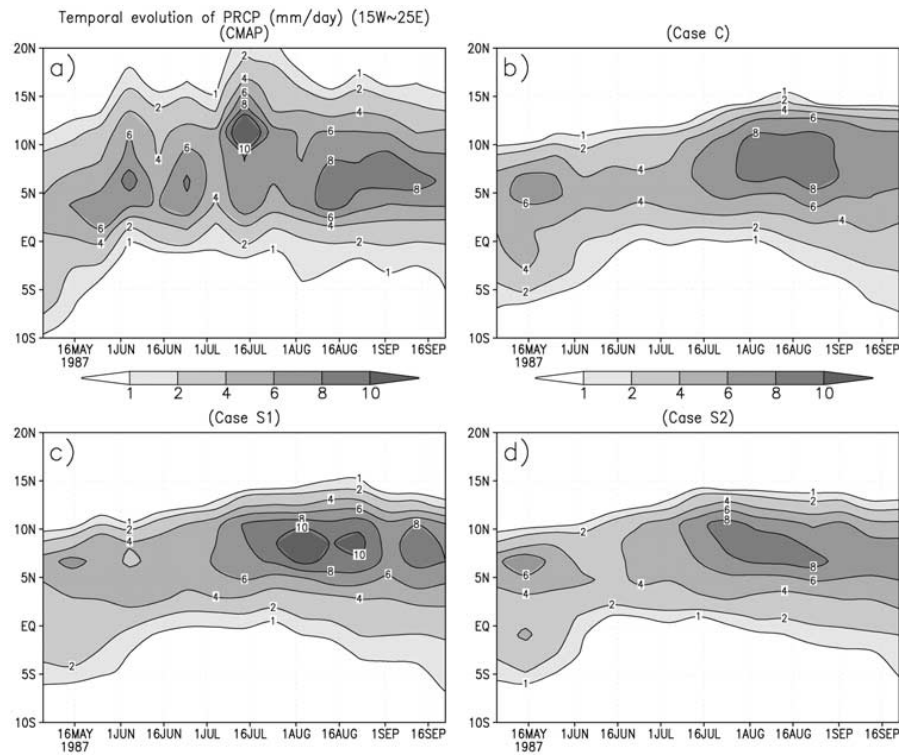


Figure 12. Temporal evolution of the 10-day mean precipitation (mm d^{-1}) averaged over 15°W to 25°E from May through September. (a) CMAP; (b) case C; (c) case S1; (d) case S2.

Sahel. Furthermore, biophysical models coupled with atmospheric models [e.g., Xue and Shukla, 1993; Xue, 1997; Clark *et al.*, 2001; Wu *et al.*, 2002; Xue *et al.*, 2003] have explored the role of land degradation in decadal Sahelian regional climate anomalies, including anomalous precipitation, higher surface temperature, lower river runoff, and the mechanisms responsible for the extended Sahel drought.

[35] In this section we present the results for the impact of two land surface parameterizations on the monsoon evolution and atmospheric circulation in northern Africa. Despite improvements in simulated total precipitation during the five-month simulation in case S1 (see Figure 3 for JJA mean) the processes of zonally mean monsoon evolution for the central and western Africa were not substantially different in cases C and S1 (Figure 12). In both the monsoon moved north in May and reached a maximum in August, but the intraseasonal variations were not well simulated. Sultan and Janicot [2000] identified a northward jump in the African monsoon from 5°N in May–June to 10°N in July–August (Figure 12a), which they attributed to African easterly waves and topographic effects. The models used here only produced a weak rainfall high in May, and showed no clear northward jump possibly because of limitation of the horizontal resolution. The May and June 1987 oscillation was a single year event, mainly because of internal variability (S. Janicot, LMD, personal communication, 2003), and was not simulated.

[36] Numerous studies have demonstrated the sensitivity of the Sahel regional climate to the land surface condition as mentioned above. Cases C and S1 had similar monthly

mean albedo, surface roughness length, and initial soil moisture. Albedo and surface roughness are the two most important land parameters influencing the climate in the Sahel region [Xue *et al.*, 1997] and, unlike the midlatitudes, the interaction between one land parameter and atmosphere may be more important than multiple interactions [Niyogi *et al.*, 2002]. The similar monthly mean surface albedo and roughness lengths in cases C and S1 could explain why there was little difference in the simulated evolution of zonal mean precipitation in the Sahel. However, the differences in spatial distributions of simulated circulation as well as precipitation were still evident.

[37] The effects of land surface processes were manifested in wind fields and divergence. In the Reanalysis (Figure 13a), the southeasterly airflow from the Indian Ocean at 850 hPa became southwesterly after crossing the equator in central Africa, and formed a northeast-southwest convergence band between 10°E and 35°E , and between 20°S and the equator. The confluence of the southwesterly and northeasterly airflows formed another convergence zone over the Sahel, which was relevant to the summer monsoon in the region. Both cases C and S1 simulated the Sahel convergence zone well, although with different intensity (Figures 13b and 13c). Case C, however, failed to simulate a southwest-northeast convergence band from 10°E to 30°E because the turning of the southeasterly from the Indian Ocean was not as strong as in the Reanalysis (Figure 13a). In fact there was a stronger heating source along 10°N between 20°E and 40°E in case S1 (not shown). After crossing the equator, the heating induced airflow

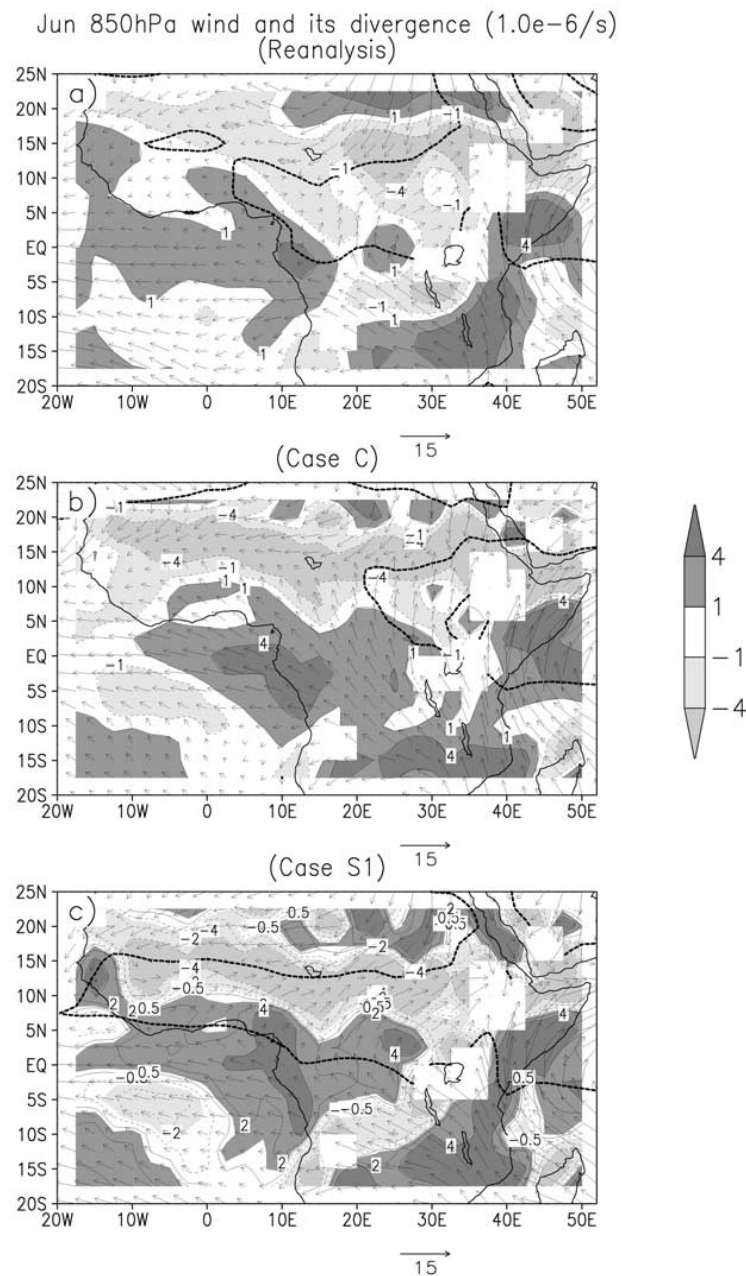


Figure 13. June 1987 wind field (m s^{-1}) and divergence ($10\text{e}-6 \text{ s}^{-1}$) at 850 hPa (a) Reanalysis; (b) case C; (c) case S1.

turned clockwise. This mechanism was consistent with the east Asian simulation (Figure 6). This heating source was missing in case C. The differences in wind fields also existed at 500 hPa. Convergence prevailed in central Africa in case C, which was opposite to that in the Reanalysis and case S1. The differences between case C and case S1 at 850 hPa and 500 hPa were also found in May and July.

[38] In addition to the circulation, cases C and S1 differed in rainfall intensity. In east Asia the differences in precipitation generally correlated with the changes in moisture flux (Figure 7) and it is interesting to examine the same

relationship in Africa. We select the differences of these variables in June and August to exhibit the extremes (Figures 14 and 15). In June, although the evaporation reduction in East Africa and the coastal area of West Africa may have contributed to the precipitation decrease, the major rainfall change in the Sahel was consistent with the changes in moisture flux. Case S1 produced stronger moisture convergence and rainfall in the Sahel. A stronger moisture divergence in the coastal area and lower evaporation also contributed to the rainfall reduction (Figures 14a and 14b). In August the effect of evaporation prevailed

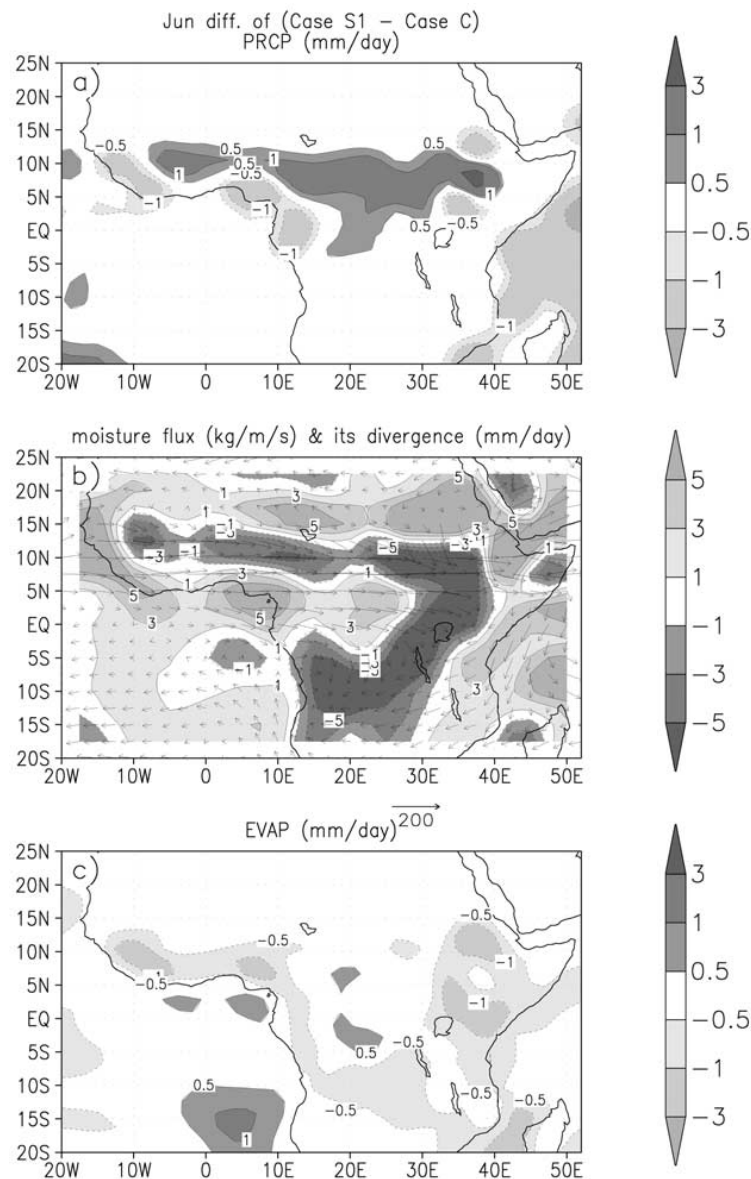


Figure 14. June differences between case S1 and case C (a) precipitation (mm d^{-1}); (b) vertically integrated moisture flux ($\text{kg m}^{-1} \text{s}^{-1}$) and divergences (mm d^{-1}); (c) evaporation (mm d^{-1}).

(Figure 15) and contributed to the large rainfall reduction in the Sahel, consistent with Xue [1997]. Compared with the east Asia, land surface evaporation played a more important role in the variation of West African monsoon. In addition to the Sahel region, the wind field and moisture convergences in case S1 and C also differed in central Africa (Figures 13c, 14b, and 15b), but there was not much precipitation in central Africa during the monsoon season and these differences did not affect the precipitation simulations there (Figures 14a and 15a).

5. Impact of Initial Soil Moisture

[39] Soil moisture is an important surface variable affecting the surface water and energy balances. Case S2

explored the effect of using the GSWP soil moisture as the initial condition for the NCEP GCM/SSiB (Figure 16). In general the soil in GSWP was drier than that in the Reanalysis except in West Africa, India and Bangladesh, East China, northwest South America, and southwest Australia (no GSWP data for Greenland and the Antarctic).

[40] It is clear that the JJA changes in precipitation and soil moisture were positively correlated, but the changes in soil moisture did not necessarily lead to changes in precipitation (Figure 17). Case S2 reduced the extra precipitation over the large continents and enhanced the Indian monsoon, but it was slightly dry in Africa (Figure 12d). Dirmeyer [2000] specified soil moisture during his entire model integration for 1987 and 1988, and found an improvement

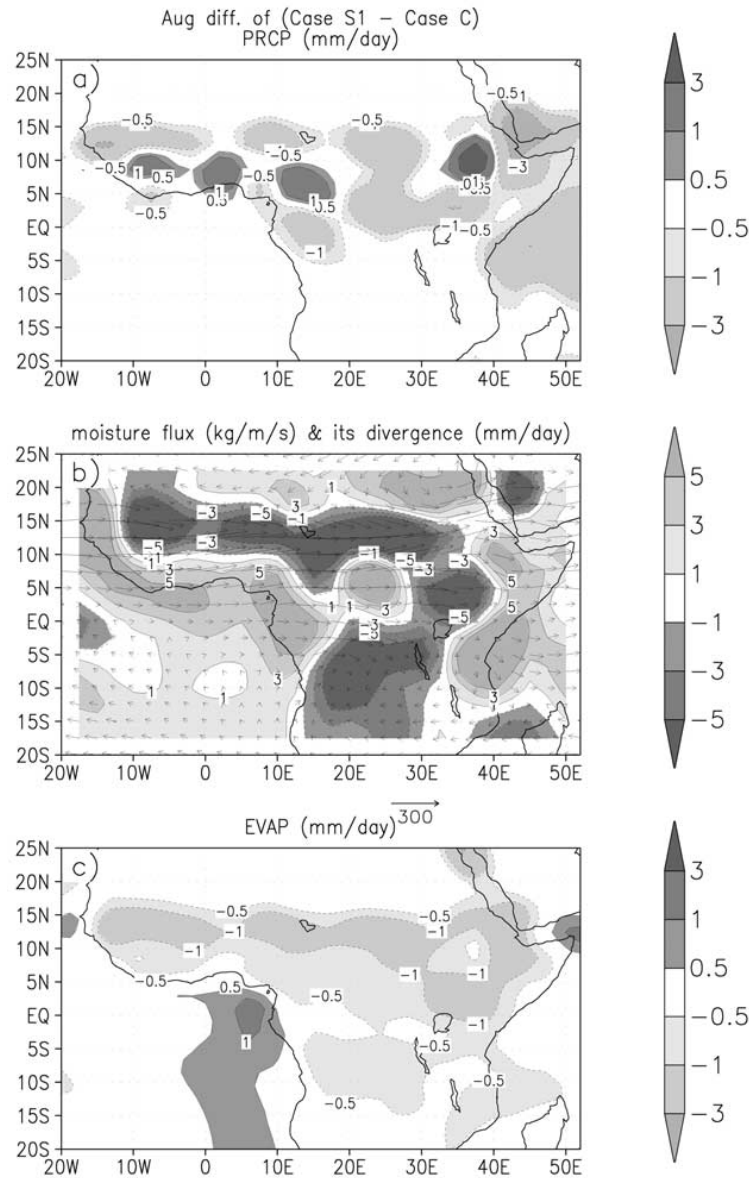


Figure 15. August differences between case S1 and case C (a) precipitation (mm d^{-1}); (b) vertically integrated moisture flux ($\text{kg m}^{-1} \text{s}^{-1}$) and divergences (mm d^{-1}); (c) evaporation (mm d^{-1}).

in the simulation of the pattern of precipitation globally and regionally, especially in monsoonal Asia.

[41] In this study the improvement in precipitation simulation in south Asia and east Asia was substantial. The year 1987 was an anomalously dry year over India. A counter-monsoon circulation anomaly at low level, associated with weaker Somali jet and Arabian Sea circulation, contributed to this summer drought [Krishnamurti *et al.*, 1989]. Cases C, S1, and S2 failed to catch the special features of the Indian monsoon for this year, but cases S1 and S2 still showed some improvements in monthly and seasonal means. For the Indian monsoon area (70°E to 85°E , 10°N to 25°N), the JJA precipitation was 6.2, $8.7 (\pm 0.4)$, $6.7 (\pm 0.2)$, and $7.6 (\pm 0.8)$ mm d^{-1} for CMAP, cases C, S1, and S2, respectively. For

the same area, but with land only, the JJA precipitation was 5.6, $6.7 (\pm 0.5)$, $4.6 (\pm 0.4)$, and $5.4 (\pm 0.8)$ mm d^{-1} , respectively. Case S1 produced the best simulation for the Indian monsoon as a whole. Case S2 had the best simulation for the Indian monsoon over land, but the standard deviations were large.

[42] The improvement in the east Asian simulation by case S2 was substantial, producing precipitation that was very similar to the May observation, with most heavy precipitation to the south of the Changjiang river and highest values near the coast of South China (Figure 4d). Case S2 also simulated the major features of the east Asian monsoon evolution processes (Figure 5d), with some differences in detail and intensity from case S1, providing further

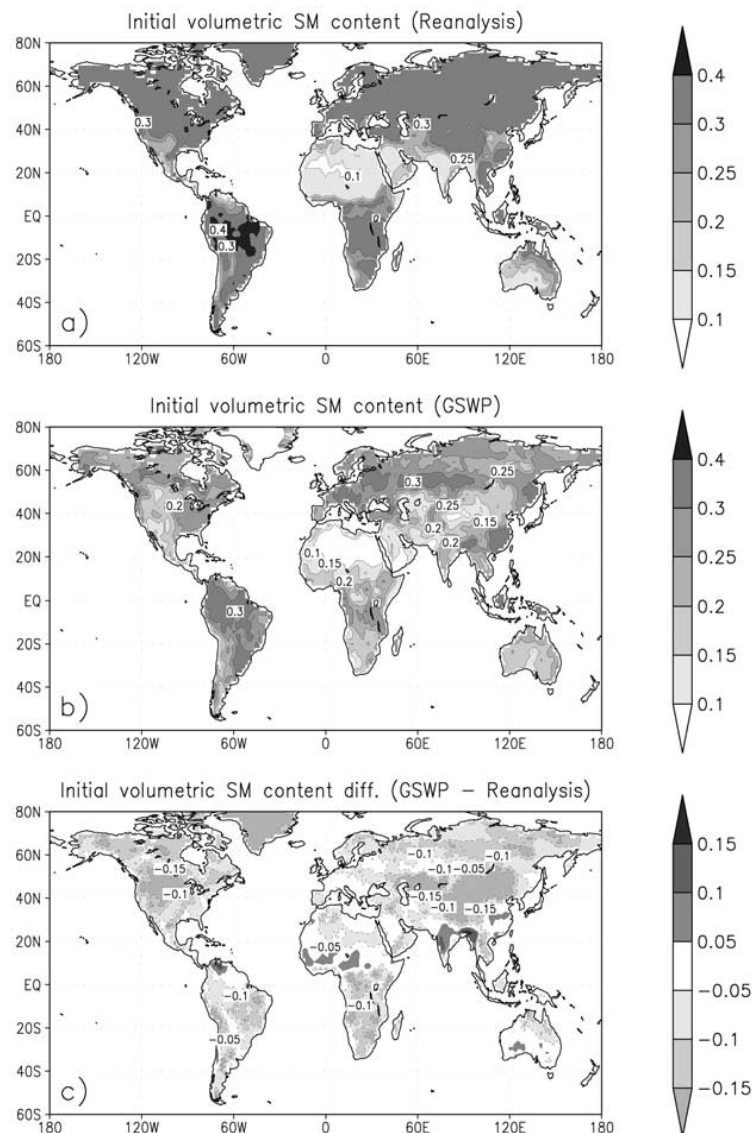


Figure 16. Initial volumetric soil water content for (a) Reanalysis; (b) GSWP; (c) GSWP–Reanalysis.

evidence that vegetation processes contribute to the northward jump of the monsoon. In fact, with the accurate initial soil moisture field, the simulated abrupt northward jump was closer to the observations (Figure 5d, dashed arrow), which suggests the soil moisture might influence the speed of the jump. A more detailed analysis revealed that most differences of precipitation between case S2 and S1 were caused by the changes in convective precipitation.

6. Impact of the Land Cover Classification Map

[43] The impact of the specification of vegetation is illustrated by a comparison of case S3, using OLD SSiB MAP, and case S1, using NEW SSiB MAP. These produced different precipitation in tropical and subtropical regions (Figures 18a and 19a). Evaporation was the main

cause of variations in precipitation on the Eurasian continent north of 40°N (Figure 18c). In the Inner Mongolian grassland and northeastern east Asia, the desert in case S3 produced less evaporation and precipitation. Divergence in northeastern east Asia was another factor that might have contributed to the reduction in precipitation (Figure 18b). On the other hand, in most parts of east Asia to the south of 40°N, the precipitation changes were associated with the dominant moisture flux convergence, consistent with the discussions in previous sections (Figure 8b). The desert in central Asia in case S3 produced a divergence region along 20°N and 30°N and a convergence region to the south. This was in general agreement with Xue [1996]. In east Asia the northward jump of the east Asian monsoon in case S3 was still evident (Figure 5e), but the timing was delayed by about

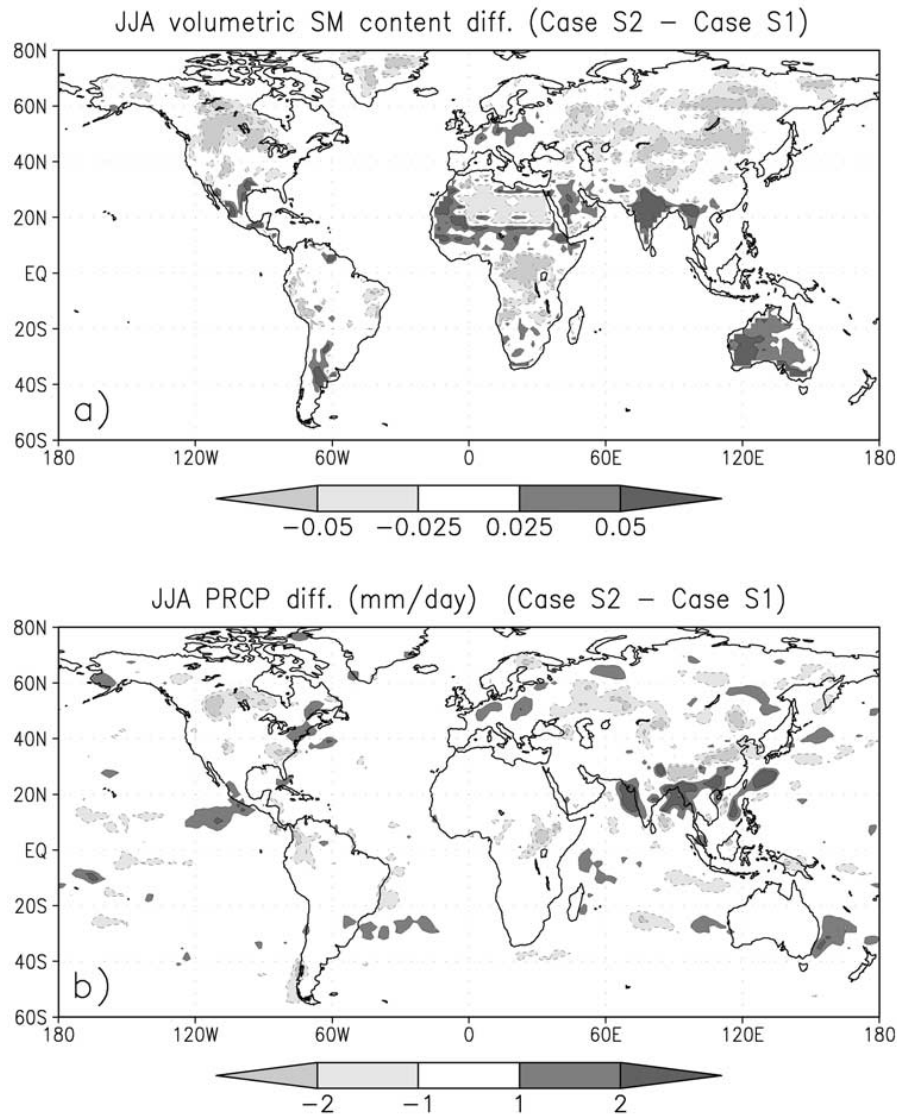


Figure 17. JJA differences between case S2 and case S1; (a) volumetric soil water content; (b) precipitation (mm d^{-1}).

one month, indicating that land cover change can modify the timing of the onset of the east Asian monsoon.

[44] In the Indian subcontinent there were no dramatic vegetation cover differences between NEW and OLD SSiB MAP. The precipitation differences between India and its surrounding ocean were mainly associated with changes in moisture fluxes, which was probably a response to land cover change in Eurasia to the north (Figures 18a and 18b). Although many studies have investigated the relationship between Eurasian snow cover and Indian monsoon, there have not been any studies investigating the relationship between Eurasian vegetation and Indian monsoon.

[45] The Sahara desert, where the vegetation classifications differed, had insufficient precipitation to exhibit any effect (Figure 19a) but there was an increase in induced

divergence over a region between 10°N and 20°N and a convergence region to the south along the coastal area (Figure 19b): a dipole type of change noted by others [e.g., Xue, 1997]. The precipitation changes in case S3 in the African continent were coincident with these divergence and convergence regions. Since there was no substantial land cover change in central Africa, the reduction in evaporation (Figure 19c) was probably a response to the reduction in precipitation.

[46] NEW SSiB MAP improved the simulation of precipitation substantially in some important monsoon regions. For example, over northern Africa (10° to 40°E , and 0°N to 10°N), the JJA precipitations were 4.65 , $4.41(\pm 0.40)$, and $3.05(\pm 0.35)$ mm d^{-1} for observation, case S1 and case S3, respectively. Even in southern China (between 110°E and 120°E and 25°N and 30°N), the error in simulation for JJA

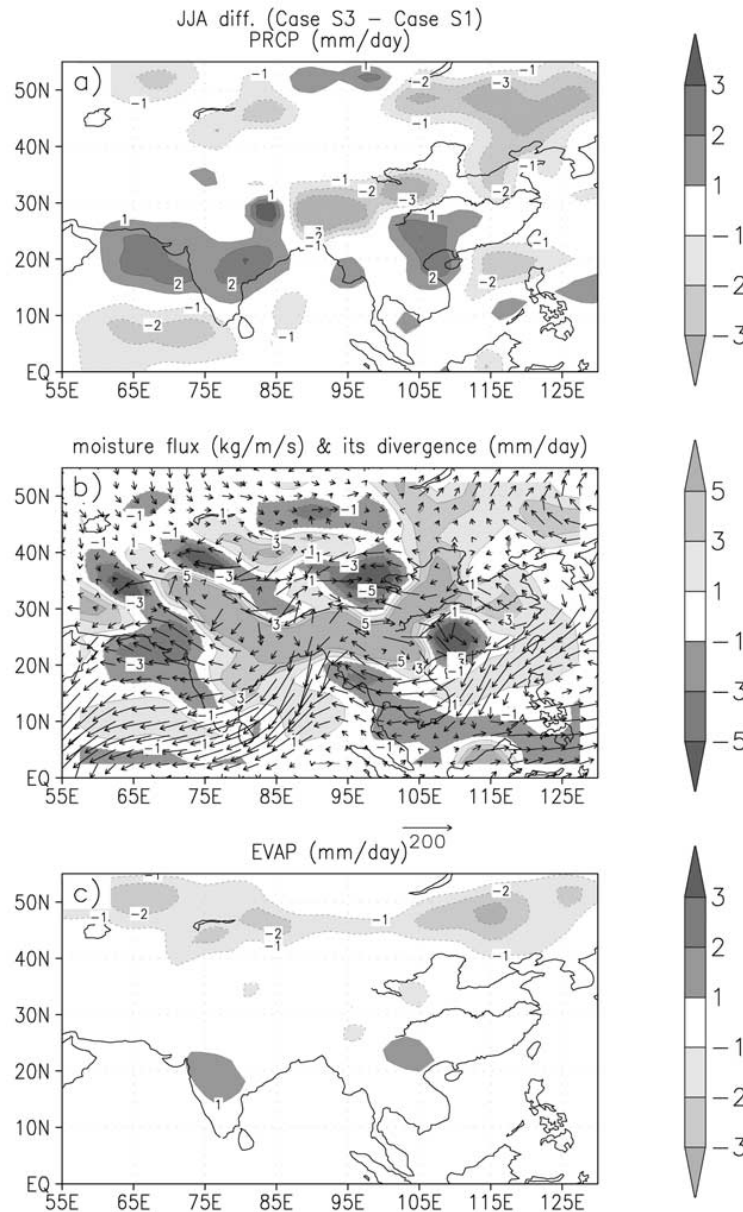


Figure 18. June differences between case S3 and case S1 in Asia (a) precipitation (mm d^{-1}); (b) vertically integrated moisture flux ($\text{kg m}^{-1} \text{s}^{-1}$) and its divergence (mm d^{-1}); (c) evaporation (mm d^{-1}).

precipitation was reduced from 1.17 mm d^{-1} in case S3 to 0.39 mm d^{-1} in case S1.

7. Discussion and Summary

[47] This study explores the impact of land surface processes on the structure and characteristics of the monsoon system with an emphasis on the evolution of precipitation. The results were obtained using the NCEP GCM coupled with two different land surface parameterizations that included or did not include vegetation processes. Because the study focused on intraseasonal variability with a temporal scale, in some cases, of only 5–10 days, three

scenarios, differing in initial soil moistures and vegetation maps, with three initial conditions were used for each land surface parameterization to evaluate the robustness of the model results.

[48] In addition to the results for east Asia and West Africa, discussed above, we also examined the global mean precipitation and the precipitation over the land (Table 3). The standard deviations were substantially smaller than the differences between case C and case Ss and were of the same order of magnitude for each month. Case C simulated the global climate with reasonable accuracy. Cases S1 and S3 provided small, but consistent improvements for each month in the simulations, which indicated that the improve-

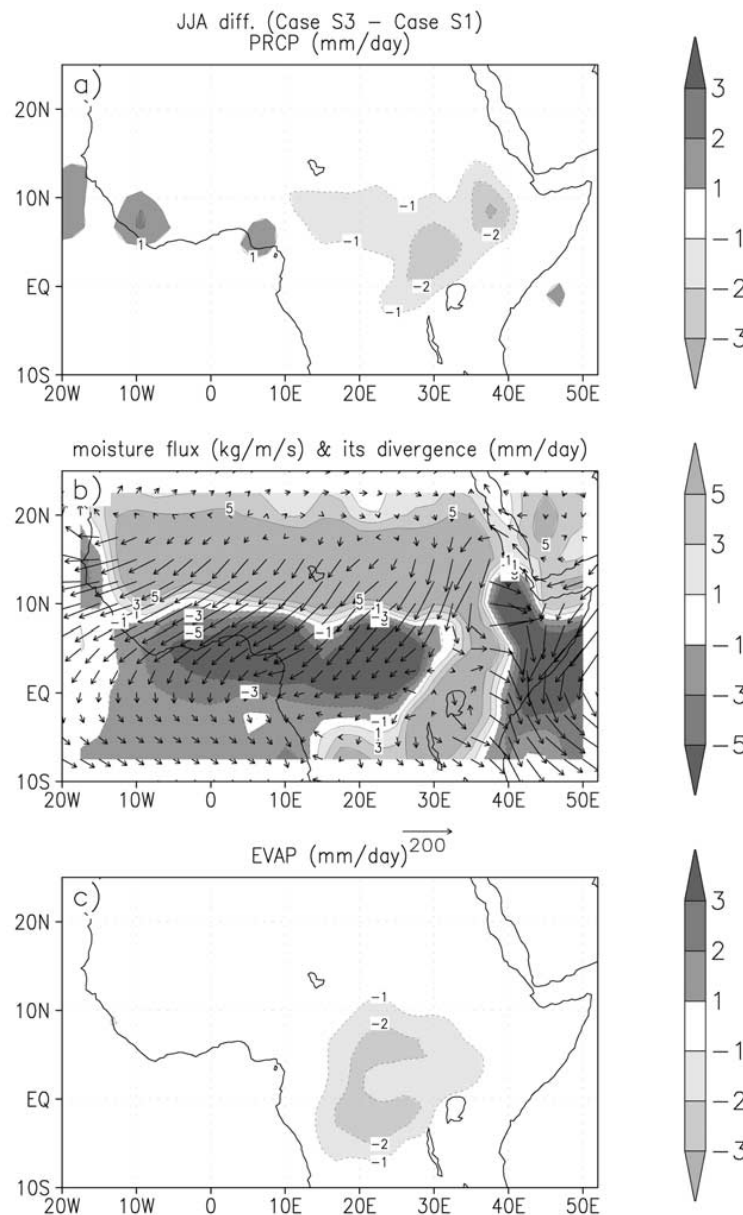


Figure 19. June differences between case S3 and case S1 in Africa (a) precipitation (mm d^{-1}); (b) vertically integrated moisture flux ($\text{kg m}^{-1} \text{s}^{-1}$) and its divergence (mm d^{-1}); (c) evaporation (mm d^{-1}).

ments at a regional scale were not at the cost of global realism. It was interesting to note that case S3 had the best simulation of the global mean (Table 3). This was mainly because, after classifying most of east Asia as desert, the wet bias in the model simulation was substantially reduced. For the area between 105°E and 125°E and 30°N and 50°N , the observed, case S1 and case S3 JJA precipitation was 3.73 , $5.89(\pm 0.35)$, and $4.08(\pm 0.33)$ mm d^{-1} , respectively. Therefore the underlying cause of the “better” global mean precipitation in case S3 was most likely due to weakness in the GCM and/or land surface model or problems in vegetation parameter specification. There are no direct

measurements of the vegetation and soil parameters for most parts of the world.

[49] Although the simulations by using the two parameterizations were compared with each other and with observations, the aim of this study was principally to understand better the influence of biophysical processes on the processes of monsoon development. The results show that at the planetary scale, two different land surface parameterizations produced similar monthly mean simulations of precipitation (Figure 3). The differences in global means were small (Table 3). However, at the continental and synoptic scales, more complete represen-

Table 3. Mean Precipitation Over Land and Over the Globe^a

	May	June	July	August	JJA	s.d. JJA
<i>Xie and Arkin</i> [1997]	1.84 (2.74)	2.10 (2.75)	2.16 (2.81)	2.23 (2.77)	2.04 (2.77)	
Case C	2.41 (2.98)	2.36 (3.07)	2.62 (3.04)	2.58 (2.99)	2.52 (3.03)	0.04 (0.01)
Case S1	2.13 (2.89)	2.30 (3.02)	2.56 (3.00)	2.45 (2.90)	2.43 (2.97)	0.04 (0.01)
Case S2	1.87 (2.88)	2.16 (2.95)	2.46 (2.97)	2.34 (2.89)	2.32 (2.94)	0.02 (0.01)
Case S3	1.73 (2.82)	2.01 (2.96)	2.27 (2.90)	2.31 (2.85)	2.20 (2.90)	0.06 (0.005)

^aMean precipitation over the globe is given in parenthesis. Values are in mm d⁻¹. Here, s.d., standard deviation.

tation of land surface processes improved the simulation of the structure and characteristics of the monsoon systems.

[50] Cases S1 and C used the similar monthly mean albedo as well as surface roughness length, and the same initial soil moisture. However, the two land surface schemes produced different surface water and energy balances, different partitioning of latent heat and sensible heat fluxes (the Bowen ratio), and different latitudinal and longitudinal thermal gradients at the surface. The effects were mainly manifested in the temporal evolution of the monsoon, its strength, the spatial distribution of precipitation, and associated circulation at continental and synoptic scales. Under the three scenarios, the GCM with a biosphere model consistently simulated the abrupt northward movement of the east Asian monsoon unlike the GCM/SOIL model runs.

[51] Furthermore, this study shows the surface processes influenced the turning of the low-level wind counterclockwise or clockwise during the premonsoon or early monsoon stages, the low-level land/sea temperature gradient, wind flow, and moisture transport, which were related to the monsoon development. This study shows that under unstable conditions, not only the low-frequency mean forcings from the land surface, but also the perturbation processes of vegetation forcing described in SSiB on much shorter timescales, such as radiative flux/canopy interaction and transpiration, may be crucial in the evolution of the monsoon. Since 1987 alone was simulated for this study, further investigations under different scenarios (such as different SSTs) will be necessary to confirm this finding.

[52] In previous studies, we found that specifications of land degradation in Sahel and east Asia allowed climate simulations to reproduce decadal anomaly patterns of precipitation and surface temperature [Xue and Shukla, 1993; Xue, 1996, 1997]. Furthermore, better representations of land surface processes in a regional model improved the short-term (24 and 48 hours) simulations of extreme climate events, such as the 1993 U.S. flood [Xue et al., 2001]. The results from this study show that land surface processes may also be important for intraseasonal simulations. However, the land/atmosphere interactions are complex and nonlinear as shown in Figure 3 and the dominant mechanisms depend on temporal and spatial scales and background climate conditions.

[53] Our findings show that better specification of the initial soil moisture improved seasonal simulation, mainly in the intensity of the simulated variables. It also suggests possible relationships between vegetation distribution in the Eurasian continent and the Indian monsoon intensity, as well as land degradation in east Asia and timing of the east

Asian monsoon onset, indicating the importance of accurate land cover maps.

[54] Both the NCEP soil submodel and SSiB are physically based models, which was evident by the fact that no empirical tuning was needed in this study when the soil model was replaced with SSiB, and the differences between the simulations by the two models could be clearly related to physical and dynamic processes, even in the complexity of a GCM. These conclusions, however, need to be evaluated using different models. Although this version of the NCEP GCM produced substantially better east Asian simulations than, for example, by Xue [1996], substantial biases were still evident (e.g., Figure 3). The accurate simulation and prediction of monsoons, especially the Asian monsoon, is a formidable task and some crucial improvements remain to be made, such as the simulation of precipitation in south Asia. This study showed that high-quality observational data of land cover and assimilated soil moisture help identify the role of land surface parameterizations in monsoon simulation and, more generally, in land/atmosphere interactions.

[55] **Acknowledgments.** The authors thank Masao Kanamitsu, UCSD, for providing the NCEP GCM code and for his support; Song-You Hong and Hua-Lu Pan, NCEP, for help in this work; Eugenia Kalnay, University of Maryland, for her support; C. Roberto Mechoso of UCLA Xiuji Zhou of the Chinese Academy of Meteorological Sciences (CAMS), and Serge Janicot of the Laboratoire de Meteorologie Dynamique, France, for discussions; Pingping Xie of NCEP for meteorological data; and Paul Dirmeyer of COLA for GSWP data. Michio Yanai of UCLA, Guoxiong Wu of the Institute of Atmospheric Physics, Chinese Academy of Sciences, Longxun Chen and Weiliang Li of CAMS, Akiyo Yatagai of the Research Institute of Humanity and Nature, Japan, and K.-M. Lau of NASA have provided comments for this manuscript. Funding was provided by NSF grants EAR 9706403 and ATM-0097260 and NASA grants NAG5-9014 and NAG5-9329. The model runs were carried out on the NCAR and NCEP supercomputers.

References

- Barnett, T. P., L. Dumenil, U. Schlese, E. Roeckner, and M. Latif (1989), The effect of Eurasian snow cover on regional and global climate variations, *J. Atmos. Sci.*, **46**, 661–685.
- Businger, J. A., J. C. Wyngaard, Y. Izumi, and E. G. Bradley (1971), Flux-profile relationships in the atmospheric surface layer, *J. Atmos. Sci.*, **28**, 181–189.
- Charney, J. G., W. K. Quirk, S.-H. Chow, and J. Kornfield (1977), A comparative study of the effects of albedo change on drought in semi-arid regions, *J. Atmos. Sci.*, **34**, 1366–1385.
- Chen, L., Q. Shu, and H. Lu (1991), *East Asian Monsoon* (in Chinese), 362 pp., China Meteorol. Press, Beijing.
- Chen, L., W. Li, and P. Zhao (2001), On the process of summer monsoon onset over east Asia, *Acta Meteorol. Sin.*, **15**, 436–449.
- Clark, D. B., Y. Xue, R. Harding, and P. J. Valdes (2001), Modeling the impact of land surface degradation on the climate of tropical North Africa, *J. Clim.*, **14**, 1809–1822.
- Deardorff, J. W. (1977), Efficient prediction of a ground surface temperature and moisture with inclusion of a layer of vegetation, *J. Geophys. Res.*, **83**, 1889–1903.

- Dey, B., and O. S. R. U. Bhanu Kumar (1982), An apparent relationship between Eurasian snow cover and the advanced period of the Indian summer monsoon, *J. Appl. Meteorol.*, 21, 1929–1932.
- Ding, Y. (1994), *Monsoon in China*, 419 pp., Kluwer Acad., Norwell, Mass.
- Dirmeyer, P. A. (2000), Using a global soil wetness dataset to improve seasonal climate simulation, *J. Clim.*, 13, 2900–2922.
- Dirmeyer, P. A., A. J. Dolman, and N. Sato (1999), The pilot phase of the Global Soil Wetness Project, *Bull. Am. Meteorol. Soc.*, 80, 851–878.
- Dorman, J. L., and P. Sellers (1989), A global climatology of albedo, roughness length and stomatal resistance for atmospheric general circulation models as represented by the Simple Biosphere Model (SiB), *J. Appl. Meteorol.*, 28, 833–855.
- Douville, H., and J.-F. Royer (1996), Sensitivity of the Asian summer monsoon to an anomalous Eurasian snow cover within the Météo-France GCM, *Clim. Dyn.*, 12, 449–466.
- Douville, H., F. Chauvin, and H. Broqua (2001), Influence of soil moisture on the Asian and African monsoons. Part I: Mean monsoon and daily precipitation, *J. Clim.*, 14, 2381–2403.
- Flohn, H. (1957), Large-scale aspects of the “summer monsoon” in South and east Asia, *J. Meteorol. Soc. Jpn.*, 75, 180–186.
- Folland, C. K., J. Owen, M. N. Ward, and A. Colman (1991), Prediction of seasonal rainfall in the Sahel region using empirical and dynamical methods, *J. Forecast.*, 1, 21–56.
- Grell, G. A., J. Dudhia, and D. R. Stauffer (1994), A description of the fifth generation Penn State/NCAR Mesoscale Model (MM5), *NCAR Tech. Note NCAR/TN-398+STR*, 121 pp., Natl. Cent. for Atmos. Res., Boulder, Colo.
- Griffiths, J. F. (1972), General climatology, in *Climates of Africa*, edited by J. F. Griffiths, pp. 1–35, Elsevier Sci., New York.
- Hahn, D. J., and J. Shukla (1976), An apparent relationship between Eurasian snow cover and Indian monsoon rainfall, *J. Atmos. Sci.*, 33, 2461–2462.
- Hansen, M. C., R. S. DeFries, J. R. G. Townshend, and R. Sohlberg (2000), Global land cover classification at 1 km spatial resolution using a classification tree approach, *Int. J. Remote Sens.*, 21, 1303–1330.
- Hastenrath, S. (1984), International variability and annual cycle: Mechanisms of circulation and climate in the tropical Atlantic sector, *Mon. Weather Rev.*, 112, 1097–1107.
- Hoskins, B. J. (1987), Diagnosis of forced and free variability in the atmosphere, in *Atmospheric and Oceanic Variability*, edited by H. Cattle, pp. 57–73, James Glaisher House, Bracknell, UK.
- Kalnay, E., M. Kanamitsu, and W. E. Baker (1990), Global numerical weather prediction at the National Meteorological Center, *Bull. Am. Meteorol. Soc.*, 71, 1410–1428.
- Kalnay, E., et al. (1996), The NMC/NCAR 40-year Reanalysis Project, *Bull. Am. Meteorol. Soc.*, 77, 437–471.
- Kanamitsu, M., J. C. Alpert, K. A. Campana, P. M. Caplan, D. G. Deaven, M. Iredell, B. Katz, H.-L. Pan, J. Sela, and G. H. White (1991), Recent changes implemented into the global forecast system at NMC, *Weather Forecast.*, 6, 425–435.
- Kistler, E., et al. (1999), The NCEP/NCAR 50-year Reanalysis, *Bull. Am. Meteorol. Soc.*, 82, 247–268.
- Kitoh, A., K. Yamazaki, and T. Tokioka (1988), Influence of soil moisture and surface albedo changes over the African tropical rain forest on summer climate investigated with the MRI-GCM-I, *J. Meteorol. Soc. Jpn.*, 66, 65–85.
- Krishnamurti, T. N., H. S. Bedi, and M. Subramaniam (1989), The summer monsoon of 1987, *J. Clim.*, 2, 321–340.
- Kuchler, A. W. (1983), World map of natural vegetation, in *Goode's World Atlas*, 16th ed., pp. 16–17, Rand McNally, New York.
- Lamb, P. J. (1978), Large-scale tropical Atlantic surface circulation patterns associated with sub-Saharan weather anomalies, *Tellus*, 30, 240–251.
- Lamb, P. J., and R. A. Pepler (1991), West Africa, in *Teleconnections Linking Worldwide Climate Anomalies*, edited by M. Glantz, R. W. Katz, and N. Nicholls, pp. 121–189, Cambridge Univ. Press, New York.
- Lau, K.-M., and W. Bua (1998), Mechanisms of monsoon-Southern Oscillation coupling: Insights from GCM experiments, *Clim. Dyn.*, 14, 759–779.
- Lau, K.-M., G. J. Yang, and S. H. Shen (1988), Seasonal and intraseasonal climatology of summer monsoon rainfall over east Asia, *Mon. Weather Rev.*, 116, 18–37.
- Laval, K., and L. Picon (1986), Effect of a change of the surface albedo of the Sahel on climate, *J. Atmos. Sci.*, 43, 2418–2429.
- Liu, C., and S. Tao (1983), Northward jumping of subtropical highs and CUSP catastrophe, *Sci. Sin., Ser. B, Engl. Ed.*, 26, 1065–1074.
- Liu, X., and M. Yanai (2002), Influence of Eurasian spring snow cover on Asian summer rainfall, *Int. J. Climatol.*, 22(9), 1075–1089.
- Louis, J.-F. (1979), A parametric model of vertical eddy fluxes in the atmosphere, *Boundary Layer Meteorol.*, 17, 187–202.
- Matthews, E. (1984), Prescription of land-surface boundary conditions in GISS GCM II: A simple method based on high-resolution vegetation data bases, *NASA Tech. Memo., NASA TM-86096*, 20 pp.
- Matthews, E. (1985), Atlas of archived vegetation, land-use and seasonal albedo data sets, *NASA Tech. Memo., NASA TM-86199*, 53 pp.
- Meehl, G. A. (1994), Influence of the land surface in the Asian summer monsoon: External conditions versus internal feedbacks, *J. Clim.*, 7, 1033–1049.
- Miao, J., and M. Ding (1985), Catastrophe theory of seasonal variation, *Sci. Sin., Ser. B, Engl. Ed.*, 28, 1079–1092.
- Miyakoda, K., and J. Sirutis (1986), Manual of the E-physics, Geophys. Fluid Dyn. Lab., Princeton Univ., Princeton, N. J.
- Nicholson, S. E. (1976), A climatic chronology for Africa: Synthesis of geological, historical, and meteorological information and data, Ph.D. thesis, Univ. of Wisc., Madison.
- Niyogi, D. S., Y. Xue, and S. Raman (2002), Hydrological feedback in land-atmosphere coupling: Comparison of a tropical and a midlatitudinal regime, *J. Hydrometeorol.*, 3, 39–56.
- Palmer, T. N., C. Brankovic, P. Viterbo, and M. J. Miller (1992), Modeling interannual variations of summer monsoons, *J. Clim.*, 5, 399–417.
- Pan, H.-L., and L. Mahrt (1987), Interaction between soil hydrology and boundary layer developments, *Boundary Layer Meteorol.*, 38, 185–202.
- Paulson, C. A. (1970), Mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer, *J. Appl. Meteorol.*, 9, 857–861.
- Robock, A., M. Mu, K. Vinnikov, and D. Robinson (2003), Land surface conditions over Eurasia and Indian summer monsoon rainfall, *J. Geophys. Res.*, 108(D4), 4131, doi:10.1029/2002JD002286.
- Rowell, D. P., C. K. Folland, K. Maskell, and M. N. Ward (1995), Variability of summer rainfall over tropical North Africa (1906–92): Observation and modeling, *Q. J. R. Meteorol. Soc.*, 121, 669–704.
- Shi, L., and E. A. Smith (1992), Surface forcing of the infrared cooling profile over the Tibetan Plateau. Part II: Cooling-rate variation over large-scale Plateau domain during summer monsoon transition, *J. Atmos. Sci.*, 49, 823–844.
- Sud, Y. C., and M. Fennesy (1982), A study of the influence of surface albedo on July circulation in semi-arid regions using the GLAS GCM, *J. Clim.*, 2, 105–125.
- Sultan, B., and S. Janicot (2000), Abrupt shift of the ITCZ over West Africa and intra-seasonal variability, *Geophys. Res. Lett.*, 27, 3353–3356.
- Vernekar, A. D., J. Zhou, and J. Shukla (1995), The effect of Eurasian snow cover on the Indian monsoon, *J. Clim.*, 8, 248–266.
- Walker, J., and P. R. Rowntree (1977), The effect of soil moisture on circulation and rainfall in a tropical model, *Q. J. R. Meteorol. Soc.*, 103, 29–46.
- Wang, G. (1986), Some factors determining the jump of the sub-tropical highs (in Chinese), *Sci. Meteorol. Sin.*, 1, 50–60.
- Wang, G., and E. A. B. Eltahir (1999), Biosphere-atmosphere interactions over West Africa. I: Development and validation of a coupled dynamic model, *Q. J. R. Meteorol. Soc.*, 126, 1239–1260.
- Webster, P. J. (1987), The elementary monsoons, in *Monsoons*, edited by J. F. Fein and P. L. Stephens, pp. 3–32, John Wiley, New York.
- Webster, P. J., V. Magana, T. N. Palmer, J. Shukla, R. A. Tomas, M. Yanai, and T. Yasunari (1998), Monsoons: Processes, predictability, and the prospects for prediction, *J. Geophys. Res.*, 103(C7), 14,451–14,510.
- Weng, H. Y., K. M. Lau, and Y. K. Xue (1999), Long term variations of summer rainfall over China and its possible link to global sea-surface temperature variability, *J. Meteorol. Soc. Jpn.*, 4, 845–857.
- Willmott, C. J., and K. Klink (1986), A representation of the terrestrial biosphere for use in global climate studies, in *ISLSCP: Proceedings of an International Conference Held in Rome, Italy, 2–6 December 1985*, pp. 109–112, Eur. Space Agency, Paris.
- Wu, G., and Y. M. Liu (2003), Summertime quadruplet heating pattern in the subtropics and the associated atmospheric circulation, *Geophys. Res. Lett.*, 30, 1201–1204.
- Wu, G., L. Sun, Y. Liu, H. Liu, S. Sun, and W. Li (2002), Impacts of land surface processes on summer climate, in *Selected Papers of the Fourth Conference on East Asia and Western Pacific Meteorology and Climate*, edited by C. P. Chang et al., pp. 64–76, World Sci., River Edge, N. J.
- Xie, P., and P. A. Arkin (1997), Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates and numerical model outputs, *Bull. Am. Meteorol. Soc.*, 78, 2539–2558.
- Xue, Y. (1996), The impact of desertification in the Mongolian and the Inner Mongolian Grassland on the regional climate, *J. Clim.*, 9, 2173–2189.
- Xue, Y. (1997), Biosphere feedback on regional climate in tropical North Africa, *Q. J. R. Meteorol. Soc.*, 123, 1483–1515.
- Xue, Y., and J. Shukla (1993), The influence of land surface properties on Sahel climate. Part I: Desertification, *J. Clim.*, 6, 2232–2245.

- Xue, Y., P. J. Sellers, J. L. Kinter III, and J. Shukla (1991), A simplified biosphere model for global climate studies, *J. Clim.*, **4**, 345–364.
- Xue, Y., H. G. Bastable, P. A. Dirmeyer, and P. J. Sellers (1996a), Sensitivity of simulated surface fluxes to changes in land surface parameterization: A study using ABRACOS data, *J. Appl. Meteorol.*, **35**, 386–400.
- Xue, Y., M. J. Fennessy, and P. J. Sellers (1996b), Impact of vegetation properties on U.S. summer weather prediction, *J. Geophys. Res.*, **101**(D3), 7419–7430.
- Xue, Y., J. Elbers, F. J. Zeng, and A. J. Dolman (1997), GCM parameterization for Sahelian land surface processes, in *HAPEX-Sahel West Central Supersite: Methods, Measurements and Selected Results, Rep. 130.HM/07.97*, edited by P. Kabat, S. Prince, and L. Prihodko, pp. 289–297, Winand Staring Cent. for Integrated Land, Soil and Water Res., Wageningen, Netherlands.
- Xue, Y., F. J. Zeng, K. Mitchell, Z. Janjic, and E. Rogers (2001), The impact of land surface processes on the simulation of the U.S. hydrological cycle: A case study of 1993 US flood using the Eta/SSiB regional model, *Mon. Weather Rev.*, **129**, 2833–2860.
- Xue, Y., R. W. A. Hutjes, R. J. Harding, M. Claussen, S. Prince, E. F. Lambin, S. J. Alan, and P. Dirmeyer (2003), The Sahelian climate, in *Vegetation, Water, Humans, and the Climate*, edited by P. Kabat et al., chap. 5, pp. 57–76, Springer-Verlag, New York.
- Yanai, M., and T. Tomita (1998), Seasonal and interannual variability of atmospheric heat sources and moisture sinks as determined from NCEP/NCAR reanalysis, *J. Clim.*, **11**, 463–482.
- Yang, S., and K.-M. Lau (1998), Influences of sea surface temperature and ground wetness on Asian summer monsoon, *J. Clim.*, **11**, 3230–3246.
- Yasunari, T., A. Kitoh, and T. Tokioka (1991), Local and remote responses to excessive snow mass over Eurasia appearing in the northern spring and summer climate: A study with the MRI GCM, *J. Meteorol. Soc. Jpn.*, **62**, 473–487.
- Yatagai, A., and T. Yasunari (1994), Trends and decadal-scale fluctuations of surface air temperature and precipitation over China and Mongolia during the recent 40 year period (1951–1990), *J. Meteorol. Soc. Jpn.*, **72**, 937–957.
- Yatagai, A., and T. Yasunari (1995), Interannual variations of summer precipitation in the arid/semi-arid regions in China and Mongolia: Their regionality and relation to the Asian summer monsoon, *J. Meteorol. Soc. Jpn.*, **73**, 909–923.
- R. DeFries and S. Prince, Department of Geography, University of Maryland, College Park, Le Frak Hall, College Park, MD 20742, USA. (rd63@umail.umd.edu; sp43@umail.umd.edu)
- Y. Jiao, Department of Earth and Atmospheric Sciences, University of Quebec at Montreal, 550 Sherbrooke West Road, Montreal, Canada H3A 1B9. (jiao@atlas.sca.uqam.ca)
- H.-M. H. Juang, NCEP/NWS/NOAA, WWBG Room 204, 5200 Auth Road, Camp Springs, MD 20746, USA. (wd20hh@ncep.noaa.gov)
- W.-P. Li, R. Vasic, and Y. Xue, Department of Geography, University of California, Los Angeles, 1255 Bunche Hall, UCLA Box 951524, Los Angeles, CA 90095-1524, USA. (wli@weber.sscnet.ucla.edu; rvasic@weber.sscnet.ucla.edu; yxue@geog.ucla.edu)

Other Key Reference Papers for Prof. Xue's Lecture:

- [1] Xue, Y., F. J. Zeng, K. Mitchell, and Z. Janjic, E. Rogers, 2001: The Impact of Land Surface Processes on Simulations of the U.S. Hydrological Cycle: A Case Study of the 1993 Flood Using the SSiB Land Surface Model in the NCEP Eta Regional Model. *Mon. Wea. Rev.*, **129**, 2833-2860.
- [2] Xue, Y., 2005: Land surface processes and monsoon. *GEWEX Newsletter*. February, 5, 6, 20.
- [3] Xue, Y., S. Sun, J.-M. Lau, J. Ji, I. Poccarrd, H.-S. Kang, R. Zhang, G. Wu, J. Zhang, J. Schaake, and Y. Jiao, 2005, Multiscale variability of the river runoff system in China and its link to precipitation and sea surface temperature. *J. Hydrometeorology*, **6**, 550-570.
- [4] Xue, Y., F. De Sales, R. Vasic, C. R. Mechooso, S. D. Prince, A. Arakawa, 2010: Global and Seasonal Assessment of Interactions between Climate and Vegetation Biophysical Processes: A GCM Study with Different Land-Vegetation Representations. *J. Climate*, **23**, 1411 -1433.

Improving the Noah Land Surface Model in Arid Regions with an Appropriate Parameterization of the Thermal Roughness Length

YINGYING CHEN, KUN YANG, DEGANG ZHOU, JUN QIN, AND XIAOFENG GUO

*Key Laboratory of Tibetan Environment Changes and Land Surface Processes, Institute of Tibetan Plateau Research,
Chinese Academy of Sciences, Beijing, China*

(Manuscript received 18 June 2009, in final form 28 March 2010)

ABSTRACT

Daytime land surface temperatures in arid and semiarid regions are typically not well simulated in current land surface models (LSMs). This study first evaluates the importance of parameterizing the thermal roughness length (z_{0h}) to model the surface temperature (T_{sfc}) and turbulent sensible heat flux (H) in arid regions. Six schemes for z_{0h} are implemented into the Noah LSM, revealing the high sensitivity of the simulations to its parameterization. Comparisons are then performed between the original Noah LSM and a revised version with a novel z_{0h} scheme against observations at four arid or semiarid sites, including one in Arizona and three in western China. The land they cover is sparse grass or bare soil. The results indicate that the original Noah LSM significantly underestimates T_{sfc} and overestimates H in the daytime, whereas the revised model can simulate well both T_{sfc} and H simultaneously. The improved version benefits from the successful modeling of the diurnal variation of z_{0h} , which the original model cannot produce.

1. Introduction

Arid and semiarid regions are an important portion of the global land surface. Many studies have indicated that desertification is increasing as a result of climatic change and human activities (Puigdefabregas 1995; Warren 1996). The arid and semiarid region of northwestern China has experienced significant environmental changes within the last half century (Ma and Fu 2006). Thus, it is crucial to understand the land–atmosphere interactions and to predict the variations of the hydrometeorological regimes in these regions. From a lack of precipitation, the surface heat transport becomes the dominant land surface process (see section 3b) in arid regions. Therefore, parameterizing the surface heat transport process in land surface models (LSMs) is vital for accurately modeling the surface energy budget.

However, the ability of current LSMs to simulate land processes in arid and semiarid regions still needs improvement. Hogue et al. (2005) found that the Noah LSM tended to overestimate the sensible heat flux (H) and

underestimate the surface temperature (T_{sfc}) during the dry season. Yang et al. (2007) evaluated seven general circulation models against coordinated enhanced observing period (CEOP) observations and found that all of the models significantly underestimated the daytime ground–air temperature differences, particularly severely in arid and semiarid regions. LeMone et al. (2008) showed that the default Noah model tended to overestimate H and underestimate T_{sfc} in relatively dry conditions. Yang et al. (2009) further evaluated three offline LSMs against observations in the area of the Tibetan Plateau (TP); their study indicated that all models significantly underestimated the daytime T_{sfc} . Chen and Zhang (2009) also found that the Noah model often overestimated the surface exchange coefficient for heat (C_h) over short vegetation.

The aforementioned modeling biases in T_{sfc} and H imply that the heat transfer resistance in the models is not appropriately parameterized. This resistance is related to both the aerodynamic roughness length (z_{0m}) and the thermal roughness length (z_{0h}). Much of the literature has focused on the parameterization of z_{0h} (Sheppard 1958; Brutsaert 1982; Zilitinkevich 1995; Zeng and Dickinson 1998; Kanda et al. 2007; Smeets and van den Broeke et al. 2008; Yang et al. 2008). Ma et al. (2002) and Yang et al. (2003) found that z_{0h} exhibits an

Corresponding author address: Yingying Chen, Institute of Tibetan Plateau Research, CAS, No. 18 Shuangqing Rd., Beijing, China 100085.

E-mail: chenyy@itpcas.ac.cn

DOI: 10.1175/2010JHM1185.1

© 2010 American Meteorological Society

jhm1185

TABLE 1. General information about the four sites used in this study.

Site	Data source	Lat, lon	Elevation (m)	Land cover
Audubon	AmeriFlux	31°35'N, 110°31'W	1469	Desert grassland
Shiquanhe	GAME-Tibet	32°30'N, 80°05'E	4279	Alpine desert
Gaize	GAME-Tibet	32°18'N, 84°03'E	4420	Alpine desert
Dunhuang	NWC-ALIEX	40°09'N, 90°41'E	1150	Desert

evident diurnal variation on the TP. Yang et al. (2008) evaluated several schemes and indicated that z_{0h} depends on the flow state and exhibits diurnal variations. Some of these z_{0h} schemes were developed and evaluated against observations in the micrometeorology community, while they need more critical evaluations and practical tests in LSMs to verify their effectiveness. That is the motivation for this study.

The objective of this work is twofold: 1) to assess the sensitivity of the land surface energy budget to different parameterizations of z_{0h} in the Noah LSM and 2) to evaluate the performance of a revised Noah LSM against field observations in arid regions. The Noah LSM is selected because it is widely used and has been adopted for operations and research in National Centers for Environmental Prediction (NCEP) weather and climate predictions models and relevant data assimilation systems.

In this paper, section 2 briefly introduces the climatic characteristics of the sites and their relevant measurements. Section 3 describes the LSM used in this work and the settings of the model parameters. Section 4 tests the sensitivity to key LSM parameters in arid regions, and section 5 evaluates the performance of the revised LSM. Concluding remarks are given in section 6.

2. Sites and observation data

In this study, simulations are conducted at one semi-arid site in Arizona and three arid sites in western China: two on the TP and one in the northwest. Some general information about these sites is given in Table 1.

T1

The semiarid site is Audubon Research Ranch (referred to as Audubon hereafter), which is located in Arizona. The mean annual precipitation at Audubon is ~300–400 mm. The land surface is characterized by sparse brown grass during the simulation period. The data were collected through the AmeriFlux network (information online at <http://public.ornl.gov/ameriflux/>). The required forcing data were measured by automatic weather stations. The observed ground truth data included the surface temperature, the soil temperature profile (2, 4, 8, 16, 32, 64, and 128 cm), the soil moisture profile (10, 20, 30, 40, 60, and 100 cm), and turbulent fluxes. The surface temperature was given by an infrared

thermometer [Apogee Infrared Thermocouple Sensor (IRTS-P), Campbell Scientific]; the soil moisture sensor (PR1/6, Delta-T Devices) and soil temperature probe with thermistors (YSI Inc.) were set up to measure the soil moisture and soil temperature, respectively; and the turbulent fluxes was measured by an eddy-covariance system (LI-COR LI-7500 for carbon dioxide and water vapor concentrations, R.M. Young 81000V for wind speed and sonic temperature). The 30-min averages were recorded for all of the measurements. At this site, the simulation period is from 15 April to 1 June 2003, when no precipitation events occurred.

The two TP sites are Shiquanhe and Gaize, both located in the western TP with elevations ~4000 m above sea level. Both sites, located in the midlatitude westerlies, belong to the alpine desert climate. The mean annual precipitation is around 200 mm and the land surface is almost bare soil. Owing to the high elevation and strong solar radiation, the surface heat fluxes and near-surface meteorological variables undergo especially evident diurnal variations. The measurements were collected through the Global Energy and Water Cycle Experiment (GEWEX) Asian Monsoon Experiment-Tibet (GAME-Tibet; Koike et al. 1999) during an intensive observing period (IOP, May–September 1998). The forcing data were recorded by automatic weather stations. The observed ground truth data included surface temperature, soil temperature profile (5, 10, 20, 40, and 80 cm), and soil moisture at 0–15 cm; turbulent fluxes, however, were not available. The surface temperature was directly measured using a thermometer, with half of the sensor buried in the soil and half exposed to the air. This technique is used routinely by the Chinese Meteorological Administration to measure the bare soil surface temperature. The surface temperature measured using this technique agrees with that converted from the measured longwave radiation (given a surface emissivity of 0.9, which is derived by assuming the thermometer measurements near sunset are reliable) at Shiquanhe, with an uncertainty of 2–3 K. The soil moisture was measured with a time domain reflectometry (TDR) soil water content hygrometer. Data averaged over 30 or 60 min were recorded. At the Shiquanhe site, the simulation period is from 1 May to 14 September 1998, when the amount of rainfall is only 25 mm. At the Gerze site, the simulation period is shorter

TABLE 2. Model parameters derived from observations at four sites. Note that a D in parentheses means a default value.

	Audubon	Dunhuang	Shiquanhe	Gaize
Albedo	0.24	0.27	0.24	0.28
Surface emissivity	0.93	0.92	0.90	0.91
Thermal diffusivity ($\text{m}^2 \text{s}^{-1}$)	2.7×10^{-7}	2.3×10^{-7}	3.7×10^{-7}	3.4×10^{-7}
Aerodynamic roughness length (m)	0.011 (D)	0.00067	0.011 (D)	0.011 (D)

(from 1 May to 15 June 1998), though no precipitation events occurred.

The last arid desert site is Dunhuang. The observed data were collected through the Field Experiment on Interaction between Land and Atmosphere in Arid Region of North-west China (NWC-ALIEX; Huang et al. 2002). Wind speed, air temperature, and humidity were measured on a tower at four heights: 1, 2, 8, and 18 m. Surface temperature was measured using the same technique as was used at Shiquanhe and Gaize. Soil temperatures were measured at six depths: 5, 10, 20, 40, 80, and 180 cm. Soil water contents were measured at depths of 5, 10, 20, and 80 cm with TDR. Downward and upward radiation flux components were measured at 1.5 m. During the IOP (10–25 June 2004), turbulent fluxes were measured by the eddy-covariance technique with 30-min bulk averaging. The simulation targets the period 18 May–25 June, when no precipitation was detected.

3. Land surface model and parameter settings

a. Model description

The Noah LSM is developed based on the Oregon State University (OSU) LSM, which includes a two-layer soil model with thermal conduction equations for soil temperature and the diffusive form of Richardson's equation for soil moisture (Mahrt and Pan 1984a), as well as a Penman approach for the calculation of the latent heat flux (Mahrt and Ek 1984b). After being widely adopted by NCEP, the Noah model has benefitted from a series of improvements. Among the major improvements are an increase from two to four soil layers, modifications to the canopy conductance formulation (Chen et al. 1996), bare soil evaporation and vegetation phenology (Betts et al. 1997), a new runoff formulation and

infiltration (Schaake et al. 1996), thermal roughness length treatment in the surface layer exchange coefficients (Chen et al. 1997), and the inclusion of cold season processes (Koren et al. 1999). A more detailed overview of the physics lineage of the Noah LSM is presented in Ek et al. (2003).

We presently employ version 2.7 of the Noah LSM. In general, the model has four soil layers (with depths of 10, 30, 60, and 100 cm from top to bottom), a single canopy layer, and a single snow layer. The vegetation types are defined according to the categories assigned from U.S. Geological Survey (USGS) database. The soil types are defined by the Food and Agriculture Organization (FAO) database. Soil moisture for each soil layer is calculated from the diffusive form of Richard's equation. Soil temperature is calculated from the heat diffusion equation. The surface temperature is determined following Mahrt and Ek (1984b) to reflect a linearly combined ground–vegetation surface. A more detailed description of the model governing equations and the parameterizations can be found in Chen and Dudhia (2001).

b. Land surface processes in arid regions

The surface energy balance (SEB) equation in Noah LSM can be written as

$$R_{\text{net}} = (1 - \alpha)S^{\downarrow} + \varepsilon(L^{\downarrow} - \sigma T_{\text{sfc}}^4) \quad \text{and} \quad (1a)$$

$$R_{\text{net}} = H + LE + G_0, \quad (1b)$$

where R_{net} is the net radiation. Equation (1a) is the radiation budget equation, where S^{\downarrow} and L^{\downarrow} are the downward solar and longwave radiation, respectively, providing

TABLE 3. The z_{0h} parameterization schemes selected for our sensitivity study: $\text{Re}_* = z_{0m}u_* / \nu$, $\text{Pr} = 0.71$, $k = 0.4$, ν is the fluid kinematical viscosity, $\alpha = 0.52$, and $C = 0.075$ in Z95.

Formula	Reference	Abbreviation
$z_{0h} = z_{0m} / (\text{Pr} \times \text{Re}_*)$	Sheppard (1958)	S58
$z_{0h} = z_{0m} \times \exp(2.0 - 2.46\text{Re}_*^{0.25})$	Brutsaert (1982)	B82
$z_{0h} = z_{0m} \times \exp(-1.0kC\text{Re}_*^{0.5})$	Zilitinkevich (1995)	Z95
$z_{0h} = z_{0m} \times \exp(-1.0k\alpha\text{Re}_*^{0.45})$	Zeng and Dickinson (1998)	Z98
$z_{0h} = z_{0m} \times \exp(2.0 - 1.29\text{Re}_*^{0.25})$	Kanda et al. (2007)	K07
$z_{0h} = (70\nu/u_*) \times \exp(-7.2u_*^{0.5} T_* ^{0.25})$	Yang et al. (2008)	Y08

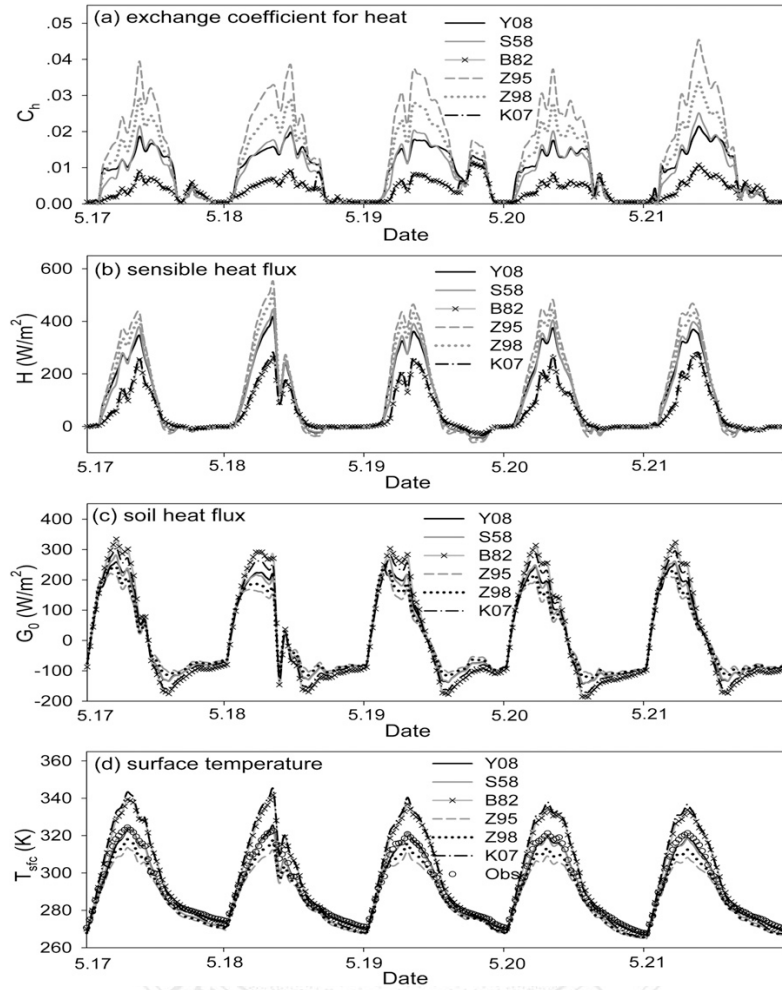


FIG. 1. Comparison of the diurnal variation of the (a) surface exchange coefficient, (b) sensible heat flux, (c) soil heat flux, and (d) surface temperature among simulations using different z_{0h} schemes: 1) Y08, 2) S58, 3) B82, 4) Z95, 5) Z98, and 6) K07 at Shiquanhe. Note that the observed T_{sfc} are available in (d).

inputs of the essential forcing; T_{sfc} is the land surface temperature; σ is the Stefan-Boltzmann constant; and α and ε are the surface albedo and the ground surface emissivity, respectively. Equation (1b) is the energy budget balance equation, where H is the turbulent sensible heat flux, LE is the turbulent latent heat flux, and G_0 is the surface soil heat flux. In this study, the negligible rainfall and the nonvegetated ground conditions at all sites lead to extremely high Bowen ratios; we thus omit LE in Eq. (1b). The other two terms, H and G_0 , should dominate the surface energy budget.

In the Noah LSM, the sensible heat flux is calculated through the bulk heat transfer equation:

$$H = -\rho c_p C_h u (\theta_{air} - \theta_{sfc}), \quad (2)$$

where ρ is the air density; c_p is the specific heat capacity of air at constant pressure; C_h is the surface exchange coefficient for heat; u is the wind speed; θ_{air} is the air temperature adjusted adiabatically for the height above the surface (z) and is approximately equal to $[T_{air} + (g/c_p) \times z]$, where g is gravitational constant; and θ_{sfc} is the corresponding variable at the surface. From Eq. (2), C_h is a crucial parameter in determining H . Usually, C_h can be obtained through the Monin-Obukhov similarity theory, and is mainly dependent on z_{0h} and z_{0m} , as well as the atmospheric stability.

TABLE 4. Determination coefficient (R^2), bias (BIAS), mean deviation (MD), and root-mean-square deviation (RMSD) between the observed T_{sfc} and the simulations using the schemes in Table 3 for Shiquanhe.

T_{sfc}	R^2	BIAS* (K)	MD* (K)	RMSD* (K)
S58	0.9454	-1.61	3.17	4.11
B82	0.9493	4.34	4.99	7.12
Z95	0.925	-6.65	6.78	8.02
Z98	0.9453	-3.8	4.44	5.37
K07	0.9493	4.34	4.99	7.11
Y08	0.9525	-1.63	2.99	3.89

* BIAS, MD, and RMSD are defined as $\text{BIAS} = [\sum_{i=1}^N (X_i - \text{Obs}_i)]/N$, $\text{MD} = [\sum_{i=1}^N |X_i - \text{Obs}_i|]/N$, and $\text{RMSD} = \sqrt{[\sum_{i=1}^N (X_i - \text{Obs}_i)^2]/N}$, where N is the sample number.

The surface soil heat flux can be written as

$$G_0 = k_T(\Theta_1) \frac{T_{\text{sfc}} - T_1}{h_1}, \quad (3)$$

where k_T is the soil thermal conductivity that depends on both soil water content (Θ) and soil type, T_1 is the soil temperature in the uppermost layer, and h_1 is equal to half of its depth. At present, Θ is very low with negligible temporal variations; therefore, the value of k_T is assumed to be a constant value at individual sites. As a result, G_0 mainly depends on the modeled T_{sfc} .

The turbulent sensible heat and ground heat fluxes are the dominant components in Eq. (1b). In addition to the input forcing data, the parameters in the radiation and energy budgets (1a) and (1b) would determine the surface energy flux partitioning and the surface temperature. Therefore, they should be set accurately.

c. Model settings

The vegetation types at four sites are prescribed as the bare soil according to the aforementioned characteristics of the ground. The soil type is derived from the FAO data. The soil hydraulic parameters are much less important for the modeling of the surface temperature and energy budget at dry sites than are the energy-related soil parameters and surface parameters (α , ε , z_{0h} , z_{0m} , and k_T). Among them, α , ε , and k_T at all sites and z_{0m} at Dunhuang can be derived from the observations, and their mean values are given in Table 2. The parameter k_T was derived from the thermal diffusivity, which was estimated from the diurnal range of observed soil temperatures profile instead of a parameterization; α was directly obtained from observed downward and upward shortwave radiation fluxes; and ε was derived from surface temperature and longwave radiation fluxes.

In addition, z_{0m} is physically related to the geometric roughness of surface elements and can be derived from

TABLE 5. Sensitivity test of the simulated surface temperature (T_{sfc}) and sensible heat flux (H) to the aerodynamic roughness length (z_{0m}). Comparative statistics are calculated between $z_{0m} = 0.011$ m and $z_{0m} = 0.001$ m for Shiquanhe.

	T_{sfc}			H		
	BIAS (K)	MD (K)	RMSD (K)	BIAS (W m^{-2})	MD (W m^{-2})	RMSD (W m^{-2})
S58	0.24	0.25	0.32	-1.31	2.04	3.17
B82	0.32	0.32	0.39	-1.75	2.19	3.5
Z95	0.53	0.54	0.73	-2.94	4.82	7.47
Z98	0.39	0.4	0.53	-2.17	3.57	5.34
K07	0.3	0.31	0.4	-1.66	2.37	3.65
Y08	0.22	0.23	0.31	-1.22	2.28	3.44

the wind speed and temperature profiles. An optimal method suggested by Yang et al. (2008) was employed to estimate z_{0m} at Dunhuang, where the profile data are available. The default value of z_{0m} prescribed by vegetation type was used at the other three sites. In the Noah LSM, z_{0h} is calculated by the Reynolds number-dependent scheme of Zilitinkevich (1995), as shown in Table 3. Yang et al. (2008) argued that this scheme overestimated z_{0h} and thus underestimated the peak values of T_{sfc} . However, considering the significance of z_{0h} and z_{0m} , the sensitivity of the surface energy budget to them will be investigated in section 4. In all simulations, the soil moisture and soil temperature are initialized with the observations.

4. Sensitivity test to the roughness lengths

Given the similarity among four arid sites, we regard Shiquanhe as being representative of the arid sites for a sensitivity analysis.

a. Thermal roughness length

To test the sensitivity of surface energy budget to different z_{0h} values, six z_{0h} schemes available in the literature were implemented into the Noah LSM. Listed in Table 3 are S58 (Sheppard 1958), B82 (Brutsaert 1982), Z95 (Zilitinkevich 1995), Z98 (Zeng and Dickinson 1998), K07 (Kanda et al. 2007), and Y08 (Yang et al. 2008). S58 and B82 were examined in detail in Verhoef et al. (1997). Z95 has been widely used in NCEP operational prediction systems since Chen et al. (1997). Z98 has been used in an LSM to unify undercanopy heat transfer processes between dense and sparse canopies. K07 was derived from urban canopy experiments. Y08 will be introduced in section 5a. In Yang et al. (2008), these schemes have been evaluated within the framework of Monin–Obukhov similarity theory by using observed T_{sfc} to parameterize H . In this study, their effectiveness is tested against independent datasets within the framework of land surface modeling.

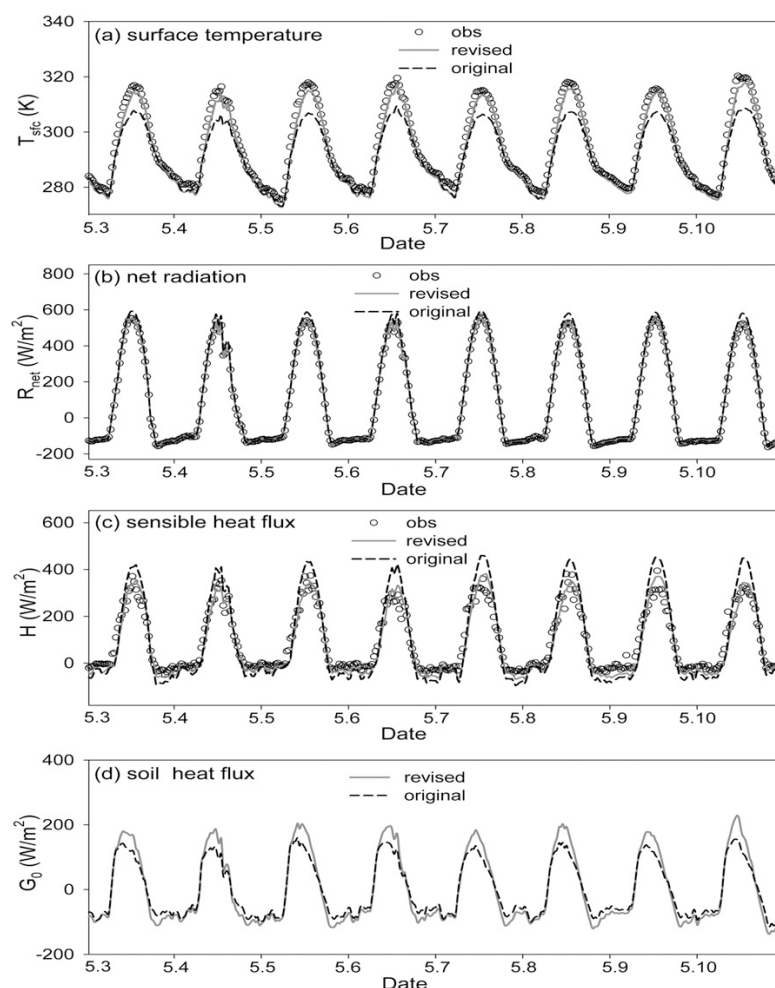


FIG. 2. Comparison of the diurnal variation of the (a) surface temperature, (b) net radiation, (c) sensible heat flux, and (d) soil heat flux between two simulations by the revised Noah LSM and the original version against the observations for the Audubon site. Circles represent the observations, the dark gray line represents the simulations by revised the Noah model, and dashed black line represents the simulations by the original model.

F1 Figure 1a compares the diurnal variation of C_h obtained using six schemes at Shiquanhe. During the simulation period, the highest C_h values of Z95 are about four times the lowest values of B82, the C_h values of Y08 are comparable to those of S58, and the results of K07 are very close to those of B82. Apparently, as in the previous studies, C_h is very sensitive to the parameterization schemes of z_{0h} in arid regions.

In Figs. 1b and 1c, we compare dominant heat flux components simulated by different z_{0h} schemes. As expected, differences in the simulated sensible heat flux arise from those in C_h values; high (low) C_h values correspond to high (low) H , and low (high) G_0 are found

accordingly. It is easy to interpret this phenomenon. As discussed in section 3, G_0 is mainly determined by the modeled T_{sfc} . Therefore, if H is overestimated, T_{sfc} , which is calculated diagnostically from surface energy balance considerations, will be underestimated and thus G_0 will be underestimated.

Figure 1d compares the diurnal variation of T_{sfc} between the observations and simulations with different z_{0h} schemes. The simulations with the Y08 and S58 schemes produce good agreement with the observed daytime surface temperature, while the simulations with other schemes produce clear biases. Table 4 gives the comparative statistics between the observed T_{sfc} and the

T4

TABLE 6. The error metrics of the difference between the half-hourly observations and the simulated results at four sites from the revised Noah model and the original version, respectively: BIAS, MD, and RMSD. The observed sensible heat flux was not available at Shiquanhe and Gaize.

		T_{sfc}			H			R_{net}		
		BIAS (K)	MD (K)	RMSD (K)	BIAS (W m ⁻²)	MD (W m ⁻²)	RMSD (W m ⁻²)	BIAS (W m ⁻²)	MD (W m ⁻²)	RMSD (W m ⁻²)
Audubon	Original	-4.05	4.06	5.16	-1.31	53.64	54.22	18.13	18.13	25.35
	Revised	-1.85	1.96	2.35	-13.57	25.45	34.64	3.94	6.41	8.69
Dunhuang	Original	-2.78	3.15	4.06	8.77	25.58	36.69	9.84	13.04	18.83
	Revised	-1.82	2.33	2.98	1.72	18.37	26.02	3.15	10.69	14.45
Shiquanhe	Original	-6.44	6.74	8.18	—	—	—	43.4	45.06	66.01
	Revised	-1.63	2.99	3.89	—	—	—	13.09	19.92	29.13
Gaize	Original	-5.58	5.62	7.07	—	—	—	25.37	26.77	41.49
	Revised	-2.16	2.67	3.44	—	—	—	4.32	10.66	16.1

simulations using various schemes; we thus confirm that Y08 reproduces T_{sfc} more consistently with observed data than do the other schemes in this case.

b. Aerodynamic roughness length

The aerodynamic roughness length is ideally determined from the wind speed profile, although there have been some successes in relating this height to the arrangement, spacing, and physical height of individual roughness elements. The lack of profile data makes it hard to precisely estimate z_{0m} . The common approach is to empirically prescribe a value of z_{0m} for a given vegetation type.

The sensitivity of H and T_{sfc} to z_{0m} is also investigated. In the simulations, we tried out two z_{0m} values: a default value of z_{0m} (0.011 m) and a lower value (0.001 m) at Shiquanhe, in combination with different z_{0h} schemes in Table 3. Table 5 shows error metrics between the simulations produced by the six schemes. Clearly, different z_{0m} values produce minor differences both in T_{sfc} and H ; H and T_{sfc} , therefore, are not reasonably sensitive to z_{0m} . Moreover, the Y08 scheme produces lower differences, indicating its particularly low sensitivity to the choice of z_{0m} .

In summary, the surface flux and temperature simulations are highly sensitive to the z_{0h} schemes and much less sensitive to the z_{0m} value; the Y08 scheme seems to be a promising scheme since it can appropriately reproduce T_{sfc} . So, we applied the Y08 scheme to update the Noah LSM, and made further evaluations against the observations.

5. Updating Noah with the Y08 scheme and evaluations

a. Brief introduction to the Y08 scheme

The parameterization of the thermal surface roughness length is crucial for directly using T_{sfc} to calculate H . Many works (see Table 3) have related z_{0h} directly to z_{0m} through the parameter kB^{-1} [defined as $\ln(z_{0m}/z_{0h})$]. Momentum transport is generally more efficient than heat transport, due to the influence of pressure fluctuation, because individual roughness elements may enhance the momentum flux through form drag with little contribution to the area-averaged heat flux (Mahrt 1996). Therefore, z_{0h} is typically less than z_{0m} , especially over a surface with bluff roughness elements, and a higher z_{0m} usually corresponds to a lower z_{0h} .

TABLE 7. As in Table 6, but for the daytime (0900–1600 local time).

		T_{sfc}			H			R_{net}		
		BIAS (K)	MD (K)	RMSD (K)	BIAS (W m ⁻²)	MD (W m ⁻²)	RMSD (W m ⁻²)	BIAS (W m ⁻²)	MD (W m ⁻²)	RMSD (W m ⁻²)
Audubon	Original	-8.08	8.09	8.39	63.42	69.26	80.22	38.39	38.62	41.47
	Revised	-2.19	2.28	2.68	-1.91	34.23	41.62	-0.23	5.96	7.69
Dunhuang	Original	-4.99	5.15	5.93	38.36	47.84	57.29	19.48	22.11	27.89
	Revised	-2.44	2.96	3.86	14.48	30.52	38.4	1.22	16.08	19.97
Shiquanhe	Original	-11.15	11.29	12.21	—	—	—	99.54	100.92	107.63
	Revised	-1.66	3.82	5.18	—	—	—	35.59	41.07	46.24
Gaize	Original	-10.18	10.19	10.91	—	—	—	64.01	64.01	68.44
	Revised	-2.62	3.73	4.76	—	—	—	15.48	20.78	25.53

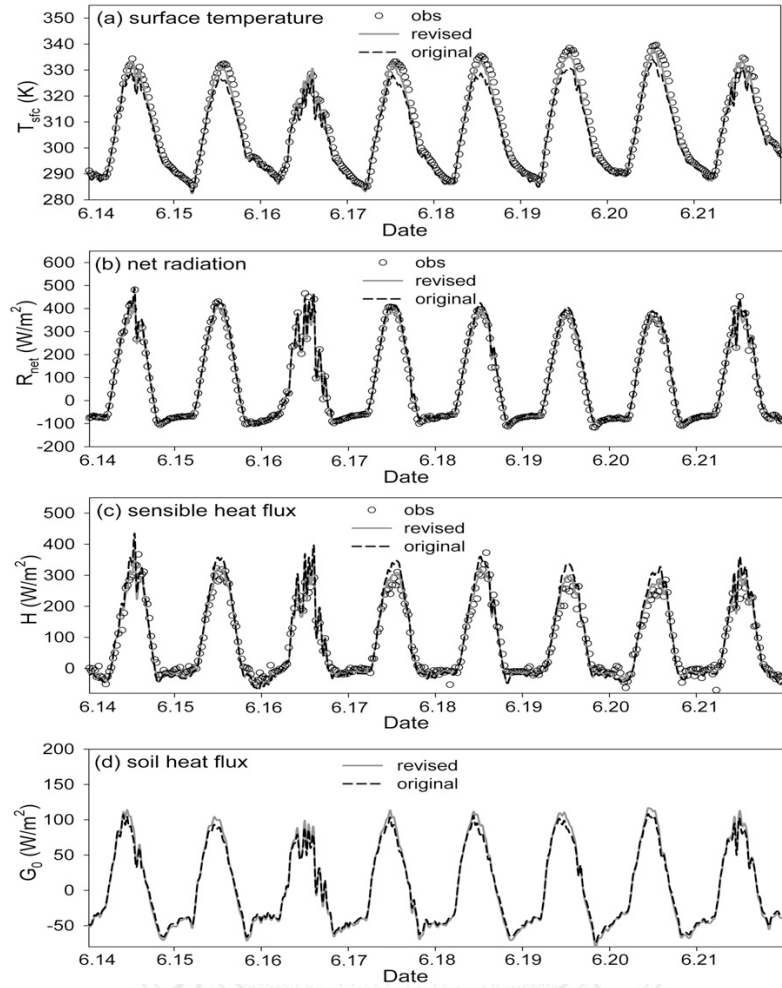


FIG. 3. As in Fig. 2, but for Dunhuang.

Following this reasoning, Yang et al. (2002) correlated z_{0h} to a physical height (h_T), which is related to a height to separate the fully turbulent layer and the transitional layer. The height h_T is determined by the critical Reynolds number (Re_{crit}):

$$h_T = \frac{\nu Re_{crit}}{u_*}, \quad (4)$$

where $Re_{crit} = 70$ in this study, ν is the fluid kinematic viscosity, and u_* is the friction velocity.

For a surface with bluff roughness elements, u_* is quite large due to form drag and, therefore, gives a small h_T . So the variation of h_T is similar to that of z_{0h} , making it reasonable to use h_T as a length to scale z_{0h} . Therefore, Yang et al. (2002) defined the parameter

$$kA^{-1} = \ln \frac{h_T}{z_{0h}}. \quad (5)$$

Based on data analysis for three TP sites, they found the typical diurnal variation of z_{0h} , and then assumed the following form for kA^{-1} :

$$kA^{-1} = \beta u_*^m |\theta_*|^n, \quad (6)$$

where β , m , and n are coefficients. Data analysis indicates that $m = 1/2$, $n = 1/4$ are reasonable values. Yang et al. (2008) suggested the use of $\beta = 7.2 \text{ m}^{-1/2} \text{ s}^{1/2} \text{ K}^{-1/4}$. Combining Eqs. (4)–(6), they obtained a general expression of z_{0h} for bare-soil or short-vegetation surfaces:

$$z_{0h} = \frac{70\nu}{u_*} \exp(-\beta u_*^{1/2} |\theta_*|^{1/4}). \quad (7)$$

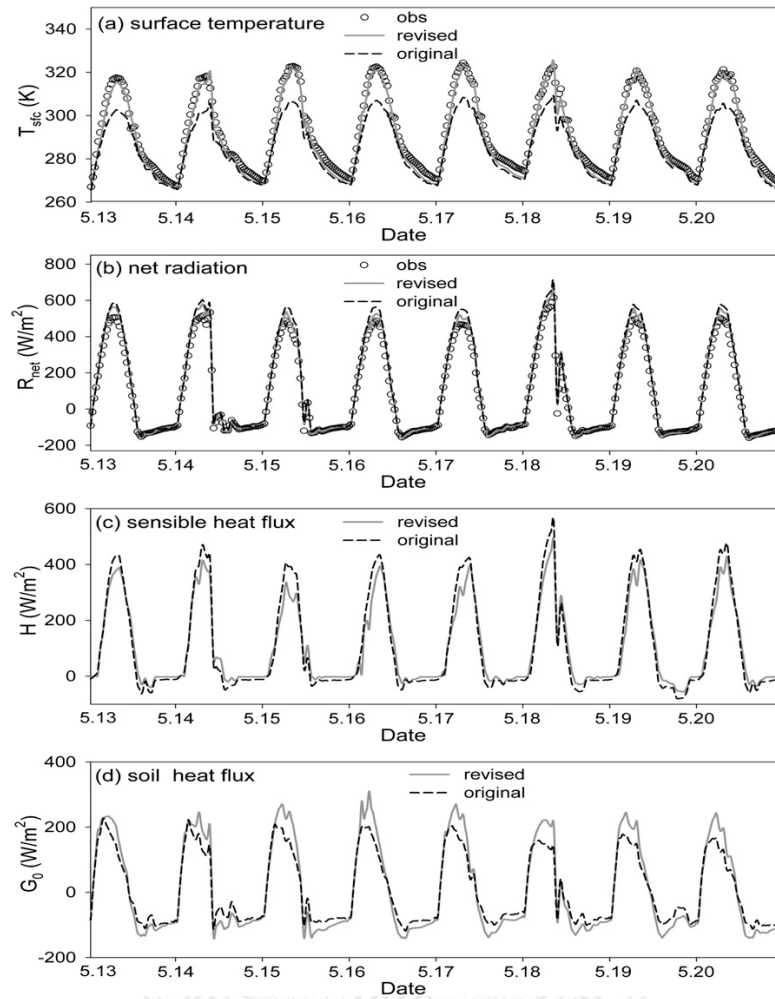


FIG. 4. Comparison of the diurnal variation of the (a) surface temperature, (b) net radiation, (c) sensible heat flux, and (d) soil heat flux from two simulations by the revised Noah and the original version for Shiquanhe.

More detailed information on Y08 and its evolution can be found in Yang et al. (2002) and Yang et al. (2008).

b. Evaluations

Evaluations of both the revised Noah LSM and the original version were performed first at Audubon and Dunhuang, where surface temperature and sensible heat flux data are available. Then, they were evaluated at Shiquanhe and Gaize, where only surface temperature data are available.

1) AUDUBON AND DUNHUANG

F2 Figures 2a–d compare the simulated diurnal variations of T_{sfc} , R_{net} , H , and G_0 , respectively, between the revised

Noah LSM and the original version at Audubon Research Ranch. Obviously, T_{sfc} and H , as well as R_{net} , were properly simulated by the revised model, while the original model produced higher H and lower T_{sfc} , and, thus, higher R_{net} . Table 6 gives error indices, which indicate that the revised model significantly reduced the simulation errors. The error metrics for T_{sfc} , R_{net} , and H are calculated using data during the whole simulation period.

The major difference between the two simulations occurs in the daytime. Table 7 shows the error metrics using data during 0900–1600 local time (LT) when the original model yields higher H , higher R_{net} , and lower T_{sfc} . The lower T_{sfc} would directly result in lower G_0 (see Fig. 2d), corresponding to higher H .

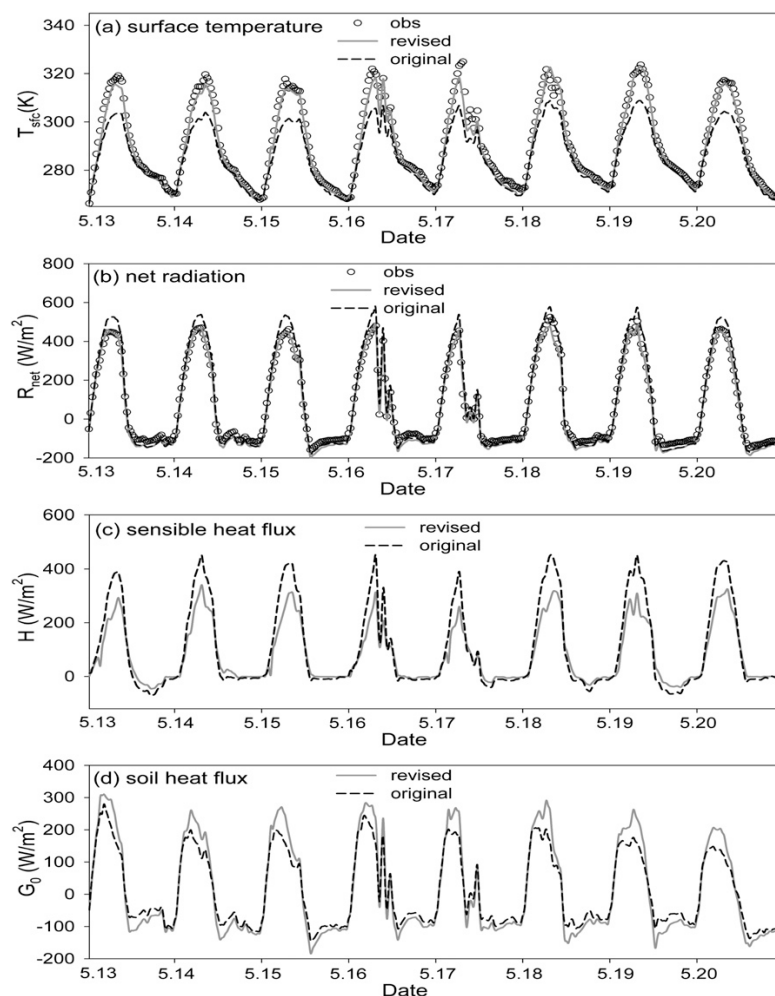


FIG. 5. As in Fig. 4, but for Gaize.

F3 Figures 3a–d compare the diurnal variations of the components in the energy balance equation at Dunhuang. As shown in Figs. 3a–c, the revised Noah LSM can better simulate T_{sfc} , R_{net} , and H than could the original version. This is confirmed by the comparative statistics in Tables 6 and 7.

2) SHIQUANHE AND GAIZE

F4|F5 Figures 4 and 5 give the comparisons at Shiquanhe and Gaize, respectively. Evidently, the original model significantly underestimated T_{sfc} and overestimated R_{net} , while the revised model simulated well both T_{sfc} and R_{net} . These results are confirmed by the error metrics in Table 6. The improvements by the revised model are especially evident in the daytime. In Table 7 we find that the original model underestimates T_{sfc} by more than 10 K

during 0900–1600 LT, while the revised model performs much better. Moreover, we found that the original model performed worse at two TP sites than at Audubon and Dunhuang; this finding will be discussed in more depth in section 5c.

c. Discussion

To understand the different levels of performance between the revised Noah LSM and the original model, Fig. 6 shows the simulated mean diurnal variation of $\ln(z_{0h})$ throughout the simulation period at four sites. It is clear that the values of z_{0h} simulated by the revised model exhibit evident diurnal variations. In fact, several studies have reported the diurnal variations of z_{0h} over the bare-soil surface and grasslands (Verhoef et al. 1997; Sun 1999; Ma et al. 2002; Yang et al. 2003). But this

F6

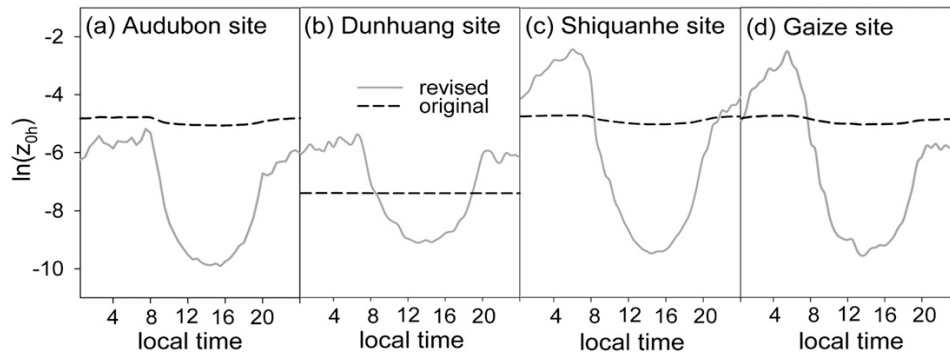


FIG. 6. Simulated mean diurnal variation of $\ln(z_{0h})$ throughout the simulation period at (a) Audubon, (b) Dunhuang, (c) Shiquanhe, and (d) Gaize.

diurnal variation was hardly simulated by the original Noah LSM. Figure 6 shows that the original model produces rather high z_{0h} , and thus fairly high C_h , compared to those produced by the revised model in the daytime. Moreover, the originally simulated T_{sfc} is lower because the too high C_h carries too much heat away from the surface. Two direct effects arise from the underestimate of T_{sfc} : one is the overestimate of R_{net} due to the reduced upward longwave radiation flux; the other is the underestimate of G_0 due to an underestimated soil temperature gradient.

Figure 6 also shows that the values of z_{0h} produced by the revised model have larger diurnal ranges at two TP sites than elsewhere. This is consistent with Yang et al. (2008), who found that the diurnal variations of z_{0h} at TP sites are more evident than at other sites. This phenomenon may be attributed to strong diurnal changes of near-surface meteorological variables and higher land-atmosphere temperature differences caused by the high elevation and thus strong solar radiation. Because the differences between the two simulated z_{0h} values at the TP sites are also larger than at Dunhuang and Audubon during the daytime, the improvements in the revised Noah are believed to be particularly meaningful at two high-elevation sites, as shown in Tables 6 and 7.

6. Conclusions

In this paper we investigated the importance of surface flux parameterization in simulating land surface processes in arid regions. Six thermal surface roughness length schemes were intercompared in the Noah LSM and then the revised version, with one promising z_{0h} scheme evaluated at four arid or semiarid sites. Our major findings are as follows.

The parameterization of z_{0h} is crucial for modeling T_{sfc} and the surface energy budget in arid regions.

Sensitivity tests for the six selected schemes confirm foregoing studies that found the daytime T_{sfc} is sensitive to the parameterization. If z_{0h} is overestimated (underestimated), T_{sfc} would be underestimated (overestimated). In addition, H would be overestimated (underestimated), while G_0 would be underestimated (overestimated).

The Noah LSM, originally using Z95 as z_{0h} its parameterization scheme, produces unseasonably high z_{0h} during the day in arid regions, which leads to overestimated H and underestimated T_{sfc} in the daytime. By implementing the Y08 z_{0h} scheme, the revised Noah model can well simulate T_{sfc} and the surface energy budget simultaneously. Given the wide usage of the Noah LSM in NCEP numerical models, further efforts are warranted to examine potential improvements in synoptic or climatic simulations adopting the presently updated land surface parameterization.

Acknowledgments. We are grateful to three reviewers whose comments have helped the authors to improve the quality of the paper. This work was supported by the National Basic Research Program of China (2009CB421405), Chinese Academy of Sciences (Innovation Project KZCX2-YW-Q11-01 and the “100-Talent” Program), and the National Natural Science Foundation of China (40875009). Audubon Research Ranch data were obtained through the AmeriFlux network. The Shiquanhe and Gaize data used in this paper were obtained through the GAME/Tibet project, which was supported by MEXT, FRSGC, NASDA of Japan, the Chinese Academy of Science, and the Asian Pacific Network. The Dunhuang data were obtained through the Field Experiment on Interaction between Land and Atmosphere in Arid Region of Northwest China. XG gratefully acknowledges the support of the K. C. Wong Education Foundation, Hong Kong.

REFERENCES

- Betts, A. K., F. Chen, K. E. Mitchell, and Z. Janjić, 1997: Assessment of the land surface and boundary layer models in the two operational versions of the NCEP Eta Model using FIFE data. *Mon. Wea. Rev.*, **125**, 2896–2916.
- Brutsaert, W. H., 1982: *Evaporation Into the Atmosphere: Theory, History, and Applications*. D. Reidel, 299 pp.
- Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the PSU–NCAR MM5 modeling system. Part I: Model description and implementation. *Mon. Wea. Rev.*, **129**, 569–585.
- , and Y. Zhang, 2009: On the coupling strength between the land surface and the atmosphere: From viewpoint of surface exchange coefficients. *Geophys. Res. Lett.*, **36**, L10404, doi:10.1029/2009GL037980.
- , and Coauthors, 1996: Modeling of land-surface evaporation by four schemes and comparison with FIFE observations. *J. Geophys. Res.*, **101**, 7251–7268.
- , Z. Janjić, and K. E. Mitchell, 1997: Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model. *Bound.-Layer Meteor.*, **85**, 391–421.
- Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah LSM advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.*, **108**, 8851, doi:10.1029/2002JD003296.
- Hogue, T. S., L. Bastidas, H. Gupta, S. Sorooshian, K. Mitchell, and W. Emmerich, 2005: Evaluation and transferability of the Noah LSM in semiarid environments. *J. Hydrometeor.*, **6**, 68–83.
- Huang, R. H., and Coauthors, 2002: The Field Experiment on Air–Land Interaction in the Arid Area of Northwest China (NWC-ALIEX) and preliminary scientific achievements of this experiment. *Interaction Workshop on the Air–Land Interaction in Arid and Semiarid Areas and Its Impact on Climate (IWALI)*, Dunhuang, China, 29–41.
- Kanda, M., M. Kanega, T. Kawai, R. Moriawaki, and H. Sugawara, 2007: Roughness lengths for momentum and heat derived from outdoor urban scale models. *J. Appl. Meteor. Climatol.*, **46**, 1067–1079.
- Koike, T., Yasunari, J. Wang, and T. Yao, 1999: GAME-Tibet IOP summary report. *Proc. First Int. Workshop on GAME-Tibet*, Xi'an, China, Chinese Academy of Sciences and Japanese National Committee for GAME, 1–2.
- Koren, V., J. C. Schaake, K. E. Mitchell, Q. Y. Duan, F. Chen, and J. Baker, 1999: A parameterization of snowpack and frozen ground intended for NCEP weather and climate models. *J. Geophys. Res.*, **104**, 19 569–19 585.
- LeMone, M. A., M. Tewari, F. Chen, J. Alfieri, and D. Niyogi, 2008: Evaluation of the Noah land surface model using data from a fair-weather IHOP_200 day with heterogeneous surface fluxes. *Mon. Wea. Rev.*, **136**, 4915–4941.
- Ma, Y.-M., O. Tsukamoto, J. Wang, H. Ishikawa, and I. Tamagawa, 2002: Analysis of aerodynamic and thermodynamic parameters over the grassy marshland surface of Tibetan Plateau. *Prog. Natl. Sci.*, **12**, 36–40.
- Ma, Z., and C. Fu, 2006: Some evidences of drying trend over North China from 1951 to 2004. *Chin. Sci. Bull.*, **51**, 2913–2925.
- Mahrt, L., 1996: The bulk aerodynamic formulation over heterogeneous surfaces. *Bound.-Layer Meteor.*, **78**, 87–119.
- , and M. Ek, 1984: The influence of atmospheric stability on potential evaporation. *J. Climate Appl. Meteor.*, **23**, 222–234.
- , and H.-L. Pan, 1984: A two-layer model of soil hydrology. *Bound.-Layer Meteor.*, **29**, 1–20.
- Puigdefabregas, J., 1995: Desertification: Stress beyond resilience, exploring a unifying process structure. *Ambio*, **24**, 311–313.
- Schaake, J. C., V. I. Koren, Q.-Y. Duan, K. E. Mitchell, and F. Chen, 1996: Simple water balance model for estimating runoff at different spatial and temporal scales. *J. Geophys. Res.*, **101**, 7461–7475.
- Sheppard, P. A., 1958: Transfer across the earth's surface and through the air above. *Quart. J. Roy. Meteor. Soc.*, **84**, 205–224.
- Smeets, C. J. P. P., and M. R. van den Broeke, 2008: The parameterisation of scalar transfer over rough ice. *Bound.-Layer Meteor.*, **128**, 339–355.
- Sun, J., 1999: Diurnal variations of thermal roughness height over a grassland. *Bound.-Layer Meteor.*, **92**, 407–427.
- Verhoef, A., H. A. R. de Bruin, and B. J. J. M. van den Hurk, 1997: Some practical notes on the parameter kB^{-1} for sparse vegetation. *J. Appl. Meteor.*, **36**, 560–572.
- Warren, A., 1996: Desertification. *The Physical Geography of Africa*, W. M. Adams, A. S. Goudie, and A. R. Orme, Eds., Oxford University Press, 342–355.
- Yang, K., T. Koike, H. Fujii, K. Tamagawa, and N. Hirose, 2002: Improvement of surface flux parameterizations with a turbulence-related length. *Quart. J. Roy. Meteor. Soc.*, **128**, 2073–2087.
- , —, and D. Yang, 2003: Surface flux parameterization in the Tibetan Plateau. *Bound.-Layer Meteor.*, **106**, 245–262.
- , and Coauthors, 2007: Initial CEOP-based review of the prediction skill of operational general circulation models and land surface models. *J. Meteor. Soc. Japan*, **85**, 99–116.
- , and Coauthors, 2008: Turbulent flux transfer over bare soil surfaces: Characteristics and parameterization. *J. Appl. Meteor. Climatol.*, **47**, 276–290.
- , Y.-Y. Chen, and J. Qin, 2009: Some practical notes on the land surface modeling in the Tibetan Plateau. *Hydrol. Earth Syst. Sci.*, **13**, 687–701.
- Zeng, X., and R. E. Dickinson, 1998: Effect of surface sublayer on surface skin temperature and fluxes. *J. Climate*, **11**, 537–550.
- Zilitinkevich, S. S., 1995: Non-local turbulent transport: Pollution dispersion aspects of coherent structure of convective flows. *Air Pollution Theory and Simulation*, H. Power, N. Moussopoulos, and C. A. Brebbia, Eds., *Air Pollution III*, Vol. I, Computational Mechanics Publications, 53–60.

Other Key Reference Papers for Prof. Yang's Lecture:

- [1] Yang, K., Y. Chen, and J. Qin (2009), Some practical notes on the land surface modeling in the Tibetan Plateau, *Hydrol. Earth Syst. Sci.*, **13**, 687–701

An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction

Newsha K. Ajami,¹ Qingyun Duan,² and Soroosh Sorooshian³

Received 20 November 2005; revised 27 August 2006; accepted 19 September 2006; published 9 January 2007.

[1] The conventional treatment of uncertainty in rainfall-runoff modeling primarily attributes uncertainty in the input-output representation of the model to uncertainty in the model parameters without explicitly addressing the input, output, and model structural uncertainties. This paper presents a new framework, the Integrated Bayesian Uncertainty Estimator (IBUNE), to account for the major uncertainties of hydrologic rainfall-runoff predictions explicitly. IBUNE distinguishes between the various sources of uncertainty including parameter, input, and model structural uncertainty. An input error model in the form of a Gaussian multiplier has been introduced within IBUNE. These multipliers are assumed to be drawn from an identical distribution with an unknown mean and variance which were estimated along with other hydrological model parameters by a Monte Carlo Markov Chain (MCMC) scheme. IBUNE also includes the Bayesian model averaging (BMA) scheme which is employed to further improve the prediction skill and address model structural uncertainty using multiple model outputs. A series of case studies using three rainfall-runoff models to predict the streamflow in the Leaf River basin, Mississippi, are used to examine the necessity and usefulness of this technique. The results suggest that ignoring either input forcings error or model structural uncertainty will lead to unrealistic model simulations and incorrect uncertainty bounds.

Citation: Ajami, N. K., Q. Duan, and S. Sorooshian (2007), An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction, *Water Resour. Res.*, 43, W01403, doi:10.1029/2005WR004745.

1. Introduction

[2] Various hydrologic rainfall-runoff models have been used to represent the watershed physical processes which control the conversion of precipitation into streamflow and water storage changes. These models include many parameters describing the properties of the watershed that need to be estimated through calibration against historical observation data. For many years, research effort has been devoted to develop techniques to find the proper estimates of the parameter values that enable the model predictions to match the watershed observations. One major weakness of this parameter-calibration approach is that the objective function used to calibrate the model parameters implicitly assumes that all sources of uncertainties in the modeling process can be attributed to parameter errors. In fact, in addition to parameter uncertainty, model predictions are affected by many other uncertainties from various sources, among them the errors in model input (forcing) data such as

the precipitation observation data, the description of boundary and initial conditions, and the model structural deficiencies. Because of the highly nonlinear nature of the hydrologic system, it is not feasible to account for all these uncertainties from different sources through model parameter adjustments.

[3] Recently, hydrologic research [Beven and Binley, 1992; Kuczera and Parent, 1998; Vrugt et al., 2003; Marshall et al., 2003; Montanari and Brath, 2004] began to analyze various uncertainty sources in hydrological modeling. New techniques have made significant progress in estimating the propagation of confidence bounds from different uncertainty sources to the model output. Among them include the use of data assimilation techniques to tackle uncertainty in boundary and initial conditions [Kitanidis and Bras, 1980a, 1980b; Beck, 1987; Evenson, 1992; Miller et al., 1994]; simultaneous data assimilation and parameter estimation [Moradkhani et al., 2005]; and simultaneous uncertainty estimation of input (forcing) data and parameter estimation [Kavetski et al., 2003]. Most of these studies focus on addressing one or two uncertainty sources based on a selected hydrologic model. However, by using a single model, those techniques (which do not change the model structures) are unable to account for the errors in model output resulting from the structural deficiencies of the specific model.

¹Berkeley Water Center, Department of Civil and Environmental Engineering, University of California, Berkeley, California, USA.

²Lawrence Livermore National Laboratory, Livermore, California, USA.

³Department of Civil and Environmental Engineering, University of California, Irvine, Irvine, California, USA.

[4] Lately a new scheme has emerged which seeks to obtain a consensus from a combination of multiple model predictions so that one model's output errors can be compensated by others'. The combination techniques can be categorized into two groups. The first group [e.g., *Shamseldin et al.*, 1997; *Abrahart and See*, 2002; *Georgakakos et al.*, 2004; *Ajami et al.*, 2005, 2006] uses a set of deterministic weights to combine multiple model outputs. Methods of simple model average (equal weights), linear regression, or artificial neural network (ANN) belong to this category. The consensus prediction from these methods is an alternative deterministic prediction without uncertainty estimates. In addition, the weights in such combination can take any arbitrary real (positive or negative) values that lack physical interpretations.

[5] The second group such as Bayesian model averaging (BMA) [*Madigan et al.*, 1996; *Hoeting et al.*, 1999] uses probabilistic techniques which derive the consensus prediction from competing predictions using likelihood measures as model weights. The likelihood measure (weight) for each member model is based on the success frequency of the predictions that an individual model has made within the observations. For this reason, BMA weights are tied directly to individual model performance. BMA has been applied in a variety of fields including statistics, management science, medicine, and meteorology [e.g., *Viallefont et al.*, 2001; *Fernandez et al.*, 2001; *Raftery et al.*, 2003, 2005; *Wintle et al.*, 2003]. In many case studies, the BMA has shown to produce more accurate and reliable predictions than other multimodel techniques [*George and McCulloch*, 1993; *Raftery et al.*, 1997; *Clyde*, 1999; *Viallefont et al.*, 2001; *Raftery and Zheng*, 2003; *Ellison*, 2004]. Very recently, the BMA method was applied to hydrologic groundwater modeling [*Neuman and Wierenga*, 2003; *Neuman*, 2003].

[6] The intent of this study is to build a hybrid framework, Integrated Bayesian Uncertainty Estimator (IBUNE), to confront the uncertainties in rainfall-runoff predictions associated with input errors, model parameters estimates, and model structural deficiencies. To accomplish this objective, the paper is divided into three major parts. First, the Shuffled Complex Evolution Metropolis (SCEM) algorithm [*Vrugt et al.*, 2003], which was developed for probabilistic parameter estimation, will be studied. We will demonstrate that not accounting for existing error in the input and model structure could lead to corrupted parameter estimations, as well as unreliable uncertainty bounds on the model predictions. The second part of the paper presents a simple approach to extend SCEM to simultaneously account for the uncertainties originating from both input precipitation data and the model parameters. This is the first step toward building IBUNE. We will demonstrate that the error incorporated within the input (forcing) data is one of the major uncertainty sources in the rainfall-runoff modeling system, and by accounting for it within our uncertainty assessment procedure, we will improve the uncertainty bounds in model prediction. We will also show that not assessing model structural uncertainty is still an important limitation of this part of the study.

[7] Finally, the intent of the third part of this paper is to consider model structural uncertainty in addition to input and parameter uncertainty. We present a hybrid approach where we merge the strengths of the Bayesian model

averaging scheme with the extended SCEM. This is the final step in building the new framework, called IBUNE. IBUNE further reduces the uncertainties caused by the deficiencies in individual models by using Bayesian model averaging, while also accounting for input and parameter uncertainty within individual models by applying extended SCEM. Finally, the IBUNE scheme will be applied to a real case study in the Leaf River basin.

2. Study Basin and Hydrological Models

[8] We have selected the Leaf River basin to demonstrate the performance of presented frameworks in this study. This 1949-km² basin is located north of Collins, Mississippi. Five years of daily historical data (1953–1957), including precipitation (millimeters per 6 hours), potential evapotranspiration (mm/d), and streamflow (m³/s) were used for calibration and uncertainty assessment. Since many other studies were conducted over the period of 1953–1957 [*Yapo et al.*, 1998; *Gupta et al.*, 1998; *Hogue et al.*, 2003; *Vrugt et al.*, 2003], for comparison purposes we selected the same period for this study. To reduce the sensitivity to initial state variables, a 365-day (through water year 1952) warm-up period was used, during which no calibration and uncertainty estimation was performed for any of the under study hydrologic models. Three hydrologic models were selected for this study including the Sacramento soil moisture accounting (SAC-SMA) model [*Burnash et al.*, 1973], the hydrologic model (HYMOD) [*Boyle*, 2001], and the simple water balance (SWB) [*Schaake et al.*, 1996] model.

[9] SAC-SMA is a nonlinear, time-continuous, and conceptual rainfall-runoff model [*Burnash et al.*, 1973] and is being used operationally by many of the U.S. National Weather Service River Forecast Centers (NWS-RFC) for flood forecasting. The model includes two soil moisture layers, an upper and lower zone (Figure 1). This model includes 16 parameters, three of which were fixed at specified values; the remaining 13 parameters need to be determined through some kind of search process.

[10] Because the Leaf River basin has been studied extensively for optimization purposes [e.g., *Yapo et al.*, 1998; *Gupta et al.*, 1998; *Thiemann et al.*, 2001; *Hogue et al.*, 2003], we have gained a very good knowledge of what SAC-SMA parameter values should be for this basin. In order to minimize the interaction between various parameters in the SAC-SMA model and hence reduce the complications due to the nonidentifiability problem, which could cast shadow over the main objectives of this work, the SAC-SMA model was simplified. First we fixed five percolation parameters to prespecified values (Table 1). Further, we maintained the relative values of the parameters associated with the lower zone and the upper zone. Consequently the number of parameters in the SAC-SMA model that need to be identified was decreased to five: upper zone tension water maximum storage (UZTWM); upper zone free-water maximum storage (UZFWM); upper zone free-water lateral depletion rate (UZK); lower zone total maximum storage (LZTM); and lower zone supplementary free-water depletion rate (LZSK). LZTM represents the summation of all lower zone storages. The lower zone primary free-water depletion rate (LZPK) is estimated to be 3% of the lower zone supplemental free-water depletion rate (LZSK).

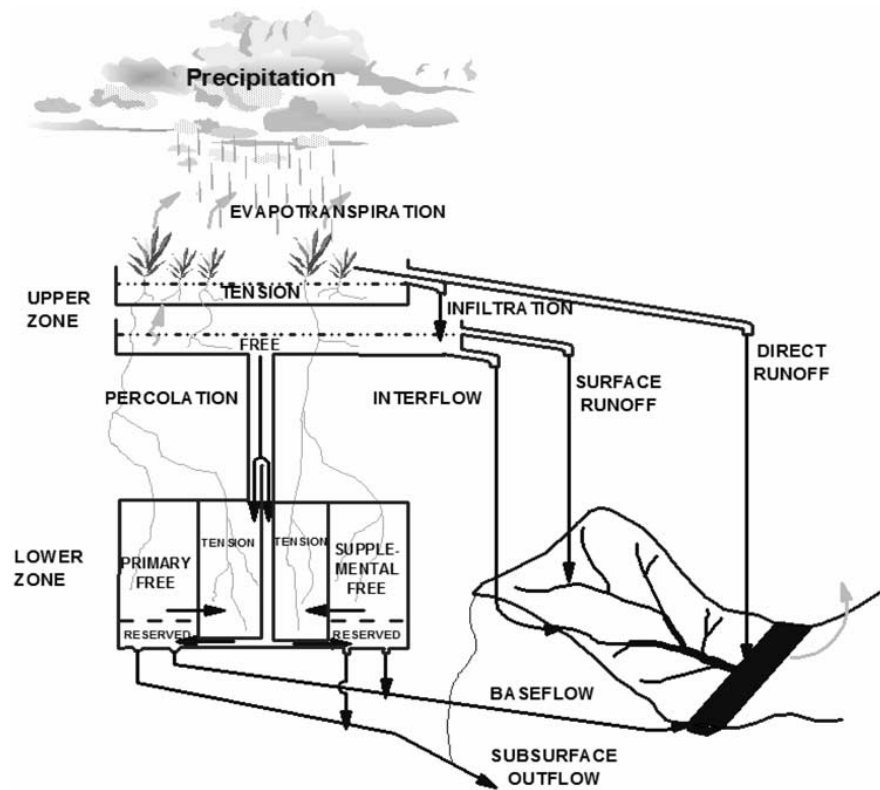


Figure 1. Schematic of the Sacramento soil moisture accounting (SAC-SMA) model [Brazil, 1988].

[11] Two simple conceptual rainfall-runoff models were also used in this study: HYMOD and SWB models. The HYMOD [Boyle, 2001] consists of a simple rainfall excess model, which is connected to two series of linear reservoirs to route surface and subsurface flow (three quick-flow reservoirs and a single slow-flow reservoir). This model includes five parameters: C_{\max} (L) is the maximum storage capacity in the catchment; $b_{\exp}(\cdot)$ is the shape factor of the main soil-water storage tank that represents the degree of spatial variability of the soil-moisture capacity within the catchment; $\alpha(\cdot)$ is the factor distributing flow between

two series of reservoirs; and $R_q(T)$ and $R_s(T)$ are the residence times of linear quick- and slow-flow reservoirs, respectively. The schematic of this model is illustrated in Figure 2. The parameters and their initial uncertainty bounds are presented in Table 2.

[12] The simple water balance (SWB) model [Schaafe *et al.*, 1996] is a conceptual, parametric water balance model which is being used as an operational model in the Nile River forecast center. This model includes two soil layers. A thin upper layer represents the vegetation canopy and the soil surface, while a lower layer represents the vegetation

Table 1. Parameters of the Modified SAC-SMA Model

Parameter	Description	Prior Range
UZWWM	upper zone tension-water capacity, mm	1.00–150.0
UZFWM	upper zone free-water capacity, mm	1.00–150.0
UZK	upper zone recession coefficient, day ⁻¹	0.10–0.5
LZTM	total lower zone water capacity, mm	1.00–1000.0
LZSK	lower zone supplementary recession coefficient, day ⁻¹	0.01–0.25
Percolation and other parameters (not optimized)		
ADIMP	additional impervious area	0.15
PCTIM	impervious fraction of the watershed area	0.025
ZPERC	minimum percolation rate coefficient	200.0
PFREE	percentage percolating from upper to lower zone free water storage	0.1
REXP	exponent of the percolation equation	3.3
RIVA	riparian vegetation area	0.01
SIDE	ratio of deep recharge to channel base flow	0.0
RSERV	fraction of lower-zone free water not transferable to tension water	0.3

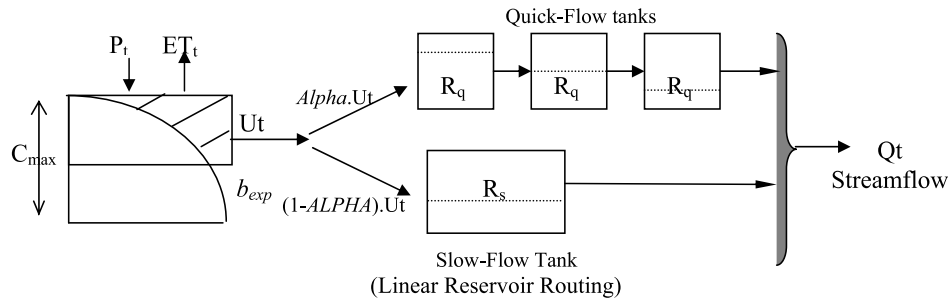


Figure 2. Schematic of the **hydrologic model (HYMOD)** [Wagener et al., 2001].

root zone and groundwater system. Five parameters controlling the SWB model processes are $D_{b,max}$, the maximum soil-moisture deficit of the bottom layer of the soil; Q_{max} , the potential subsurface runoff; $\beta = Q_{max}/S_{max}$, the ratio of the lower level posture that produces subsurface flow (S_{max} is the minimum threshold that guarantees subsurface flow); $\alpha = D_{u,max}/D_{b,max}$, the upper layer deficit proportion ($D_{u,max}$ is the maximum soil-moisture deficit of the upper layer); and K_{dt} , the timescale factor that controls infiltration into the bottom layer and the surface runoff amount. The schematic of this model is illustrated in Figure 3. The SWB model parameters and their initial uncertainty bounds are listed in Table 3.

3. Traditional Uncertainty Assessment in Hydrological Modeling

3.1. Derivation of Likelihood Function for Assessment of Parameter Uncertainty

[13] A typical hydrologic model, M , can be represented as follows:

$$y = M(\tilde{X}, \theta), \quad (1)$$

where y represents the response matrix of the catchment (e.g., streamflow), $M(\cdot)$ denotes the nonlinear hydrologic model, θ is a set of model parameters, and \tilde{X} stands for the observed forcing input matrix (e.g., precipitation). In the traditional approach, the uncertainty in the catchment response is attributed to parameter estimation uncertainty, while input and model structural uncertainty is not addressed explicitly. Assuming that the residuals are additive,

$$\tilde{y} = M(\tilde{X}, \theta) + e(\theta). \quad (2)$$

[14] The Bayesian statistics treats hydrologic model parameter, θ , as probabilistic variables, with the joint posterior probability distribution $P(\theta | \tilde{X}, \tilde{y})$, which presents the probabilistic characteristic of the θ conditioned on the observed data, \tilde{X} and \tilde{y} . Under Bayes statistics, $P(\theta | \tilde{X}, \tilde{y})$ is proportional to the product of likelihood function and the prior distribution function, $P(\theta)$. The prior probability density function explains the information about the θ , before any data are collected. Here we use a noninformative (uniform) prior over the feasible parameter space (which consists of realistic upper and lower bound for each of the parameters), $\theta \in \Theta \subset \mathbb{R}^n$.

[15] Assuming that the residuals are additive, independent (uncorrelated), and normally distributed noise with mean equal to zero and constant unknown variance, σ_y , Box and Tiao [1973] described the likelihood of parameter set describing the observed data over the number of time steps (T) can be estimated as follows:

$$L(\theta, \sigma_y | \tilde{X}, \tilde{y}) = \frac{1}{\sigma_y^T} \exp \left(-\frac{1}{2\sigma_y^2} \left(\sum_{t=1}^T (e(\theta)_t)^2 \right) \right). \quad (3)$$

[16] Further assuming noninformative prior, then $P(\sigma_y) \propto \sigma_y^{-1}$, σ_y can be integrated out of the posterior density yielding the following expression [Box and Tiao, 1973]:

$$p(\theta | \tilde{X}, \tilde{y}) \propto \left[\sum_{t=1}^T (e(\theta)_t)^2 \right]^{-\left(\frac{T}{2}\right)}. \quad (4)$$

[17] In practice, it is easier to maximize the logarithm of the likelihood function. It will identify a set of plausible parameter values given the available observed data. There are several Bayesian approaches tailored for hydrologic modeling, including the Generalized Likelihood Uncertainty

Table 2. Parameters of the HYMOD Model and Their Initial Uncertainty Ranges

Parameter	Description	Prior Range
C_{max}	maximum storage capacity in catchment, mm	1.0–500.0
b_{exp}	factor distributing flow between two series of reservoirs (·)	0.1–2.0
$ALPHA$	shape factor for the main soil water storage tank (·)	0.1–0.990
R_s	residence time of linear slow flow reservoirs, days	0.0–0.1
R_q	residence time of linear quick flow reservoirs, days	0.1–0.99

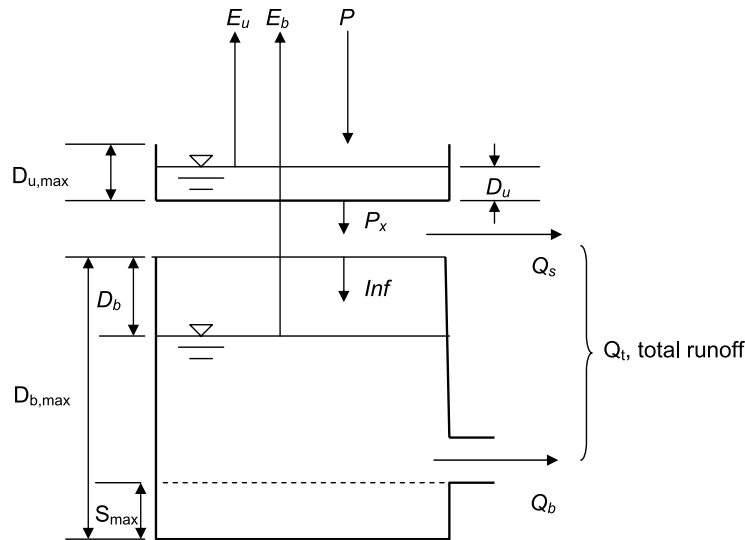


Figure 3. Schematic of the simple water balance (SWB) model.

Estimation (GLUE) framework [Beven and Binley, 1992] and the Shuffled Complex Evolution Metropolis (SCEM-UA) algorithm [Vrugt *et al.*, 2003] that consider model parameters in equation (1) as probabilistic variables and estimate their uncertainty bound based on the posterior pdf. In this study, we will further explore the SCEM-UA algorithm for estimating model parameters and their associated uncertainty bounds.

3.2. The Shuffled Complex Evolution Metropolis

[18] The Shuffled Complex Evolution Metropolis (SCEM) was built upon the principles of the effective and efficient global optimization technique, the Shuffled Complex Evolution (SCE-UA) developed by Duan *et al.* [1992]. Vrugt *et al.* [2003] combined the strengths of the Monte Carlo Markov Chain (MCMC) sampler with the concept of complex shuffling from SCE-UA to form an algorithm that not only provides the most probable parameter set, but also estimates the uncertainty associated with estimated parameters. The main difference between SCEM and SCE is that the downhill simplex method in SCE was replaced by the Metropolis-Hastings search algorithm [Metropolis *et al.*, 1953; Hastings, 1970]. Thus SCEM in every model run is able to simultaneously identify both the most likely parameter set and its associated posterior probability distribution. SCEM-UA is explained in detail by Vrugt *et al.* [2003]. The convergence of the algorithm was monitored using the Gelman-Rubin criterion [Gelman and Rubin, 1992], which

is a scale reduction score that quantitatively diagnoses if each parameter converges to a stationary distribution.

3.3. Case Study: Use of SCEM for Calibration and Uncertainty Assessment of Hydrologic Model Parameters

[19] In this section we demonstrate the performance and applicability of SCEM-UA to identify and estimate model parameters and their associated uncertainty bounds, by application to three above mentioned hydrologic models: SAC-SMA [Burnash *et al.*, 1973], HYMOD [Boyle, 2001], and SWB [Schaake *et al.*, 1996].

[20] Input-forcing data and model structures were assumed perfect in this section, and all of the uncertainty in the streamflow simulation was attributed to parameter estimation uncertainty. Uniform prior distributions were assumed on the parameter ranges of all three models. The marginal posterior probability distribution for the estimated SAC-SMA model parameters is given in Figure 4. As we mentioned earlier, the number of unknown parameters in this model was reduced to five major parameters. The distributions are generated using 20,000 samples after the algorithm converged to the final posterior distribution. Figure 4 illustrates two points. The first point is that the posterior distributions for three of the five parameters (UZTWM, LZTM, and LZSK) are approximately normal; however, the posterior distribution of UZFWM depicts the existence of two modes (multimodality). The posterior

Table 3. Parameters of the SWB Model and Their Initial Uncertainty Ranges

Parameters	Description	Prior Ranges
$D_{b,max}$	maximum soil moisture deficit of bottom layer of the soil, mm	10.0–800.0
Q_{max}	potential subsurface runoff, mm/d	5.0–100.0
β	ratio of the lower level posture that produces subsurface flow (-)	0.1–0.90
α	upper layer deficit proportion (-)	0.01–0.5
K_{dt}	timescale factor, days	1.0–20.0

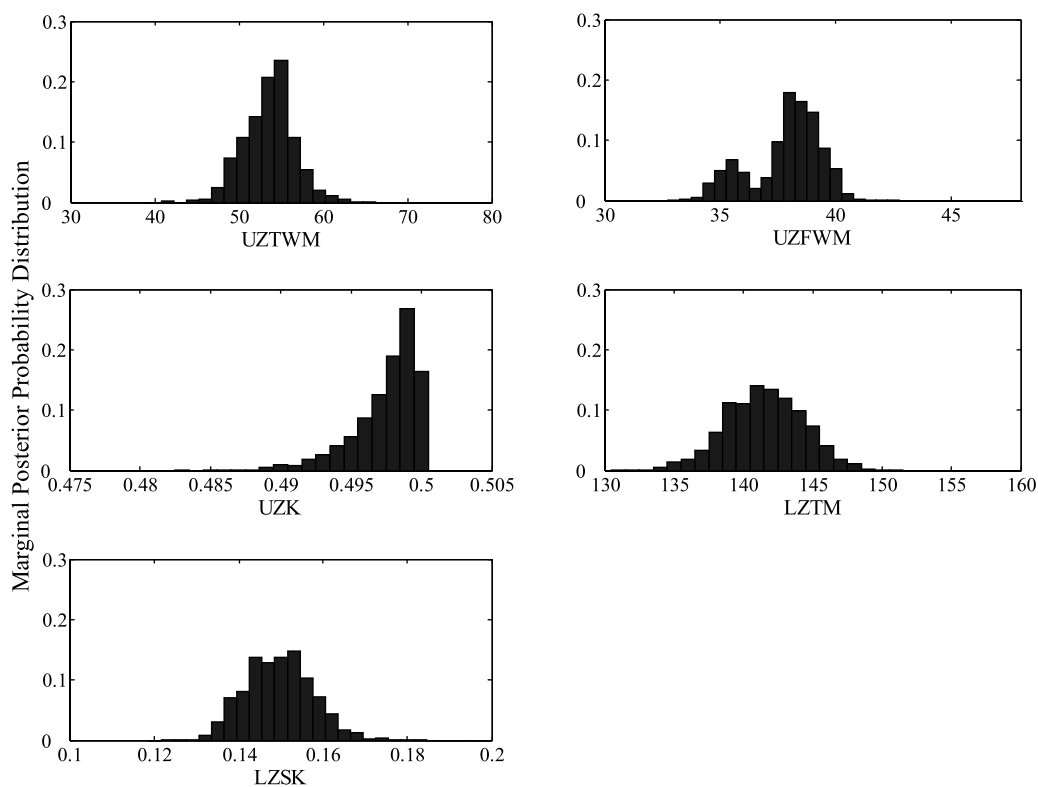


Figure 4. Marginal posterior probability distribution of the SAC-SMA parameters, using 20,000 samples generated after convergence of the SCEM-UA algorithm.

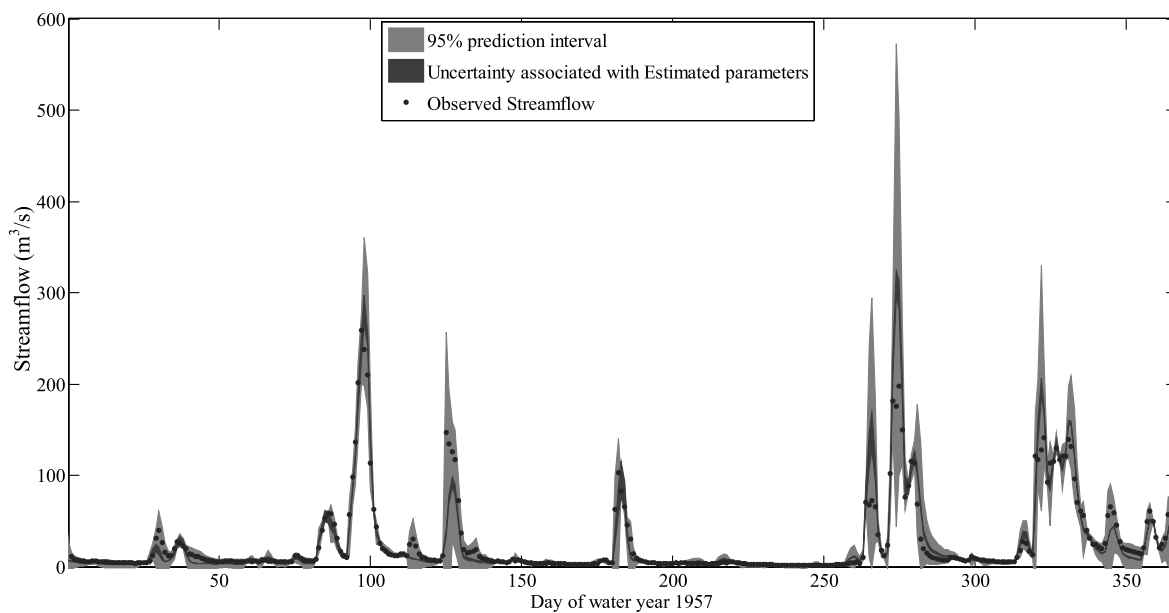


Figure 5. Streamflow hydrograph prediction uncertainty associated with estimated parameters (shown in darker gray) for the SAC-SMA model and 95% confidence interval for prediction of observed streamflow (shown in lighter gray) for water year 1957.

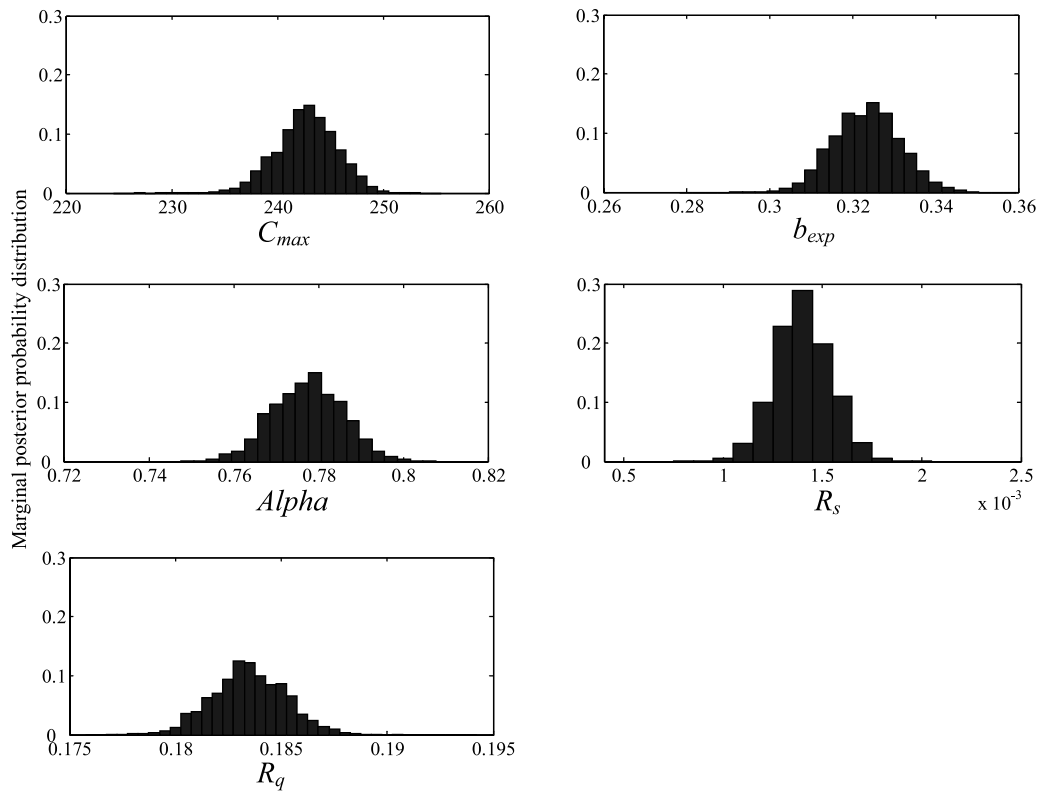


Figure 6. Marginal posterior probability distribution of the HYMOD parameters, using 20,000 samples generated after convergence of the SCEM-UA algorithm.

distribution of UZK is very close to the upper boundary of the National Weather Service predefined probable parameter range. This can be an indication of an inherent model structural uncertainty and/or other sources of uncertainty within the system which are not being considered here. The second observation is that the final converged samples for all the parameters capture only a small space of the predefined range for the parameters (Table 1). However, the hydrograph uncertainty bounds (Figure 5) associated with these parameter ranges do not cover the expected number of observed streamflow values (dark gray region in Figure 5). This can be argued as a problem of over-conditioning the selected relationships between observed and simulated (modeled) output. The light gray region in Figure 5 shows the 95% hydrograph prediction uncertainty associated with the total error in the hydrologic system in terms of model residuals (calculated based on predictive variance of SCEM). Even though the 95% total prediction uncertainty range captures all the observations, it is very wide compared with uncertainty bounds associated with parameter uncertainty, revealing a considerable amount of uncertainty in both the structure of the model under study and the data used to condition the model.

[21] To further demonstrate the applicability of SCEM, we used this algorithm to estimate optimal parameter sets and assess their associated uncertainty boundaries for two other hydrologic models, HYMOD [Boyle, 2001] and SWB [Schaake *et al.*, 1996].

[22] The final estimated marginal posterior distributions of the HYMOD model parameters, after 20,000 samples, are given in Figure 6. The results reveal that the distributions for all HYMOD parameters are approximately normal. These parameter distributions cover a very small range of predefined parameter ranges. However, in Figure 7 we can see that even though the algorithm shows high probability for these parameter sets, the estimated hydrograph prediction uncertainty bounds (dark gray) does not include many of the observed streamflow values. Similar results are presented in Figures 8 and 9 for the SWB model.

[23] The examples presented above reveal that attributing all uncertainties in hydrologic models to model parameters and ignoring input and model structural uncertainties leads to an inaccurate, biased, and inconsistent simulation of the system processes and their associated uncertainty bounds.

4. Extended SCEM-UA to Include the Input Error Model: Simultaneous Parameter and Input Uncertainty Estimation

[24] Results from the previous section indicate that dealing only with model parameter uncertainty is not enough to accurately estimate the true uncertainty in hydrologic simulation. Uncertainties from other sources must be dealt with more directly. There have been a few studies in hydrological modeling that explicitly account for input uncertainty within the system through input error models. One such approach is the Bayesian total error analysis (BATEA) by Kavetski *et*

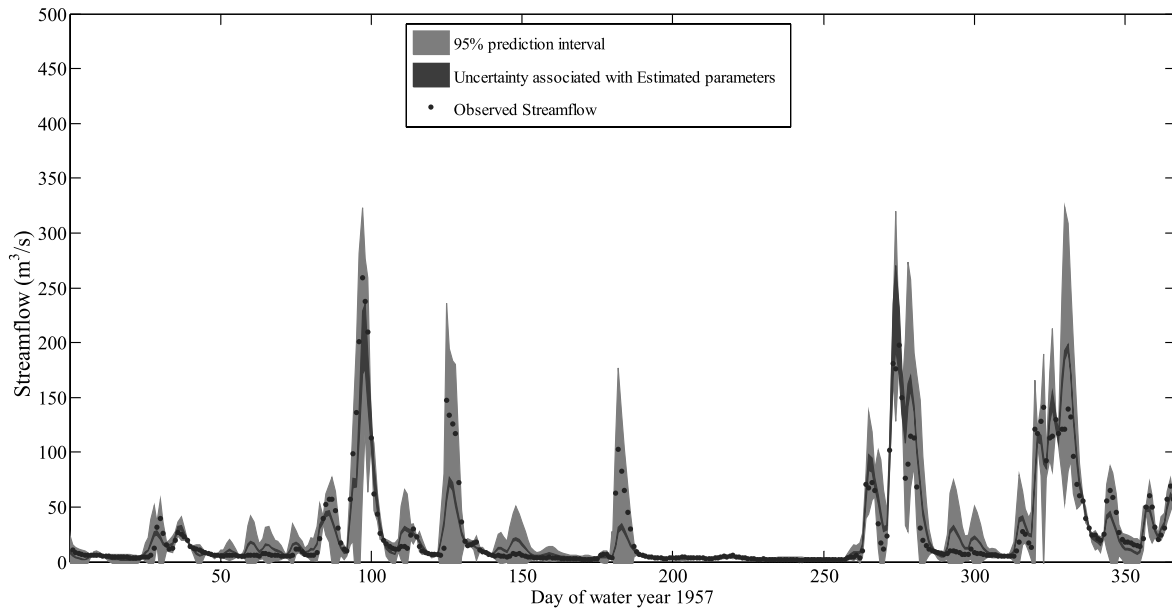


Figure 7. Streamflow hydrograph prediction uncertainty associated with estimated parameters (shown in darker gray) for the HYMOD model and 95% confidence interval for prediction of observed streamflow (shown in lighter gray) for water year 1957 (calibration period).

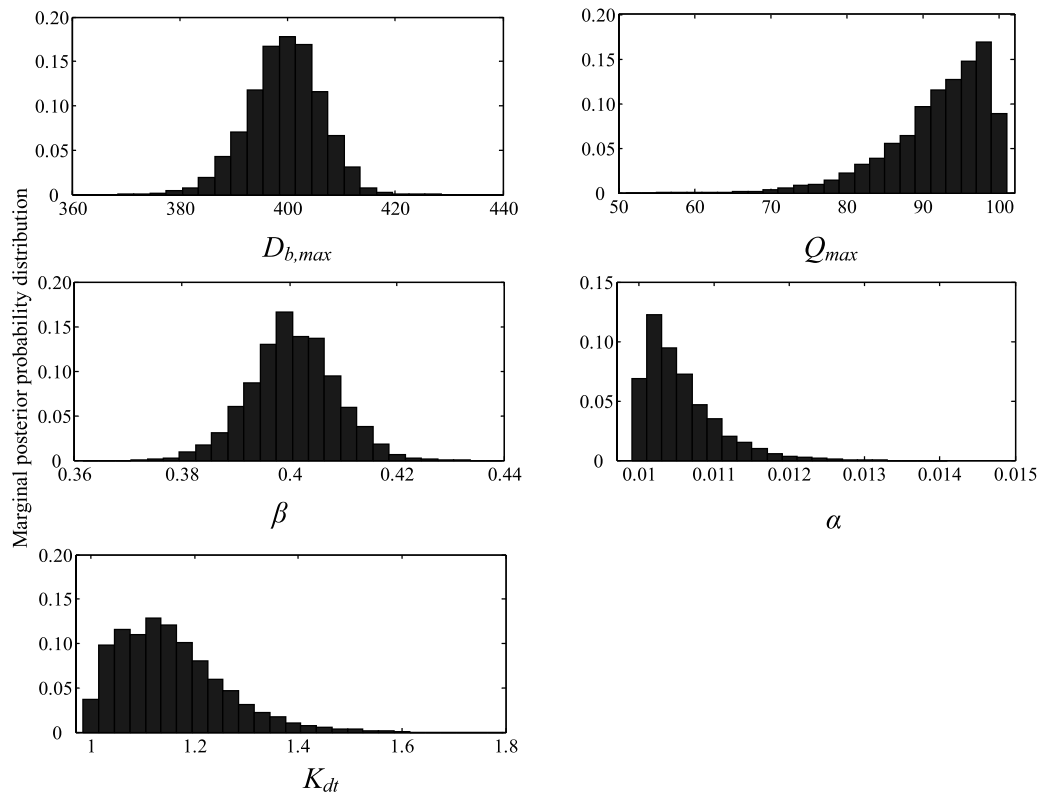


Figure 8. Marginal posterior probability distribution of the SWB parameters, using 20,000 samples generated after convergence of the SCEM-UA algorithm.

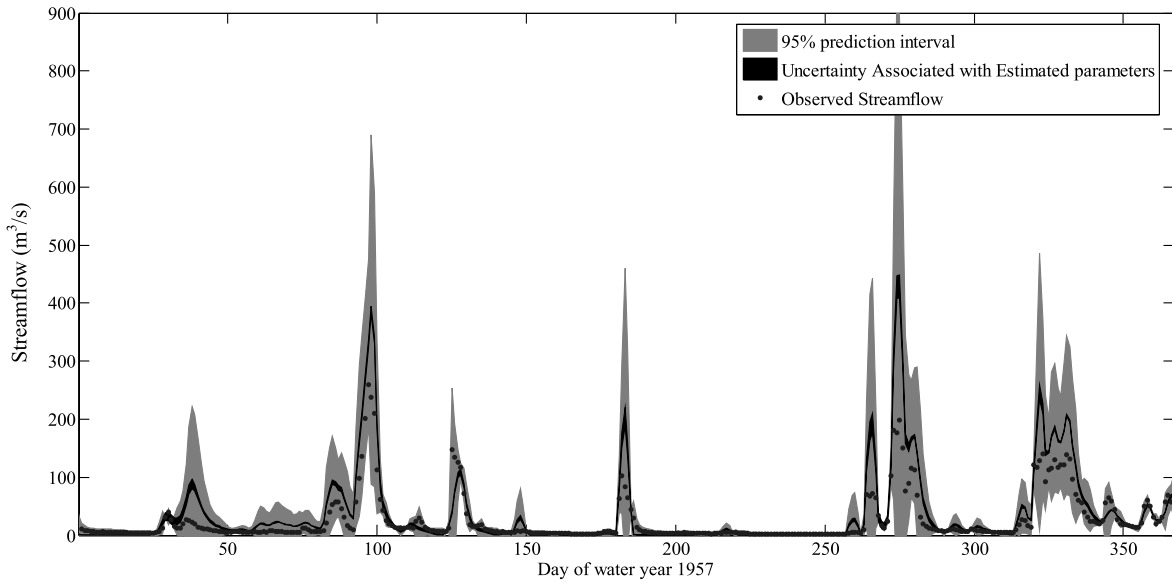


Figure 9. Streamflow hydrograph prediction uncertainty associated with estimated parameters (shown in darker gray) for the SWB model and 95% confidence interval for prediction of observed streamflow (shown in lighter gray) for water year 1957 (calibration period).

al. [2003]. BATEA is one of the few techniques which explicitly considers input error in the development of the likelihood function in hydrological modeling. The rainfall events are predefined and each is given a unique multiplying constant. These multipliers allow the pattern of rainfall as well as the event magnitude to change. *Kavetski et al.* [2003] introduced rainfall depth multipliers as some “latent variables” to the system and introduced an explicit term to the likelihood function to estimate these variables. If \tilde{r}_t represents the true rainfall depth $\hat{X} = [\tilde{r}_1, \tilde{r}_2, \dots, \tilde{r}_t, t = 1:T]$, and r_t is the observed rainfall depth, their input error model has the following form:

$$r_j = m_j \tilde{r}_j; \quad m \sim N(1, \sigma_m^2), \quad (5)$$

where j indicates the storms within the rainfall series and m_j is the random noise from a normal distribution with mean equal to one and known (prespecified) variance σ_m^2 in the form of a multiplier that corrupts the true rainfall depth and yields the observed rainfall depth. *Kavetski et al.* [2003] assumed the rainfall multipliers, m_t , as latent variables and estimated both them and the model parameters through their probabilistic calibration procedure called BATEA. By considering the multipliers just for the predefined rainfall events, they decreased dimensions of the system. Considering Bayes’ law, and assuming that (1) \hat{X} (observed input) and \hat{y} (observed catchment response) are statistically independent because catchment response \hat{y} depends only on the true input forcing \hat{X} , not necessarily on observed forcing, and (2) \hat{X} is statistically independent of θ (model parameter set), because observed input is uncorrelated to the hydrologic model parameters, *Kavetski et al.* [2003] derived the final form of their likelihood function as follows:

$$p(\theta, \hat{X} | \hat{y}, \tilde{y}) \propto L(\tilde{y} | \theta, \hat{X}) \times L(\hat{X} | \hat{X}) \times p(\theta, \hat{X}), \quad (6)$$

where $L(\tilde{y} | \theta, \hat{X})$ is the likelihood of observing \tilde{y} given a parameter set θ , and the true input forcing \hat{X} . $L(\hat{X} | \hat{X})$ is the likelihood based on input error model, and $p(\theta, \hat{X})$ represents the prior distribution of parameters and true input forcing.

[25] *Kavetski et al.* [2003] applied their BATEA framework to a series of synthetic case studies and demonstrated that considering an input error model explicitly and adding a new term to the likelihood function can improve the response surface and assessment of uncertainty bounds. Nonetheless, even though equation (6) allows the use of explicit input error models, it has two drawbacks. First, it is impossible to know what the true input forcing is in a real-world problem, and therefore it is impossible to assess the input error model likelihood, $L(\hat{X} | \hat{X})$. Second, in some cases the number of these “latent variables” can increase considerably and cause some dimensionality issues. To circumvent these two problems, in this study the input error model was changed as follows:

[26] 1. Instead of introducing latent variables to the system, we considered a multiplier in the following form:

$$\tilde{r}_t = \phi_t r_t; \quad \phi \sim N(m, \sigma_m^2), \quad (7)$$

where ϕ_t represents a random multiplier at time step t with mean equal to m , $m \in [0.9, 1.1]$ and variance equal to σ_m^2 , $\sigma_m^2 \in [1e-5, 1e-3]$. In this implementation we assume true rainfall depth \tilde{r}_t is corrupted at all times by random multipliers from the identical distribution with unknown mean, m , and variance, σ_m^2 . Thus, instead of searching for every single multiplier as a latent variable, we introduce two new parameters to the system including mean and variance of error model multiplier (instead of additive) distribution, $\eta = \{m, \sigma_m^2\}$. Considering the error term in the form of the multiplier helps to maintain the heteroscedastic (nonhomo-

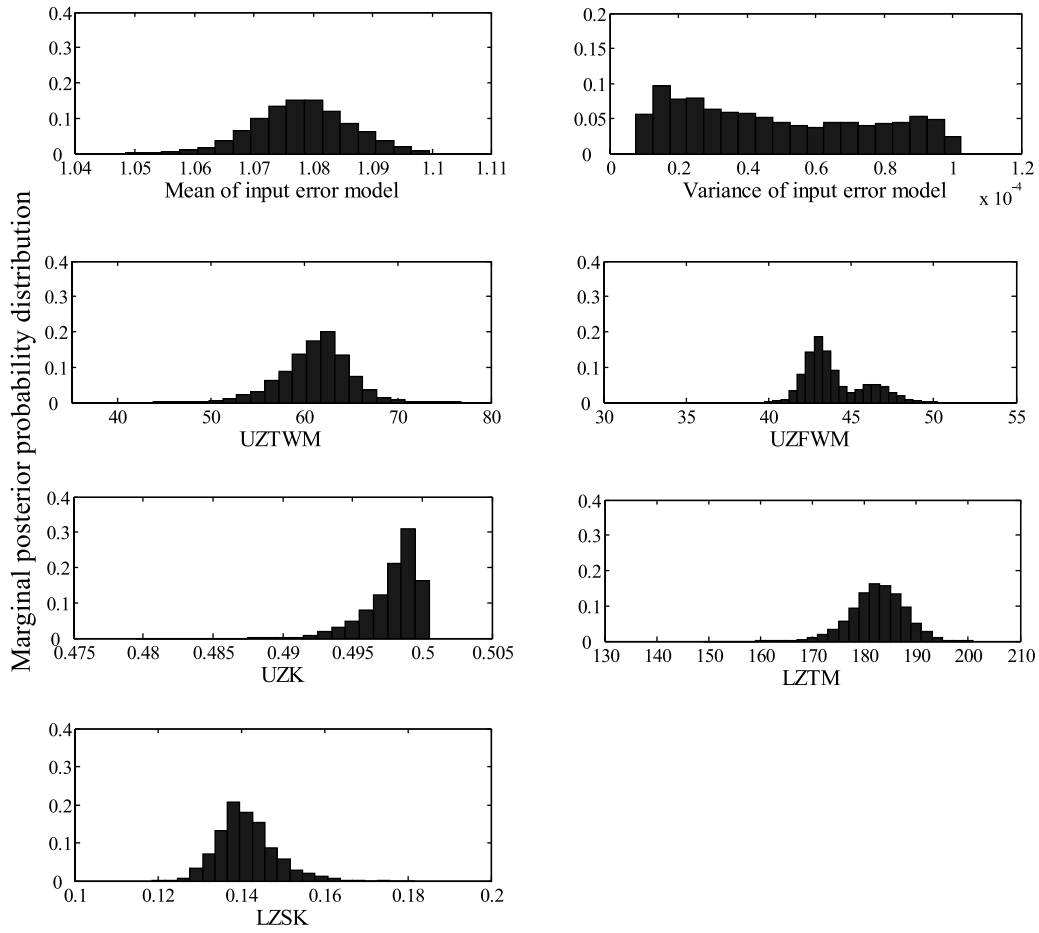


Figure 10. Marginal posterior probability distribution of the input error model parameters and the SAC-SMA model parameters using 20,000 samples generated after convergence of the SCEM-UA algorithm.

geneous) nature of the error (higher deviation in higher rainfall depths) [Sorooshian and Dracup, 1980].

[27] 2. To deal with the issue of not having true observations of input forcing data, it was decided to integrate the input error model into the model error term:

$$e(\theta) = y(\theta, \eta) - \tilde{y} \quad (8)$$

Therefore the likelihood function will have the following form:

$$p(\theta, \eta | \tilde{X}, \tilde{y}) \propto L(\tilde{y} | \theta, \eta, \tilde{X}) \times p(\theta, \eta). \quad (9)$$

[28] In brief the implemented changes into the hydrologic input-output system included introduction of a random multiplier to each time step, drawn from the same normal distribution with unknown mean and variance (m and σ_m^2). These two variables of the input error model (mean and variance of the distribution) were added as two unknown parameters to the system. The SCEM-UA was used to estimate the model parameters and input error model parameters simultaneously. Later the uncertainty associated with input error model parameters and hydrologic model

parameters were propagated through the system to estimate associated uncertainty with streamflow simulations and predictions.

4.1. Case Study: Use of Extended SCEM for Calibration and Uncertainty Assessment of Hydrologic Model Parameters and Input Error Model Parameters

[29] By means of a case study, we illustrate the performance of the SCEM-UA while considering an input error model to specify the hydrologic system. Again, we applied SCEM-UA to calibrate and assess uncertainty bounds for SAC-SMA, HYMOD, and SWB model parameters along with input error model parameters on the Leaf River basin. The idea is to compare the results from this part of the study to those from section 2.3.

[30] Figure 10 shows the new marginal posterior distribution estimated for each parameter of the SAC-SMA model while considering an input error model's first two moments as two additional parameters in the system, using SCEM-UA. Looking at Figure 10 and comparing the results with Figure 4, two observations can be made. One is that considering input error model, the final estimated marginal distribution for the model parameters moved over the possible parameter ranges and assigned the mode of the

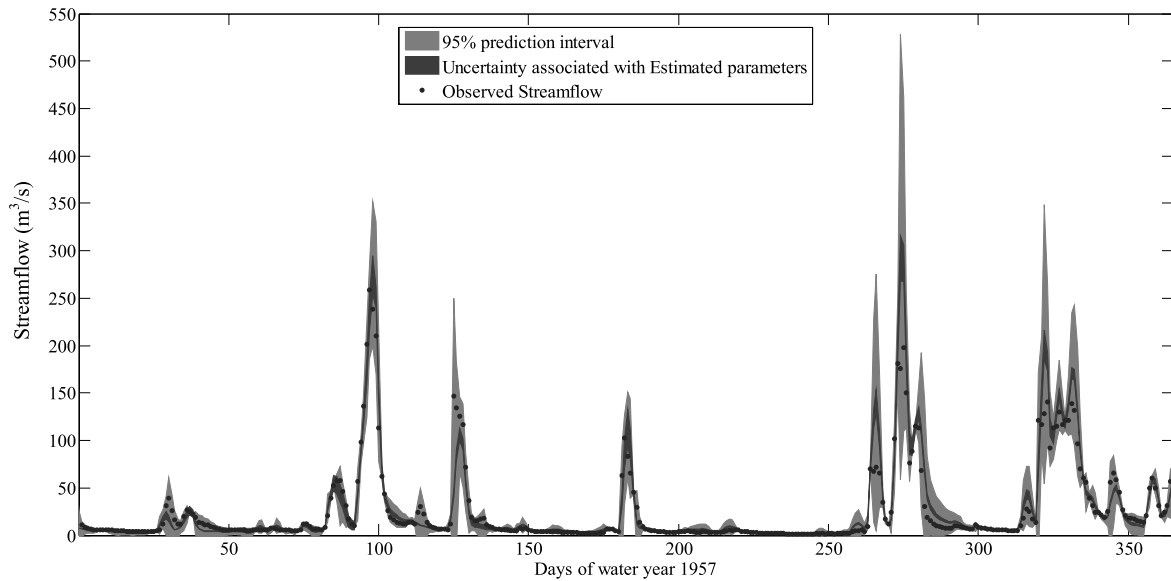


Figure 11. Streamflow hydrograph prediction uncertainty associated with estimated parameters and input error model parameters (shown in darker gray) for the SAC-SMA model and 95% confidence interval for prediction of observed streamflow (shown in lighter gray) for water year 1957 (calibration period).

probability distribution to different parameter values. The second observation is that the mean of the input error model has a mode different than one. If the input forcing was correct, the mean of the input error model would concentrate around one, and the final marginal distribution of the parameters would be the same as if we did not account for input uncertainty. This can also be the indication that the input error model is somehow compensating for the existing model structural deficiencies. The estimated uncertainty bounds for the hydrograph associated with input and model parameter uncertainty are shown in Figure 11. The 95% prediction intervals are narrower here compared with the original case (only considering uncertainty in model parameters). This reveals that the final uncertainty bounds associated with both input and model parameters are more accurate and that variance of the residuals at each point is smaller compared with the original scenario.

[31] These above mentioned results for the SAC-SMA model are confirmed in Table 4. The observation coverage for the estimated 95% uncertainty bounds for the simulation has increased by almost 70% when we account for input uncertainty. The same results are presented in the table for the HYMOD and SWB models, which reveals that accounting for input uncertainty improved the final streamflow simulation of these models as well. These results illustrate that not accounting for input uncertainty can lead to biased parameter estimates, which are compensating for other sources of uncertainty. Accounting for input uncertainty improves the daily root-mean-square (DRMS) error for all three models across all of their ensembles, as seen in Figure 12. We can also see that this improvement is more significant for the SWB and HYMOD models and less significant for the SAC-SMA model.

[32] One of the important observations from the set of experiments presented in this section was that the estimated mean and variance of input error model and their associated

uncertainty bound are different from one hydrological model to the other one. This is an inevitable result since we are still ignoring model structural uncertainty. Therefore all the model parameters as well as input error model parameter are still compensating for model structural uncertainty. The next section focuses on this important source of uncertainty in hydrologic system simulation.

5. Uncertainty Assessment in Hydrological Modeling: Simultaneous Parameter and Input and Model Structural Uncertainty Estimation

5.1. Classical Model Structural Error

[33] The dominant approach in hydrological modeling and streamflow forecasting has been the use of a single model. However, dependence on a single hydrological model, which presumably does not adequately represent all of the physical processes of the watershed well, results in unreliable, uncertain, and overconfident forecasts. This is the case even if we account for all other sources of uncertainty such as parameter estimation and input forcing uncertainty [Georgakakos *et al.*, 2004]. To date, all of the approaches set forth to identify model structural inadequacy focused on a single-model structure and how it can be improved to more adequately represent the system [e.g., *Vrugt et al.*, 2005].

[34] A new kind of approach that has recently emerged to identify model structural uncertainty is to use multimodel combination techniques, which provide a better understand-

Table 4. Percentage of Observations Being in 95% Uncertainty Bounds.

	SAC-SMA HYMOD SWB		
SCEM (hydrologic model parameters)	15%	10%	5.6%
SCEM (hydrologic + input model parameters)	25%	17%	10%

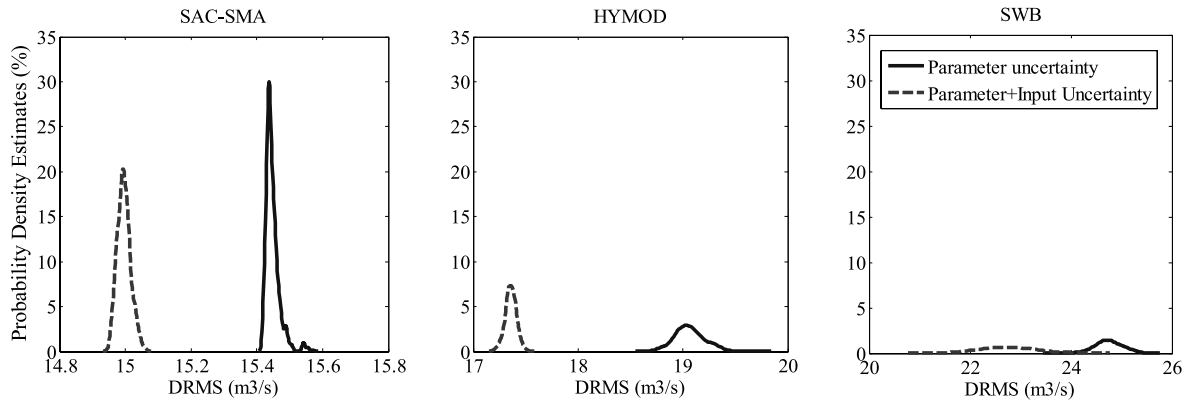


Figure 12. Distribution of daily root-mean square (DRMS) of SAC-SMA, HYMOD, and SWB considering only parameter uncertainty compared to parameter and input uncertainty.

ing of the watershed processes by investigating multiple model structures.

5.2. Bayesian Model Averaging

[35] Bayesian model averaging is a probabilistic scheme for model combination. It is a coherent technique for accounting for model structural uncertainty [Madigan *et al.*, 1996]. Below is a brief description of the essence of the BMA scheme. Consider a quantity \tilde{y} as the observed output variable to be forecasted and $M = [M_1, M_2, \dots, M_K]$ the set of all considered models. The $p_k(y_k | M_k, \tilde{X}, \tilde{y})$ is the posterior distribution of y_k which represents the quantity to be forecasted under model M_k , given a discrete data set, \tilde{X} (input forcing data) and \tilde{y} (observed system processes, here streamflow). The posterior distribution of the BMA prediction, y_{bma} , is thus given as

$$p(y_{bma} | M_1, \dots, M_K, \tilde{X}, \tilde{y}) = \sum_{k=1}^K p(M_k | \tilde{X}, \tilde{y}) \cdot p_k(y_k | M_k, \tilde{X}, \tilde{y}), \quad (10)$$

where $p(M_k | \tilde{X}, \tilde{y})$ is the posterior probability of model M_k . This term is also known as the likelihood of model M_k being the correct model. If we denote $w_k = p(M_k | \tilde{X}, \tilde{y})$, we should obtain $\sum_{k=1}^K w_k = 1$. The $p_k(y_k | M_k, \tilde{X}, \tilde{y})$ is represented by the normal distribution with mean equal to the output of model M_k and standard deviation σ_k . Suppose that y_k is a prediction made by model M_k . Weights can be estimated through the expectation-maximization algorithm [Dempster *et al.*, 1977] which will be discussed in the next section. The posterior mean and variance of the BMA prediction for variable y_{bma} are

$$E[y_{bma} | y_1, \dots, y_K, \tilde{X}, \tilde{y}] = \sum_{k=1}^K w_k y_k \quad (11)$$

$$Var[y_{bma} | y_1, \dots, y_K, \tilde{X}, \tilde{y}] = \sum_{k=1}^K w_k \left(y_k - \sum_{i=1}^K w_i y_i \right)^2 + \sigma^2, \quad (12)$$

where σ^2 is the variance of the time series shaped based on one of the model predictions (ensembles) being the best at each time step. Suppose if we build a time series that at each

time step includes the best prediction (closest to the observation) from one of the K models; σ^2 represents the variance of such time series considering observations.

[36] In essence, the BMA prediction is the average of predictions weighted by the likelihood that an individual model is correct. There are several attractive properties to the BMA predictions. First, the BMA prediction receives higher weights from better performing models, as the likelihood of a model is essentially a measure of the agreement between the model predictions and the observations. Second, the BMA variance is a measure of the uncertainty of the BMA prediction. This measure is a better description of predictive uncertainty than that in a non-BMA scheme, which estimates uncertainty based only on the model ensemble spread (i.e., only the between-model variance is considered), and consequently results in under-dispersive predictions [Raftery *et al.*, 2003, 2005].

5.3. Combination of Global Optimization and Bayesian Multimodel Combination: An Integrated Bayesian Uncertainty Estimator

[37] Because the Bayesian multimodel combination framework offers an excellent statistical approach to account for model structural uncertainty, the BMA framework was combined with the SCEM-UA to form a hybrid framework to exploit the strengths of these two techniques for integrated schemes for quantification of input, parameter estimation, and model structural uncertainty. This framework should provide a more precise measure of uncertainty in system simulations. Throughout the remainder of this paper we will refer to this integrated Bayesian uncertainty estimator framework as IBUNE.

[38] IBUNE first estimates the two terms in the right-hand side of equation (10), $p_k(y_k | M_k, \tilde{X}, \tilde{y})$ and $p(M_k | \tilde{X}, \tilde{y})$, for each model. The $p_k(y_k | M_k, \tilde{X}, \tilde{y})$, which represents the posterior distribution of estimated hydrologic response (e.g., streamflow), y_k , under model M_k , is directly related to the input and parameter uncertainty under model M_k , expressed as follows:

$$p(y_k | M_k, \tilde{X}, \tilde{y}) \propto p(\theta_k, \eta_k | M_k, \tilde{X}, \tilde{y}). \quad (13)$$

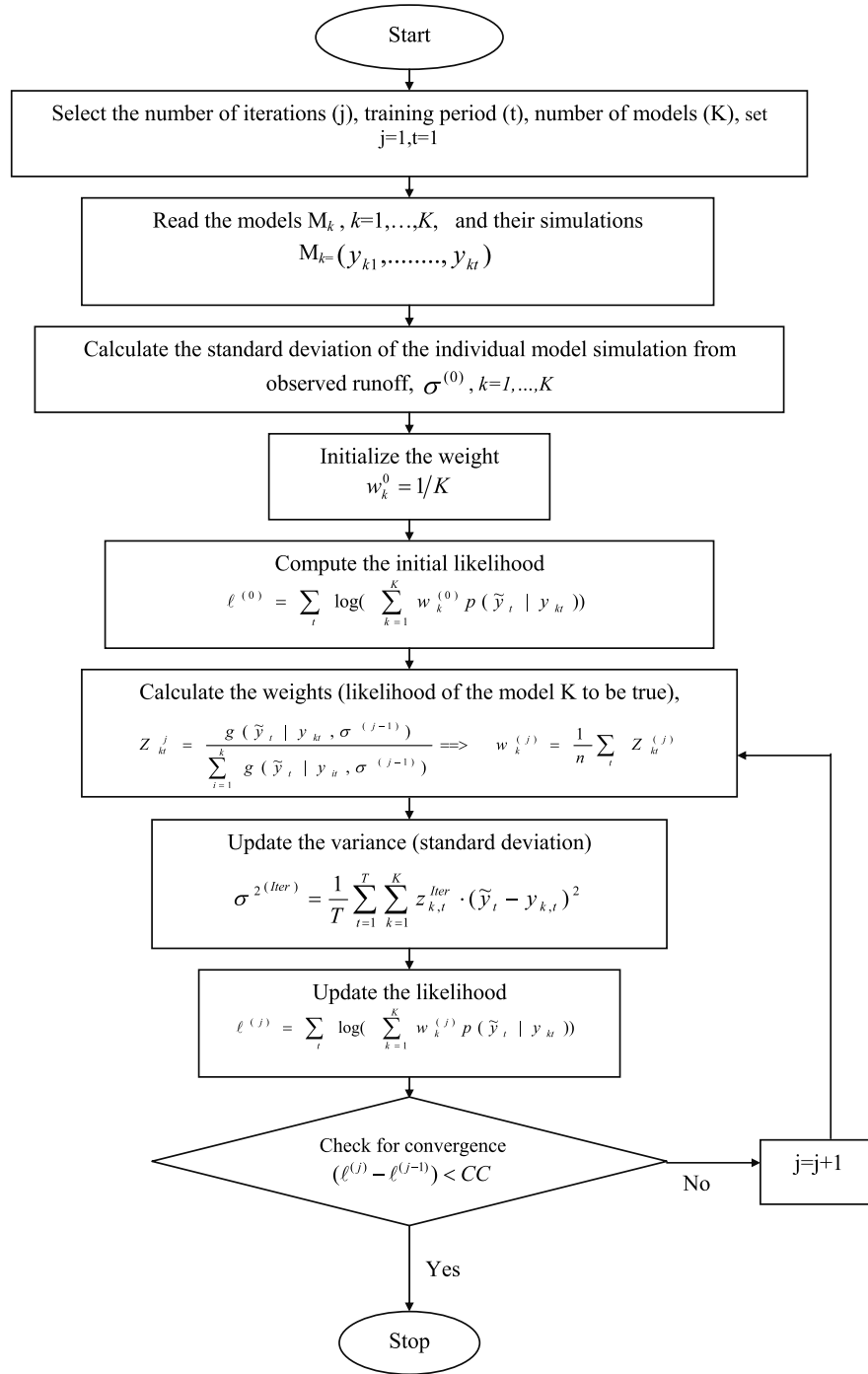


Figure 13. Expectation-maximization (EM) flowchart.

[39] We can substitute the left-hand side of equation (13), which is the outcome of SCEM-UA, into equation (10) directly. The first term of equation (10), $p(M_k|\tilde{X}, \tilde{y})$, which represents the posterior probability of the model M_k being a correct model, reflects how well model M_k matches the observed quantity of interest. To estimate $p(M_k|\tilde{X}, \tilde{y})$ or as

we mentioned earlier in the previous section w_k and σ^2 (the variance of the best time series shaped based on one of the model predictions (ensembles) being the best at each time step), we used the maximum likelihood approach. The idea is to estimate w_k and σ^2 by maximizing the likelihood of occurrence of the observed data, \tilde{y} . As we mentioned earlier,

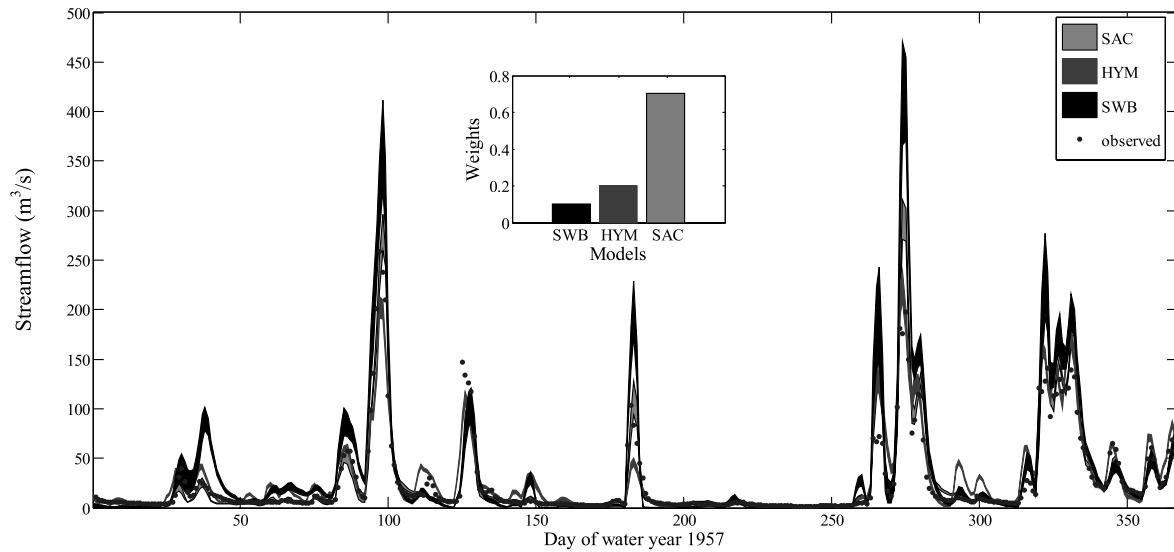


Figure 14. Streamflow hydrograph prediction uncertainty associated with estimated parameters and input error model parameters for all three models for water year 1957 and estimated combination weights for each model using Integrated Bayesian Uncertainty Estimator (IBUNE) (training/calibration period).

it is easier to maximize the logarithm of likelihood function, and therefore we define the logarithm of likelihood function as follows:

$$L(w_1, \dots, w_k, \sigma^2) = \sum_{t=1}^T \log \left(\sum_{k=1}^K w_k \cdot p(\tilde{y}_t | y_{kt}) \right). \quad (14)$$

[40] Because of the high dimensionality of this problem, it is hard and inefficient to maximize equation (14) through direct nonlinear maximization methods such as Newton-Raphson or its variants [Raftery *et al.*, 2003]. In this study, to approximate equation (14) and estimate the model weights as Raftery *et al.* [2003] suggested, we maximized the equation (14) by performing the expectation-maximization technique. All of the conditional densities were described as Gaussian distribution for computational simplicity; however, the BMA scheme can be applied by assuming other probability distributions. The EM algorithm is applied to estimate w_k and σ^2 for each model. In brief, the expectation-maximization [Dempster *et al.*, 1977] algorithm casts the maximum likelihood problem as a “missing data” problem. The missing data here are introduced as a latent variable $Z_{k,t}$ that needs to be estimated. If the k th model ensemble is the best prediction at time t , $Z_{k,t} = 1$; otherwise $Z_{k,t} = 0$. At any time t , there is only one $Z_{k,t}$ equal to 1 and the rest are equal to 0. The EM algorithm starts with an initial guess for w_k and σ^2 (the variance of the time series shaped based on one of the models being best at each time step) and then alternates between the E (or expectation) step, which estimates $Z_{k,t}$ based on the current value of w_k and σ^2 , and the M (or maximization) step, where new values for w_k and σ^2 are estimated based on the current value of $Z_{k,t}$. The EM algorithm is described in Figure 13. For more detail description of the EM algorithm, readers are referred to McLachlan and Krishnan [1997].

[41] After convergence of this algorithm, we will have specified weights for each model. Therefore equation (10) can be derived and the posterior mean and variance of the forecast can be estimated through equations (11) and (12), respectively.

[42] In brief, the IBUNE framework can be implemented as follows:

[43] 1. Select the number of hydrologic models.

[44] 2. Assign prior probability to each model (we assume noninformative prior which gives uniform weights to all of the models).

[45] 3. Define an input error model.

[46] 4. Obtain posterior distribution of model parameters and input error model parameters for each model using SCEM [Vrugt *et al.*, 2003].

[47] 5. Generate a prespecified number of streamflow ensembles for each model, using probabilistic parameter estimates obtained from steps 2–4.

[48] 6. Estimate the model weight and variance of each ensemble member using the EM algorithm [Dempster *et al.*, 1977].

[49] 7. Compute the model weights by summing the weights for all ensemble members of each model.

[50] 8. Assess predictive mean and variance using equations (11) and (12).

[51] A case study on the applicability and robustness of IBUNE for reliable assessment of predictive uncertainty propagated through the system from all the important sources of uncertainty is provided in the next section.

5.4. Use of IBUNE: Uncertainty Assessment of Hydrologic Model Parameters and Input Error Model Parameters and Model Structure

[52] The IBUNE scheme promises better assessment of total uncertainty because it accounts for model parameters, input, and model structural uncertainty. In this section we will present the results for IBUNE and compare it with all

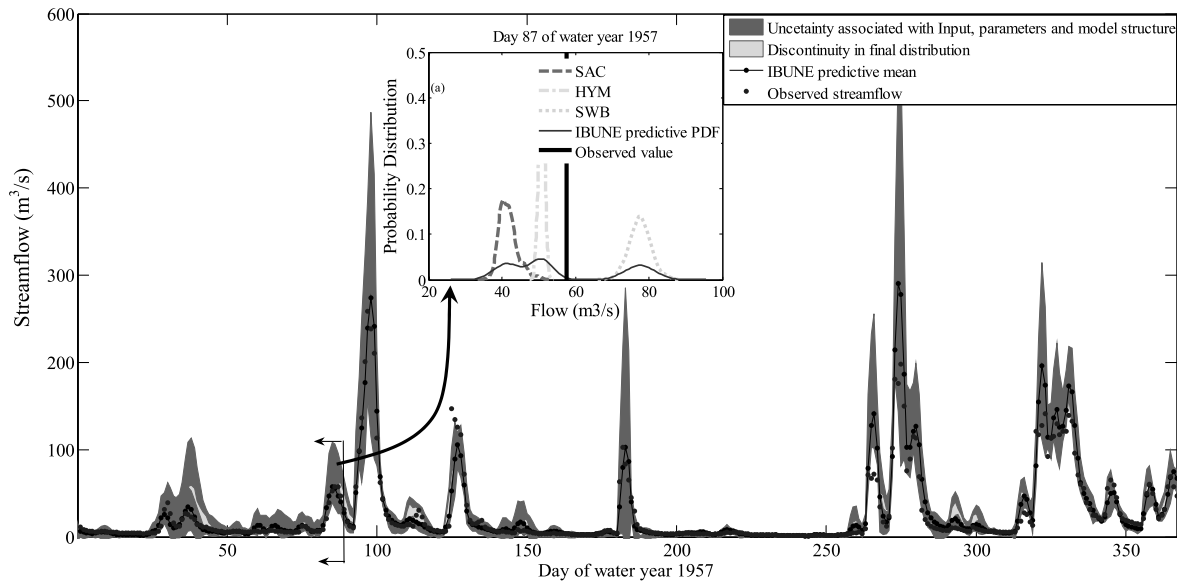


Figure 15. Streamflow hydrograph prediction uncertainty associated with estimated parameters and input error model parameters as well as model structural uncertainty (in shaded gray) for water year 1957 (calibration period). The lighter patches in the uncertainty bounds represent the discontinuity of the final model distributions. (inset) Profile of the selected cross section which includes the final distribution of each member model and the final IBUNE predictive probability density function.

other scenarios. Figure 14 illustrates the estimated uncertainty bound using SCEM, associated with input and model parameters for the three above mentioned models for the water year 1957. The solid dots in Figure 15 represent observed streamflow. Notice that different models include different observation values from various parts of the observed hydrograph, which can be interpreted as skill of the model to capture various processes within the watershed. On the basis of step 5 of IBUNE (presented in the previous section), the posterior probability distribution of each model in capturing observations (i.e., the weight) for each model was estimated. The weights are presented in Figure 14. As expected, the model with the higher skill (SAC-SMA) was assigned the highest weight, while the model with the lowest skill (SWB) was assigned the lowest weight. Both HYMOD and SWB gain very small weights. However, their contribution to the final results is considerable because they represent variety of the watershed processes which were not well represented in the SAC-SMA.

[53] The final IBUNE predictive probability which was estimated based on the probability of contributing model in the combination is given in Figure 15. The width of this final probability can be calculated through equation (12); however, the shape and intensity of the distribution can be captured through summation of the posterior probability distribution of contributing models in the combination (Figure 15a). The connected dots depict the IBUNE predictive mean which was estimated through equation (11) using the estimated weights and model simulations at each point. Another interesting observation from Figure 15 is that in some parts of the hydrograph, the final posterior probability of the three contributing model does not meet and therefore causes discontinuity in the final posterior probability distribution at these parts of the hydrograph (Figure 15). These

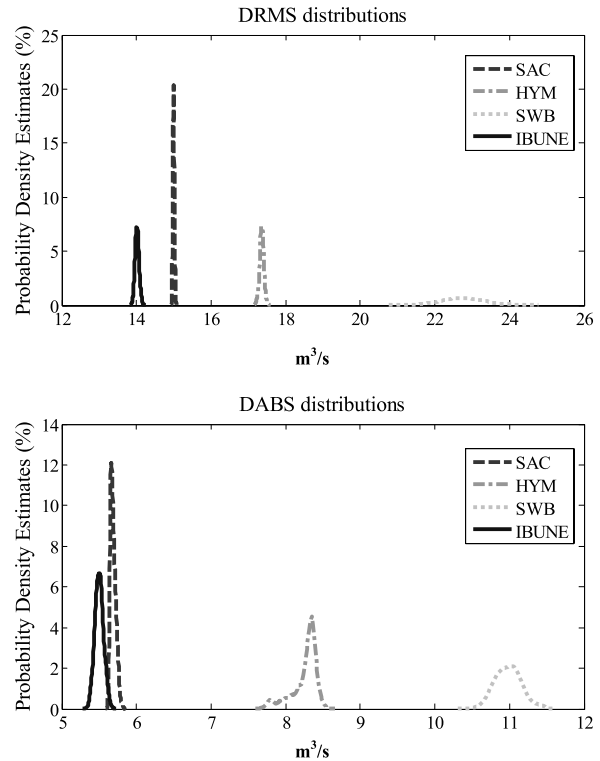


Figure 16. Distribution of DRMS and DABS for individual models and IBUNE.

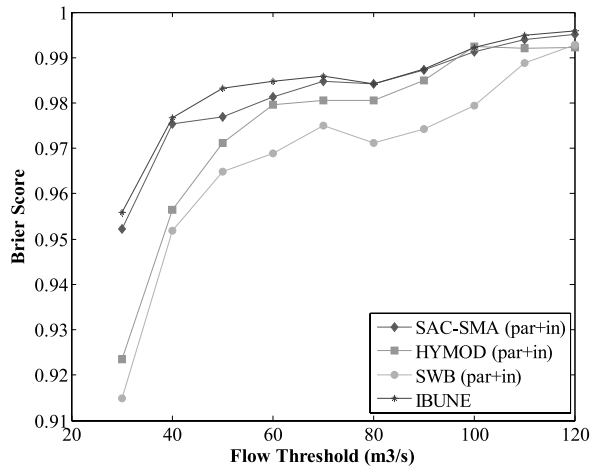


Figure 17. Brier score for IBUNE and three member models.

discontinuity areas are presented by the light gray color in Figure 15. Also shown in Figure 15a is the profile of a cross section in the hydrograph for clarification. Notice that the posterior probability distribution at this time step is discontinuous with three distinct modes. This is a clear indication that these three models do not represent the model space well and that more models are needed to avoid this problem. This also suggests that just looking at the uncertainty bounds as set of percentiles can be misleading in some cases. Figure 16 shows the distribution of daily root-mean-square error as well as the daily absolute error for all three contributing models and simulations generated through IBUNE. These distributions were estimated based on the ensemble of simulations generated by each model through their input and model parameter distributions. Figure 16 illustrates that IBUNE improved DRMS more than DABS, and these results indicate that IBUNE improved simulation of the high-flow values more than the low-flow values.

[54] The Brier score (BS) was also used to compare the skill of the individual model ensembles (considering both parameter and input uncertainty) with IBUNE. The Brier score is a scalar measure of the quality of probabilistic forecast and has been commonly used in literature. BS is defined as follows [Georgakakos *et al.*, 2004]:

$$BS = 1 - \frac{1}{N} \sum_{t=1}^N (f(t) - o(t))^2, \quad (15)$$

where $f(t)$ is frequency of target event at time step t estimated by the fraction of model ensemble simulations which are larger than prespecified threshold; $o(t)$ is equal to 1 if the observation at that time step is larger than threshold and equal to zero otherwise; and N is the number of time steps in the record. Here BS is a positively oriented score, and therefore in Figure 16 the higher the BS the better. Figure 17 shows the BS for all the models and IBUNE. The findings in Figure 16 that IBUNE produces superior predictions than individual member models are confirmed in Figure 17. One can see that IBUNE gained a higher score in most of the thresholds. Another observation from this

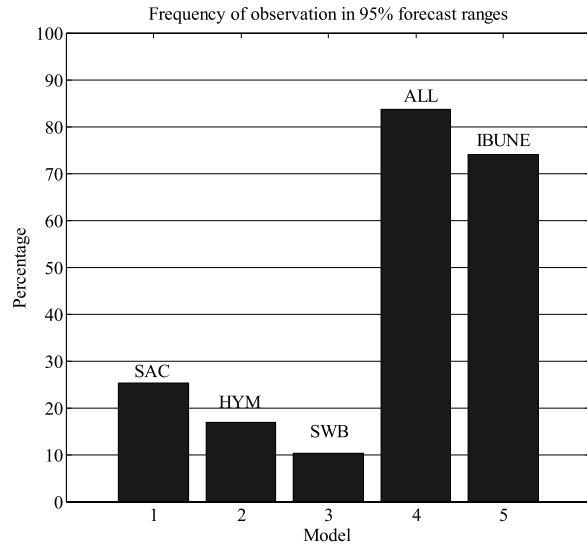


Figure 18. Percentage of observation in 95% uncertainty bounds of different models and IBUNE.

figure is that IBUNE outperformed other models over the low-flow periods as well as the high-flow periods. This suggests that IBUNE is a promising flood forecasting framework because it has higher skills in capturing higher flows.

[55] The percentage of observations which are bracketed by the estimated 95% uncertainty bounds is given in Figure 18. Ninety-five percent uncertainty bounds estimated through IBUNE cover 74% of the observation over the whole study period, which is significantly higher than any single model. This 74% excludes the points which are in the discontinuity sections with zero probability. However, only considering whatever point falls within the upper and lower uncertainty bounds at each time step will give us a percent convergence equal to 83% (Figure 18).

[56] To demonstrate usefulness of IBUNE as a streamflow prediction framework, we evaluated its performance using data from an independent 3-year validation period (1958–1960). Table 5 presents summary statistics of the validation results comparing all three scenarios, including SCEM, Extended SCEM, and IBUNE. The results in Table 5 indicate that IBUNE consistently provides better values of

Table 5. Summary of Statistics of the Streamflow Prediction During the Validation Period (Water Year 1958–1960) for All Three Scenarios Presented

Statistic	SCEM (par)			Extended SCEM (par + input)			IBUNE
	SAC	HYM	SWB	SAC	HYM	SWB	
DRMS	16.68	15.48	24.05	16.09	13.23	22.27	12.59
Percent bias	−12.38	−7.83	18.89	−0.45	−3.67	10.64	−2.6
Correlation	0.89	0.89	0.86	0.90	0.92	0.87	0.94
Percent of Observations ^a	11.5	17.5	7.6	22	20.5	14.5	76.3

^aPercentage of observations being in 95% uncertainty bounds.

the DRMS error, bias (percent bias), correlation, and percent of observations fall within 95% uncertainty bounds (percent of observations) statistics than both SCEM and Extended SCEM for all three models. It is also interesting to point out that as we account for input uncertainty along with parameter uncertainty (extended SCEM) in the hydrological models, the statistics tend to improve compared with the conventional SCEM approach, which just accounts for parameter uncertainty. Therefore explicitly accounting for input, parameter, and model structural uncertainty during calibration period can lead to improved assessment of predictive uncertainty as well as model forecasts.

6. Summary and Conclusions

[57] The prevailing approach in hydrological modeling and the assessment of related uncertainty has been the use of sophisticated calibration techniques to estimate an optimal set of parameters for a single model. Through these processes, all other sources of uncertainty, including input and model structural uncertainty, are generally ignored, and the uncertainty in the model estimation of the system is primarily assigned to the uncertainty in model parameters. Nevertheless, we know that a single-model structure is incapable of representing all of the hydrological processes within a watershed and all of the system observations including input forcing contain measurement error. Consequently, these assumptions lead to an incorrect estimation of total uncertainty in the model predictions.

[58] The objectives of this paper were threefold: (1) to demonstrate that the classic uncertainty assessment approach in hydrology which relays all of the uncertainty within the system on the parameter estimation is not reliable and accurate, (2) to introduce a new approach to simultaneously address model parameter estimation and input forcing uncertainty, and (3) to propose a new framework that tackles three major sources of uncertainty, including uncertainty inherited in input forcings, parameter estimation, and model structure. The conclusion of this work can be summarized as follows:

[59] 1. The underlying approach for uncertainty assessment in hydrological modeling has been to treat the model and observation data unbiased and precise and treat the uncertainty in the modeling processes as being explicitly attributed to the uncertainty in the parameter estimates. In this study we verified that such an assumption will lead to biased and corrupted parameter estimates. Hence the result is unrealistic model simulations and their associated uncertainty bounds which does not consistently capture and represent the real-world behavior of the watershed. This was demonstrated through two separate case studies using Shuffled Complex Evolution Metropolis (SCEM) [Vrugt *et al.*, 2003], the newly developed probabilistic parameter estimation algorithm, to calibrate three selected hydrologic models for the Leaf River basin in Mississippi. The under-study models included the Sacramento soil moisture accounting (SAC-SMA) model, the soil water balance (SWB) model, and the hydrologic model (HYMOD).

[60] 2. In the second attempt to estimate more accurate and less corrupt uncertainty bounds for the hydrologic model simulation, we proposed a new approach to account for associated uncertainty in the input forcings. We simply introduced an input error model which assumed a random

Gaussian error as a multiplier for every input observation. The common ground for these multipliers is that they are all from an identical distribution with unknown first two moments (mean and variance). Therefore we extended SCEM to estimate these two new unknown parameters with the hydrologic model parameters and their associated uncertainty. We demonstrated that undertaking such a simple approach to address input uncertainty improved the accuracy and reliability of the hydrologic simulations and their associated uncertainty bounds significantly.

[61] 3. Although accounting for the input uncertainty generated more reliable results, these results were still suffering from a very common limitation in hydrologic modeling attitudes that the model under study is the best model in hand. However, the most sophisticated models are still simple representations of the real world and cannot capture all of the processes with the catchments. In order to take into account this source of uncertainty, we exploit the newly developed technique, called Bayesian model averaging (BMA) [Hoeting *et al.*, 1999]. BMA disregards the traditional belief in hydrological modeling and explores multiple model structures to represent the processes within the system. We merged this method (BMA) with the extended SCEM presented in this paper which accounts for both input and parameter uncertainty and proposed a new hybrid framework entitled, Integrated Bayesian Uncertainty Estimator (IBUNE). IBUNE combines and exploits the strengths of the SCEM as an efficient and effective probabilistic model parameter estimator algorithm and the introduced input error model, as well as Bayesian model combination techniques, to provide an integrated assessment of uncertainty propagating through the system from parameter estimation, input forcing, and model structure.

[62] 4. The usefulness and applicability of IBUNE has also been demonstrated via a validation study over a 3-year-period from 1958 to 1960. The results confirmed that both extended SCME and IBUNE framework are convincing tools to improve the accuracy and reliability of the predictions and their associated uncertainty bounds even during validation period. The results presented here were obtained through simulation experiment; however, it would be very interesting to test the performance of this framework through a set of forecast experiments which uses forecasted inputs such as precipitation to force the hydrologic models.

[63] To demonstrate the usefulness and applicability of IBUNE, we used the same hydrologic models considered earlier. The strength of these three models was combined through IBUNE. We showed that IBUNE is a very useful and applicable technique which accounts for all of the different sources of uncertainty within the hydrologic system and results in improved model prediction uncertainty bounds that bracket higher percentage of system observations.

[64] IBUNE is a flexible framework which can be expanded by including many more hydrologic models. All three major components of the framework, SCEM, input error model, and BMA, investigate different limitations in hydrologic modeling processes and provide more precise estimation of uncertainty bounds by confronting all of these different sources of uncertainty.

[65] Although accounting for all sources of uncertainty is very important in forecasting future devastating events, all

of the at-hand techniques including the work presented here are still too expensive to be used for real-time operational application. However, the ever increasing pace of computational power will soon provide the opportunity for operational communities to take advantage of these state-of-the-art methods to address uncertainty associated with their forecasts in a more reliable and accurate manner.

[66] **Acknowledgments.** This work was supported by the NSF Sustainability of Semi-Arid Hydrology and Riparian Areas (SAHRA) Science and Technology Center (NSF EAR-9876800) and HyDIS project (NASA grant NAG5-8503). The work of the second author was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under contract W-7405-Eng-48. The authors acknowledge Jasper Vrugt for providing the MATLAB code of SCEM and also SAC-SMA which was revised for this paper. This work was completed at the Department of Civil and Environmental Engineering, University of California, Irvine, as part of Newsha K. Ajami's dissertation work.

References

- Abrahart, R. J., and L. See (2002), Multi-model data fusion for river flow forecasting: An evaluation of six alternative methods based on two contrasting catchments, *Hydrol. Earth Syst. Sci.*, 6(4), 655–670.
- Ajami, N. K., Q. Duan, H. Moradkhani, and S. Sorooshian (2005), Recursive Bayesian model combination for streamflow forecasting, paper presented at the American Meteorological Society meeting, San Diego, Calif., Jan.
- Ajami, N. K., Q. Duan, X. Gao, and S. Sorooshian (2006), Multi-model combination techniques for hydrological forecasting: Application to Distributed Model Intercomparison Project results, *J. Hydrometeorol.*, 7(4), 755–768.
- Beck, M. B. (1987), Water quality modeling: Review of the analysis uncertainty, *Water Resour. Res.*, 23, 1393–1442.
- Beven, K., and A. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Processes*, 6, 279–298.
- Box, G. E. P., and G. C. Tiao (1973), *Bayesian Inference in Statistical Analysis*, Addison-Wesley, Boston, Mass.
- Boyle, D. (2001), Multicriteria calibration of hydrological models, Ph.D. dissertation, Univ. of Ariz., Tucson.
- Brazil, L. E. (1988), Multilevel calibration strategy for complex hydrologic simulation models, Ph.D. dissertation, 217 pp., Dep. of Civ. Eng., Colo. State Univ., Fort Collins.
- Burnash, R. J., R. L. Ferral, and R. A. McGuire (1973), A generalized streamflow simulation system conceptual modeling for digital computers, report, U. S. Dep. of Commer. Natl. Weather Serv. and State of Calif. Dep. of Water Resour.
- Clyde, M. A. (1999), Bayesian model averaging and model search strategies, in *Bayesian Statistics*, vol. 6, edited by J. M. Bernardo et al., pp. 157–185, Oxford Univ. Press, New York.
- Dempster, A. P., N. M. Laird, and D. B. Rubin (1977), Maximum likelihood from incomplete data via the EM algorithm, *J. R. Stat. Soc., Ser. B*, 39, 1–39.
- Duan, Q., S. Sorooshian, and V. K. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, 28(4), 265–284.
- Ellison, A. M. (2004), Bayesian inference in ecology, *Ecol. Lett.*, 7, 509–520.
- Evenson, G. (1992), Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model, *J. Geophys. Res.*, 97, 17,905–17,924.
- Fernandez, C., E. Ley, and M. Steel (2001), Benchmark priors for Bayesian model averaging, *J. Econometrics*, 100, 381–427.
- Gelman, A., and D. B. Rubin (1992), Inference from iterative simulation using multiple sequences, *Stat. Sci.*, 7, 457–472.
- Georgakakos, K. P., D. J. Seo, H. V. Gupta, J. Schake, and M. B. Butts (2004), Characterizing streamflow simulation uncertainty through multi-model ensembles, *J. Hydrol.*, 298(1–4), 222–241.
- George, E. I., and R. E. McCulloch (1993), Variable selection via Gibbs sampling, *J. Am. Stat. Assoc.*, 88, 881–889.
- Gupta, V. H., S. Sorooshian, and P. O. Yapo (1998), Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resour. Res.*, 34(4), 751–763.
- Hastings, W. K. (1970), Monte-Carlo sampling methods using Markov Chains and their applications, *Biometrika*, 57, 97–109.
- Hoeting, J. A., D. Madigan, A. E. Raftery, and C. T. Volinsky (1999), Bayesian model averaging: A tutorial, *Stat. Sci.*, 14(4), 382–417.
- Hogue, T. S., H. V. Gupta, S. Sorooshian, and C. D. Tomkins (2003), A multi-step automatic calibration scheme for watershed models, in *Calibration of Watershed Models*, *Water Sci. Appl. Ser.*, vol. 6, edited by Q. Duan et al., pp. 165–174, AGU, Washington, D. C.
- Kavetski, D., S. W. Franks, and G. Kuczera (2003), Confronting input uncertainty in environmental modeling, in *Calibration of Watershed Models*, *Water Sci. Appl. Ser.*, vol. 6, edited by Q. Duan et al., pp. 49–68, AGU, Washington, D. C.
- Kitanidis, P. K., and R. L. Bras (1980a), Adaptive filtering through detection of isolated transient errors in rainfall-runoff models, *Water Resour. Res.*, 16, 740–748.
- Kitanidis, P. K., and R. L. Bras (1980b), Real-time forecasting with a conceptual hydrological model: 1. Analysis of uncertainty, *Water Resour. Res.*, 16, 1025–1033.
- Kuczera, G., and E. Parent (1998), Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The metropolis algorithm, *J. Hydrol.*, 211, 69–85.
- Madigan, D., A. E. Raftery, C. Volinsky, and J. Hoeting (1996), Bayesian model averaging, in *Proceedings of the AAAI Workshop on Integrating Multiple Learned Models*, pp. 77–83, AAAI Press, Portland, Ore.
- Marshall, L., A. Sharma, and D. J. Nott (2003), Efficient Bayesian model selection in hydrological modeling, in *Proceedings of the International Congress on Modelling and Simulation (MODSIM2003)*, edited by D. A. Post, pp. 831–836, Modell. and Simul. Soc. of Australia and N. Z., Townsville, Australia, Australia.
- McLachlan, G. J., and T. Krishnan (1997), *The EM Algorithm and Extensions*, John Wiley, Hoboken, N. J.
- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953), Equation of state calculations by fast computing machines, *J. Chem. Phys.*, 21, 1087–1091.
- Miller, R. N., M. Ghil, and F. Ghautiez (1994), Advanced data assimilation in a strongly nonlinear dynamical system, *J. Atmos. Sci.*, 51, 1037–1055.
- Montanari, A., and A. Brath (2004), A stochastic approach for assessing the uncertainty of rainfall-runoff simulations, *Water Resour. Res.*, 40, W01106, doi:10.1029/2003WR002540.
- Moradkhani, H., K. L. Hsu, H. V. Gupta, and S. Sorooshian (2005), Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, *Water Resour. Res.*, 41, W05012, doi:10.1029/2004WR003604.
- Neuman, S. P. (2003), Maximum likelihood Bayesian averaging of uncertain model predictions, *Stochast. Environ. Res. Risk Assess.*, 17, 291–305.
- Neuman, S. P., and P. J. Wierenga (2003), A comprehensive strategy of hydrologic modeling and uncertainty analysis for nuclear facilities and sites, *NUREG/CR-6805*, U. S. Nucl. Regul. Comm., Washington, D. C.
- Raftery, A. E., and Y. Zheng (2003), Discussion: Performance of Bayesian model averaging, *J. Am. Stat. Assoc.*, 98(464), 931–938.
- Raftery, A. E., D. Madigan, and C. T. Volinsky (1997), Model selection and accounting for model uncertainty in linear regression models, *J. Am. Stat. Assoc.*, 92, 179–191.
- Raftery, A. E., F. Balabdaoui, T. Gneiting, and M. Polakowski (2003), Using Bayesian model averaging to calibrate forecast ensembles, *Tech. Rep. 440*, Dep. of Stat., Univ. of Wash., Seattle.
- Raftery, A. E., T. Gneiting, F. Balabdaoui, and M. Polakowski (2005), Using Bayesian model averaging to calibrate forecast ensembles, *Mon. Weather Rev.*, 133, 1155–1174.
- Schaake, J. C., V. I. Koren, Q. Y. Duan, K. Mitchell, and F. Chen (1996), Simple water balance model for estimating runoff at different spatial and temporal scales, *J. Geophys. Res.*, 101(D3), 7461–7475.
- Shamseldin, A. Y., K. M. O'Connor, and G. C. Liang (1997), Methods for combining the outputs of different rainfall-runoff models, *J. Hydrol.*, 197, 203–229.
- Sorooshian, S., and J. A. Dracup (1980), Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: Correlated and heteroscedastic error cases, *Water Resour. Res.*, 16(2), 430–442.
- Thiemann, M., M. Trosset, H. Gupta, and S. Sorooshian (2001), Bayesian recursive parameter estimation for hydrologic models, *Water Resour. Res.*, 37(10), 2521–2535.
- Viallefond, V., A. E. Raftery, and S. Richardson (2001), Variable selection and Bayesian model averaging in epidemiological case-control studies, *Stat. Med.*, 20, 3215–3230.
- Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003), A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, 39(8), 1201, doi:10.1029/2002WR001642.

- Vrugt, J. A., C. G. H. Diks, H. V. Gupta, W. Bouten, and J. M. Verstraten (2005), Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, *Water Resour. Res.*, *41*, W01017, doi:10.1029/2004WR003059.
- Wagener, T., D. P. Boyle, M. J. Lees, H. S. Wheeler, H. V. Gupta, and S. Sorooshian (2001), A framework for development and application of hydrological models, *Hydrol. Earth Syst. Sci.*, *5*(1), 13–26.
- Wintle, B. A., M. A. McCarthy, C. T. Volinsky, and R. P. Kavanagh (2003), The use of Bayesian model averaging to better represent uncertainty in ecological models, *Conserv. Biol.*, *17*(6), 1579–1590.
- Yapo, P. O., H. V. Gupta, and S. Sorooshian (1998), Multi-objective global optimization for hydrologic models, *J. Hydrol.*, *204*, 83–97.
- N. K. Ajami, University of California at Berkeley, Department of Civil and Environmental Eng., University of California, Berkeley, 626 Davis Hall, Berkeley, CA 94720, USA. (newshaajami@berkeley.edu)
- Q. Duan, Lawrence Livermore National Laboratory, P.O. Box 808, L0631, 7000 East Avenue, Livermore, CA 94551, USA.
- S. Sorooshian, Center for Hydrometeorology and Remote Sensing, Department of Civil and Environmental Engineering, University of California, Irvine, E-4130 Engineering Gateway, Irvine, CA 92679, USA.

Other Key Reference Papers for Prof. Duan's Lecture:

- [1] Duan, Q., and T.J. Phillips, 2010. Bayesian Estimation of Local Signal and Noise in Multi-Model Simulations of Climate Change, *J. Geophys. Res.*, 2009D013654, in press.
- [2] Q. Duan, N. Ajami, X. Gao, S. Sorooshian, 2006, “Multi-model Hydrologic Ensemble Predictions Using Bayesian Model Averaging”, *Advances in Water Resources*, doi:10.1016/j.advwatres.2006.11.014
- [3] Montgomery, D.C., 2005. *Design and Analysis of Experiments*, John Wiley & Sons, New York
- [4] Hastie, T., R. Tibishirani, J. Friedman, 2004. *The elements of statistical learning*. Springer-Verlag,

Model–data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications

M. R. RAUPACH*, P. J. RAYNER†, D. J. BARRETT‡, R. S. DEFRIESS§, M. HEIMANN¶, D. S. OJIMA||, S. QUEGAN** and C. C. SCHMULLIUS††

*CSIRO Earth Observation Centre, GPO Box 3023, Canberra, ACT 2601, Australia, †LSCE-CEA de Saclay Orme des Merisiers, 91191 Gif/Yvette, France, ‡CSIRO Land and Water, Canberra, ACT 2601, Australia, §Department of Geography, University of Maryland, College Park, MD 20742, USA, ¶Department of “Biogeochemical Systems”, Max-Planck-Institut für Biogeochemie, D-07701, Jena, Germany, ||Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO 80523-1499, USA, **Centre for Terrestrial Carbon Dynamics, University of Sheffield, Sheffield S37RH, UK, ††Institute für Geographie, Friedrich-Schiller-Universität, D-07743 Jena, Germany.

Abstract

Systematic, operational, long-term observations of the terrestrial carbon cycle (including its interactions with water, energy and nutrient cycles and ecosystem dynamics) are important for the prediction and management of climate, water resources, food resources, biodiversity and desertification. To contribute to these goals, a terrestrial carbon observing system requires the synthesis of several kinds of observation into terrestrial biosphere models encompassing the coupled cycles of carbon, water, energy and nutrients. Relevant observations include atmospheric composition (concentrations of CO₂ and other gases); remote sensing; flux and process measurements from intensive study sites; *in situ* vegetation and soil monitoring; weather, climate and hydrological data; and contemporary and historical data on land use, land use change and disturbance (grazing, harvest, clearing, fire).

A review of model–data synthesis tools for terrestrial carbon observation identifies ‘nonsequential’ and ‘sequential’ approaches as major categories, differing according to whether data are treated all at once or sequentially. The structure underlying both approaches is reviewed, highlighting several basic commonalities in formalism and data requirements.

An essential commonality is that for all model–data synthesis problems, both nonsequential and sequential, data uncertainties are as important as data values themselves and have a comparable role in determining the outcome.

Given the importance of data uncertainties, there is an urgent need for soundly based uncertainty characterizations for the main kinds of data used in terrestrial carbon observation. The first requirement is a specification of the main properties of the error covariance matrix.

As a step towards this goal, semi-quantitative estimates are made of the main properties of the error covariance matrix for four kinds of data essential for terrestrial carbon observation: remote sensing of land surface properties, atmospheric composition measurements, direct flux measurements, and measurements of carbon stores.

Received 8 June 2004; accepted 25 August 2004

Introduction

Systematic earth observation implies the collection and interpretation of multiple kinds of data about the evolving state of the earth system across wide spatial domains and over extended time periods. Three factors have caused a massive acceleration in earth observation activities over recent years. The first is need: global change is raising issues – such as greenhouse-induced climate change, water shortages and imbalances, land degradation, soil erosion, loss of biodiversity – which require informed human responses at both global and regional levels. Second, technological advances in sensors, satellite systems and data storage and processing capabilities are making possible observations and interpretations which were out of reach only a few years ago and unimaginable a few decades ago. Third, the synthesis of formerly discrete disciplines into a unified Earth System Science is driving new hypotheses about the dynamics of the earth system and the interconnectedness of its components, including humans. Systematic earth observation motivates and tests these hypotheses.

The focus of this paper is observation of the carbon cycle, and in particular its land-atmosphere components, as one part of an integrated earth observation system. It is a significant part because of the coupling between the carbon cycle and the terrestrial cycles of water, energy and nutrients, and the connections of all these biospheric processes with global climate and human activities (Field & Raupach, 2004; Raupach *et al.*, 2004). The carbon cycle is integral to the growth and decay of vegetation, maintains the water cycle through transpiration and provides habitat for maintaining biodiversity. Thus, terrestrial carbon observation is important for climate observation and prediction, for the management of water resources, nutrients and biodiversity, and for monitoring and managing the enhanced greenhouse effect.

It is increasingly recognized that strategies for earth observation (including terrestrial carbon observation) require methods for combining data and process models in systematic ways. This is leading to research towards the application in terrestrial carbon observation (and in earth observation more generally) of 'model–data synthesis', the combination of the information contained in both observations and models through both parameter-estimation and data-assimilation techniques. Motivations for model–data synthesis approaches include (1) model testing and data quality control (through systematic checks for agreement within specified uncertainty bands for both data and model); (2) interpolation of spatially and temporally sparse observations; (3) inference from available ob-

servations of quantities which are not directly observable (such as carbon stores and fluxes over large areas) and (4) forecasting (prediction forward in time on the basis of past and current observations).

The present paper arose from a workshop held in Sheffield, UK, 3–6 June 2003, to further the development of a Terrestrial Carbon Observation System (TCOS) with a particular emphasis on model–data synthesis. Antecedents for this effort were (1) preliminary steps toward a TCOS (Cihlar *et al.*, 2002a, b, c); (2) a wider concept for an Integrated Global Carbon Observing Strategy including atmosphere, oceans, land and human activities (Ciais *et al.*, 2004) and (3) the research program of the Global Carbon Project (Global Carbon Project, 2003).

The paper is founded on three themes arising from the Sheffield workshop. First, model–data synthesis, based on terrestrial biosphere models constrained with multiple kinds of observation, is an essential component of a TCOS. Second, from the standpoint of model–data synthesis, data uncertainties are as important as data values themselves and have a comparable role in determining the outcome. Third, and consequently, there is an urgent need for soundly based uncertainty specifications for the main kinds of data used in terrestrial carbon observation. These themes are developed as follows: the next section summarizes major purposes and attributes of a TCOS. 'Model–data synthesis: methods' provides an overview of model–data synthesis in the context of terrestrial carbon observation, by briefly describing some of the main methods, indicating their common characteristics, and highlighting the key role of data uncertainty. 'Model–data synthesis: examples' provides some examples. 'Data characteristics: uncertainty in measurement and representation' undertakes a survey of the uncertainty characteristics of the main kinds of relevant data.

Purposes and attributes of a TCOS

A succinct statement of the overall purpose of a TCOS might be: to operationally monitor the cycles of carbon and related entities (water, energy, nutrients) in the terrestrial biosphere, in support of comprehensive, sustained earth observation and prediction, and hence sustainable environmental management and socio-economic development. These words are congruent with the Framework Document emerging from the Second Earth Observation Summit, Tokyo, April 2004 (<http://earthobservations.org/docs/Framework%20Doc%20Final.pdf>), which calls for a 'Global Earth Observation System of Systems' to serve nine areas of socio-economic benefit. A TCOS is a contributor to such a system with relevance to at least six of these areas:

- Understanding climate, and assessing and mitigating climate change impacts;
- Improving global water resource management and understanding of the water cycle;
- Improving weather information and prediction;
- Monitoring and managing inland ecosystems, including forests, and land use change;
- Supporting sustainable agriculture and combating desertification;
- Understanding, monitoring and preventing loss of biodiversity.

To make these contributions effectively, a TCOS must have a number of attributes (see also Running *et al.*, 1999; Cihlar *et al.*, 2002a; Ciais *et al.*, 2004). First, *scientific credibility* is needed to maintain methodological and observational rigour, and to include procedures for estimating uncertainties or confidence limits. Second, *consistency with global budgets* is necessary to respect constraints from global-scale carbon and related budgets incorporating terrestrial, atmospheric and oceanic pools and anthropogenic sources such as fossil fuel burning. Third, sufficient *spatial resolution* is necessary to resolve spatial variations in patterns of land use (typically tens of metres, consistent with high-resolution remote sensing). Fourth, enough *temporal resolution* is needed to resolve the influence of weather, inter-annual climate fluctuations and long-term climate change on carbon and related cycles. Fifth, the system needs to encompass a broad *range of entities*, eventually including CO₂, CH₄, CO, volatile organic carbons (VOCs) and aerosol black carbon. Of these, the highest priority is CO₂. Water is also a high priority because of its importance in modulating other terrestrial GHG fluxes. Sixth, a sufficient *range of processes* must be encompassed. A high priority is resolution of net land-air fluxes of greenhouse gases in which all terrestrial sources and sinks are lumped together. However, there is an equally high demand for identification of the terms contributing to the net fluxes, for example to partition a net flux between vegetation and soil storage changes. Finally, *quantification of uncertainty* is required. The 'demand side' of the uncertainty issue is: what level of uncertainty is acceptable for a TCOS to offer useful information? The answer is not simple and depends on the application, for example, from the areas mentioned above. This paper does not attempt to answer the demand-side question, but rather concentrates on the 'supply side' of uncertainty: that is, how uncertainty can be determined in a TCOS based on model-data synthesis and multiple observation sources, each with its own specified uncertainty.

Model-data synthesis: methods

In this section, we survey a range of model-data synthesis methods potentially applicable in a TCOS. More detail and further references can be found in a growing number of excellent sources, for instance Tarantola (1987) and Evans & Stark (2002) for high-level treatments of the general statistical problem of inverse estimation, Grewal & Andrews (1993) and Drécourt (2003) for introductions to the Kalman Filter, Reichle *et al.* (2002) for hydrological applications with an emphasis on the Kalman Filter and Enting (2002) and Kasibhatla *et al.* (2000) for applications of a range of methods to biogeochemical cycles.

Overview

The central problem is: using appropriate observations and models, we must determine the spatial distributions and temporal evolutions of the terrestrial stores and fluxes of carbon and related entities (water, nutrients, energy) across the earth. Important fluxes include land-air exchanges (atmospheric sources and sinks), exchanges with rivers and groundwater, and exchanges between terrestrial pools such as biomass and soil. We also need to determine the main processes influencing the fluxes, including those under human management. No single model or set of observations can supply this amount of information – hence the need for a synthesis approach. The task of combining observations and models can be carried out in many ways, encompassed by the umbrella terms 'model-data synthesis' or 'model-data fusion'. The general principle is to find an 'optimal match' between observations and model by varying one or more 'properties' of the model. (Words in quotes have specific meanings defined below). The optimal match is a choice of model properties, which minimizes the 'distance' between the model representations of a system and what we know about the real biophysical system from observational and prior 'data'. At this high level of generality, model-data synthesis encompasses both 'parameter estimation' and 'data assimilation'. All applications rest on three foundations: a model of the system, data about the system, and a synthesis approach.

Model. For a TCOS, the model is a terrestrial biosphere model describing the evolving stores and fluxes of carbon, water, energy and related entities. This *dynamic model* has the form

$$\begin{aligned}\frac{dx}{dt} &= f(\mathbf{x}, \mathbf{u}, \mathbf{p}) + \text{noise} \quad \text{or} \\ \mathbf{x}^{n+1} &= \boldsymbol{\phi}(\mathbf{x}^n, \mathbf{u}^n, \mathbf{p}) + \text{noise} = \mathbf{x}^n + \Delta t f(\mathbf{x}^n, \mathbf{u}^n, \mathbf{p}) + \text{noise},\end{aligned}\tag{1}$$

where \mathbf{x} is a vector of *state variables* (such as stores of carbon, water and related entities, or store attributes such as age class distributions); \mathbf{f} is a vector of rates of change (net fluxes where components of \mathbf{x} are stores); $\boldsymbol{\phi}$ is the discrete analogue for \mathbf{f} ; \mathbf{u} is a set of externally specified time-dependent forcing variables (such as meteorological variables and soil properties) and \mathbf{p} is a set of time-independent model parameters (such as rate constants and partition ratios). In the discrete formulation, time steps are denoted by superscripts. The noise terms account for both imperfections in model formulation and stochastic variability in forcings (\mathbf{u}) or parameters (\mathbf{p}). Once the model function $\mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{p})$ or $\boldsymbol{\phi}(\mathbf{x}^n, \mathbf{u}^n, \mathbf{p})$ is specified, then the system evolution $\mathbf{x}(t)$ can be determined by integrating Eqn (1) in time (with zero noise), from initial conditions $\mathbf{x}(0)$, with specified external forcing $\mathbf{u}(t)$ and parameters \mathbf{p} .

Data. These are generally of two broad kinds: (1) observations or measurements of a set of quantities \mathbf{z} and (2) prior estimates for model quantities (\mathbf{x} , \mathbf{u} and \mathbf{p}). Both include uncertainty, through errors and noise. In this paper, the term ‘data’ includes both observations and prior estimates, and incorporates the uncertainty inherent in each.

The measured quantities (\mathbf{z}) are related to the system state and external forcing variables by an *observation model* of the form

$$\mathbf{z} = \mathbf{h}(\mathbf{x}, \mathbf{u}) + \text{noise}, \quad (2)$$

where the operator \mathbf{h} specifies the deterministic relationship between the measured quantities and the system state. The noise term accounts for both ‘measurement error’ (instrumental and processing errors in the measurements \mathbf{z}), and ‘representation error’ (errors in the model representation of \mathbf{z} , introduced by shortcomings in the observation model \mathbf{h}). In the rare case where we can observe all state variables directly, \mathbf{h} reduces to the identity operator, so $\mathbf{z} = \mathbf{x} + (\text{measurement}) \text{ noise}$. In time-discrete form, Eqn (2) becomes $\mathbf{z}^n = \mathbf{h}(\mathbf{x}^n, \mathbf{u}^n) + \text{noise}$. Note the interpretation of the time-step superscripts: \mathbf{x}^n and \mathbf{u}^n are simply the model state and forcings at time step n , whereas \mathbf{z}^n is the set of new observations introduced at time step n , whatever the actual time of its measurement. However, no observations may be used more than once.

Examples of potential observations in a TCOS include (1) atmospheric composition (concentrations of CO_2 and other gases); (2) remote sensing of terrestrial and atmospheric properties; (3) fluxes of carbon and related entities, with supporting process observations, at intensive study sites; (4) vegetation and soil stores of carbon from forest and ecological inventories; (5)

hydrological data on river flows, groundwater, and concentrations of C, N and other entities; (6) soil properties and topography; (6) disturbance records (both contemporary and historical) including land management, land use, land use change and fire and (8) climate and weather data (precipitation, solar radiation, temperature and humidity). Of these, some (especially the first five) typically provide observational constraints (\mathbf{z}), while others provide model drivers (\mathbf{u}). Examples of observation models (Eqn (2)) include radiative transfer models to map modelled surface states into the radiances observed by satellites; atmospheric transport models to transform modelled surface fluxes to measured atmospheric concentrations; and allometric relations to transform modelled biomass to observed tree diameters.

Synthesis. The final requirement is a synthesis process, or a systematic method for finding the optimal match between the data (including observations and prior estimates) and the model. This process needs to provide three kinds of output: optimal estimates for the model properties to be adjusted, uncertainty statements about these estimates, and an assessment of how well the model fits the data, given the data uncertainties. In any synthesis process, there are three basic choices: (1) the model properties to be adjusted or ‘target variables’, (2) the measure of distance between data and model or ‘cost function’ and (3) the search strategy for finding the optimum values. Search strategies can be classified broadly into (3a) ‘nonsequential’ or ‘batch’ strategies in which the data are treated all at once, and (3b) ‘sequential’ strategies in which the data arrive in a time sequence and are incorporated into the model–data synthesis step by step. The rest of this section explores the choices (1), (2), (3a) and (3b).

Target variables

The target variables are the properties of the model to be adjusted in the optimization process. They include any model property considered to be sufficiently uncertain as to benefit from constraint by the data. Model properties which can be target variables include: (1) model parameters (\mathbf{p}); (2) forcing variables (\mathbf{u}^n), if there is substantial uncertainty about them; (3) initial conditions on the state variables (\mathbf{x}^0) and (4) time-dependent components of the state vector \mathbf{x}^n . The inclusion of the state vector \mathbf{x}^n as a possible target variable is for the following reason: in a purely deterministic model the trajectory \mathbf{x}^n is determined by the dynamical model (\mathbf{f} or $\boldsymbol{\phi}$), the values of \mathbf{p} and \mathbf{u}^n , and the initial value \mathbf{x}^0 . It might seem sufficient, therefore, to estimate these and allow integration of

the model to take care of \mathbf{x}'' . However, the model itself may not be perfect, as indicated by the noise term in Eqn (1), so there may be advantage in adjusting values of \mathbf{x}'' through the model integration.

To maintain generality, we denote the vector of target variables by \mathbf{y} . This vector may or may not be a function of time, and will usually be a subset of all model variables ($\mathbf{x}'', \mathbf{u}'', \mathbf{p}$). Broadly speaking, parameter estimation problems are those where the target variables are restricted to model parameters (\mathbf{p}), while data assimilation problems may include any model property as a target variable, usually with an emphasis on state variables (\mathbf{x}'').

Cost function

The cost or objective function J (a function of the target variables \mathbf{y}) defines the mismatch or distance between the model and the data. It can take a wide range of forms, but must have certain properties (for example, it must be monotonic in the absolute difference between data and model-predicted values). A common choice is the quadratic cost function:

$$J(\mathbf{y}) = (\mathbf{z} - \mathbf{h}(\mathbf{y}))^T [\text{Cov } \mathbf{z}]^{-1} (\mathbf{z} - \mathbf{h}(\mathbf{y})) + (\mathbf{y} - \hat{\mathbf{y}})^T [\text{Cov } \hat{\mathbf{y}}]^{-1} (\mathbf{y} - \hat{\mathbf{y}}), \quad (3)$$

where $\hat{\mathbf{y}}$ is the vector of 'priors' (*a priori* estimates) for the target variables, and $[\text{Cov } \mathbf{z}]$ and $[\text{Cov } \hat{\mathbf{y}}]$ are covariance matrices for \mathbf{z} and $\hat{\mathbf{y}}$, respectively ($[\text{Cov } \mathbf{z}]_{mm} = \langle z'_m z'_m \rangle$, with $z'_m = z_m - \langle z_m \rangle$, angle brackets denoting the expectation operator). The first term in Eqn (3) is a sum of the squared distances between measured components of the observation vector (\mathbf{z}) and their model predictions ($\mathbf{h}(\mathbf{y})$), while the second is a corresponding sum of distances between target variables and their prior estimates. The matrices $[\text{Cov } \mathbf{z}]^{-1}$ and $[\text{Cov } \hat{\mathbf{y}}]^{-1}$ represent the weights accorded to the observations and the priors, and thus scale the confidences accorded to each. Their role can be clarified by considering the simple case in which components z_m of the observation vector \mathbf{z} are independent, with variances σ_m^2 ; then $[\text{Cov } \mathbf{z}]^{-1}$ is the diagonal matrix $\text{diag}[1/\sigma_m^2]$ and the squared departures of the measurements (z_m) from the predictions ($h_m(\mathbf{y})$) are seen to be weighted by the confidence measure $1/\sigma_m^2$ for each component.

The model-data synthesis problem now becomes: vary \mathbf{y} to minimize $J(\mathbf{y})$, subject to the constraint that $\mathbf{x}(t)$ must satisfy the dynamic model, Eqn (1). The value of \mathbf{y} at the minimum is the *a posteriori* estimate of \mathbf{y} , including information from the observations as well as the priors. We denote it by $\tilde{\mathbf{y}}$ (so frowns and smiles respectively designate prior and posterior estimates).

Equation (3) defines the generalized least squares cost function minimized by the *minimum-variance* estimate ($\tilde{\mathbf{y}}$) for \mathbf{y} . For any distribution of the errors in the data (observations \mathbf{z} and priors $\hat{\mathbf{y}}$), this estimate is unbiased, and has the minimum error covariance among all linear (in \mathbf{z}), unbiased estimates (Tarantola 1987). Use of Eqn (3) has another, additional foundation: provided that the probability distributions for data errors are Gaussian, it yields a *maximum-likelihood* estimate for \mathbf{y} , conditional on the data and the model dynamics (Press *et al.*, 1992, p. 652; Todling 2000). Outside the restriction of Gaussian distributions, $\tilde{\mathbf{y}}$ as defined by minimizing a quadratic J is not exactly the maximum-likelihood estimate, but it is often not far from it. A quadratic J is widely used even when the data errors are not Gaussian; see Press *et al.* (1992, p. 690) for discussion. There are alternative cost functions J in which model-measurement differences ($\mathbf{z} - \mathbf{h}(\mathbf{y})$) are raised to powers other than 2, the choice in Eqn (3) (Tarantola, 1987; Gershensfeld, 1999). For example, in flood event modelling, the absolute maximum error is needed to capture peak flow rates, while for modelling base flow rates, the mean absolute deviation ($|\mathbf{z} - \mathbf{h}(\mathbf{y})|$ to the power 1) has the desirable property of being less sensitive to outliers than a power 2. Different powers for $|\mathbf{z} - \mathbf{h}(\mathbf{y})|$ produce maximum-likelihood estimates for $\tilde{\mathbf{y}}$ with different distributions for data errors; for example, a power 1 J yields a maximum-likelihood estimate when the data errors are distributed exponentially, and a high-power J preferentially weights outliers such as peak flows. Here, we use a power 2 J exclusively.

Search strategies for nonsequential problems

In nonsequential or batch problems, all data are treated simultaneously and the minimization problem is solved only once. A familiar case is least-squares parameter estimation.

Example. Some of the attributes of these problems are demonstrated by considering a simple linear example, which extends the parameter-estimation problem. Although mathematically straightforward, this case finds important application in the atmospheric inversion methods used to estimate trace gas sources from atmospheric composition observations (see 'Model-data synthesis: Examples'). Here the target variables (\mathbf{y}) are a set of surface-air fluxes, averaged over suitable areas; there is no dynamic model relating fluxes at different times and places to each other; and the observation operator (\mathbf{h}) is a model of atmospheric transport. From the linearity of the conservation equation for an inert trace gas, it follows that \mathbf{h} is linear and can hence be represented by a matrix \mathbf{H}

(Raupach, 2001), thus, $\mathbf{z} = \mathbf{H}\mathbf{y} + \text{noise}$. For now, the noise is assumed to be Gaussian with zero mean and no temporal correlation, and thus completely characterised by an observation error covariance matrix $[\text{Cov}\mathbf{z}]$. By minimizing J analytically, one obtains the expression (Tarantola, 1987, p. 196; Enting, 2002):

$$\tilde{\mathbf{y}} = \hat{\mathbf{y}} + [\text{Cov}\tilde{\mathbf{y}}]\mathbf{H}^T[\text{Cov}\mathbf{z}]^{-1}(\mathbf{z} - \mathbf{H}\hat{\mathbf{y}}), \quad (4)$$

where $[\text{Cov}\tilde{\mathbf{y}}]$, the estimated error covariance of the a posteriori estimate $\tilde{\mathbf{y}}$, is given by

$$[\text{Cov}\tilde{\mathbf{y}}]^{-1} = [\text{Cov}\hat{\mathbf{y}}]^{-1} + \mathbf{H}^T[\text{Cov}\mathbf{z}]^{-1}\mathbf{H}. \quad (5)$$

These expressions already tell us some important things. The posterior estimates are given by the prior estimates plus a term depending on the mismatch between the experimental observations and the observations as predicted by the prior estimates. This mismatch is weighted by our confidence in the observations, $[\text{Cov}\mathbf{z}]^{-1}$. Thus, observations with little weight hardly shift the posterior estimate from the prior, and *vice versa*. Furthermore, the transpose of the observation operator (\mathbf{H}^T) multiplies the weighted mismatch. If this operator is very weak, that is if the available observations are only weakly related to the target variables, then the update to the initial estimate is also small. Finally the posterior covariance $[\text{Cov}\tilde{\mathbf{y}}]$ (Eqn (5)) is bounded above, in some sense, by the prior covariance $[\text{Cov}\hat{\mathbf{y}}]$. If the prior covariance is small (suggesting substantial confidence in the initial estimate) then the increment $\tilde{\mathbf{y}} - \hat{\mathbf{y}}$ (the difference between the posterior and prior estimates, a measure of the information added by the observations \mathbf{z} , and equal to the second term in Eqn (4) in the present case) is also small.

All the above is reasonable. More surprising is the relationship between the data, its uncertainty and the cost function. We can decompose the (positive definite) matrix $[\text{Cov}\mathbf{z}]^{-1}$ into a matrix product $\mathbf{A}^T\mathbf{A}$, using the Cholesky decomposition for a positive definite matrix. For a diagonal covariance matrix, $\text{diag}[\sigma_m^2]$, the decomposition is trivial: $\mathbf{A} = \text{diag}[1/\sigma_m]$. Likewise, we can write $[\text{Cov}\hat{\mathbf{y}}]^{-1} = \mathbf{B}^T\mathbf{B}$. The cost function, Eqn (3), can then be rewritten as

$$J(\mathbf{y}) = (\mathbf{a} - \mathbf{A}\mathbf{h}(\mathbf{y}))^T(\mathbf{a} - \mathbf{A}\mathbf{h}(\mathbf{y})) + (\mathbf{b} - \hat{\mathbf{b}})^T(\mathbf{b} - \hat{\mathbf{b}}), \quad (6)$$

where $\mathbf{a} = \mathbf{A}\mathbf{z}$, $\mathbf{b} = \mathbf{B}\mathbf{y}$ and $\hat{\mathbf{b}} = \mathbf{B}\hat{\mathbf{y}}$. Thus the cost function, and thence the entire minimization, takes a form in which neither the observations nor the prior estimates appear; they are replaced by quantities \mathbf{a} and \mathbf{b} scaled by the square roots of the inverse covariance matrices, which are measures of confidence. This is no mathematical nicety; rather it demonstrates that the

data and the uncertainties are completely inseparable in the formalism. To put the point provocatively, providing data and allowing another researcher to provide the uncertainty is indistinguishable from allowing the second researcher to make up the data in the first place. This realization informs the emphasis on uncertainty throughout this paper.

Algorithms for nonsequential problems

The task in general is to find the target variables \mathbf{y} which minimize $J(\mathbf{y})$. Clearly, the shape of $J(\mathbf{y})$ is all important: it may have a single minimum or multiple separated local minima, only one of which is the true global minimum. Near the minimum, J may be shaped like a long, narrow ellipsoidal valley. If this valley has a flat floor tracing out some line in \mathbf{y} space, then all points along that line are equally acceptable and these \mathbf{y} coordinates cannot be distinguished in terms of optimality, so such combinations of target variables cannot be resolved by model-data synthesis with the available data and model. Diagnostic indicators about these issues are provided by the Hessian or curvature matrix $\mathbf{D} = \partial^2 J / \partial y_j \partial y_k$, a measure of the local curvature of $J(\mathbf{y})$. The degree of orthogonality among columns of \mathbf{D} indicates the extent to which it is possible to find a unique local minimum to $J(\mathbf{y})$ in the vicinity of the point at which \mathbf{D} is evaluated. A high 'condition number' (ratio of largest to smallest eigenvalue) for \mathbf{D} indicates that some linear combination(s) of the columns of \mathbf{D} are nearly zero, that is, that the curvature is nearly zero in some direction(s), so that the minimization problem is ill-conditioned, as in the case of a valley with a flat floor.

Given these considerations, classes of method for finding the minimum in $J(\mathbf{y})$ include the following.

1. *Analytic solution* is possible when the observation operator $\mathbf{h}(\mathbf{y})$ is linear ($\mathbf{z} = \mathbf{H}\mathbf{y} + \text{noise}$). In this case $J(\mathbf{y})$ is a quadratic form shaped like a parabolic bowl, and the minimization can be carried out analytically as in the example of Eqns (4) and (5). This 'direct' or 'one-step' solution is highly efficient when applicable; however, most problems are nonlinear and require a nonlinear method.

2. *Gradient descent algorithms* are the most familiar search algorithms for nonlinear optimization. They include (for example) steepest-descent, conjugate-gradient, quasi-Newton and Levenberg-Marquardt algorithms (Press *et al.*, 1992). Gradient-descent methods are easily implemented, provided that the gradient vector $\nabla_{\mathbf{y}}J = \partial J / \partial y_k$ can be calculated. The main advantages of gradient-descent algorithms are relative simplicity and low cost; the main disadvantage is that if the surface $J(\mathbf{y})$ has multiple minima, they tend to find local

minima near the starting value of \mathbf{y} rather than the global minimum.

3. *Global search methods* find the global minimum in a function $J(\mathbf{y})$ by searching (effectively) the whole of \mathbf{y} space. They overcome the local-minimum pitfall (so to speak) of gradient-descent methods, but have the disadvantage of higher computational costs. Simulated annealing and genetic algorithms are two examples. These methods are efficient at finding the vicinity of a global minimum where there may be multiple local minima, but do not locate an exact local minimum. They may be combined with gradient-descent methods for finding an exact global minimum once in the right vicinity.

Search strategies for sequential problems

In sequential problems, the task is to solve for a set of target variables \mathbf{y}^n associated with a particular time step, usually including the state variables of the dynamic model (\mathbf{x}^n). The process is then repeated sequentially to give a time history for \mathbf{y}^n . Information about \mathbf{y}^n can come from two sources: evolution of the dynamic model from the previous time step, and comparison between the observations at the current time step (\mathbf{z}^n) and the model predictions ($\mathbf{h}(\mathbf{y}^n)$).

Kalman filter. Introduced by Kalman (1960), the Kalman filter is by now a group of algorithms for the sequential combination of dynamic and observational information, using a 'prediction' step and an 'analysis' step. In the prediction step, the dynamic model is used to calculate prior estimates $\hat{\mathbf{y}}^n$ for the target variables at time step n , from the best (posterior) estimates $\hat{\mathbf{y}}^{n-1}$ at the previous step. In the analysis step, posterior estimates $\hat{\mathbf{y}}^n$ at step n are obtained by 'improving' the prior estimates with data. The model state is then ready for evolution to the next ($n+1$) time step. A key point is that the confidence in the current state, embodied in the error covariance for the target variables \mathbf{y} , is also evolved with the dynamic model and improved with observations. A schematic diagram of the information flow in the Kalman filter is given in Fig. 1.

In the prediction step, the task of evolving \mathbf{y} is common across all implementations of the Kalman filter since it involves only a normal forward step of the dynamic model: $\hat{\mathbf{y}}^n = \Phi(\hat{\mathbf{y}}^{n-1})$. The prior estimate for the covariance at time step n evolves according to

$$[\text{Cov } \hat{\mathbf{y}}^n] = \Phi[\text{Cov } \hat{\mathbf{y}}^{n-1}]\Phi^T + \mathbf{Q}, \quad (7)$$

where $\Phi = \partial\phi/\partial\mathbf{y}$, the Jacobian matrix of the dynamic model $\phi(\mathbf{y})$. The first term on the right represents the propagation of the error covariance in the target

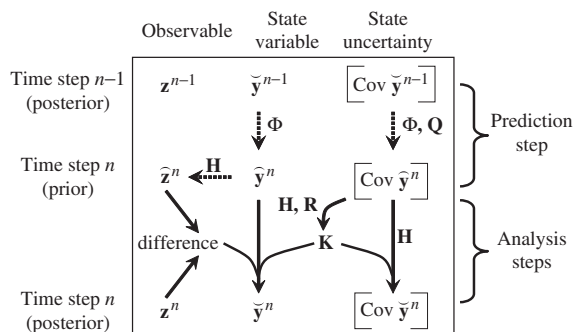


Fig. 1 Information flow in the linear Kalman filter. Linear operators are indicated next to arrows. Operations in prediction and analysis steps are shown as dashed and solid lines, respectively.

variables \mathbf{y} from one time step to the next, by a linearized version of the model. The second term (\mathbf{Q}) is the covariance of the noise term in the dynamic model, Eqn (1), which includes both model imperfections and stochastic variability in forcings and parameters. This term plays a crucial role in the Kalman filter: it quantifies our lack of confidence in the ability of the dynamic model to propagate the model state, and is usually referred to as model error. In most implementations of the Kalman filter the model error is assumed to be Gaussian with zero mean and no temporal correlation, and thus completely characterized by the covariance matrix \mathbf{Q} .

In the analysis step, the prior estimates are refined by the inclusion of data. This is done using the prior estimate for the predicted observation vector, $\hat{\mathbf{z}}^n = \mathbf{h}(\hat{\mathbf{y}}^n)$, and its covariance

$$[\text{Cov } \hat{\mathbf{z}}^n] = \mathbf{H}[\text{Cov } \hat{\mathbf{y}}^n]\mathbf{H}^T + \mathbf{R}, \quad (8)$$

where $\mathbf{H} = \partial\mathbf{h}/\partial\mathbf{y}$ is the Jacobian matrix of the observation model $\mathbf{h}(\mathbf{y})$, and \mathbf{R} is the data covariance matrix $[\text{Cov } \mathbf{z}]$, indicating lack of confidence in the data and often called the data error. Again it is usually assumed that the data error is Gaussian with zero mean and no temporal correlation, and thus completely characterized by $\mathbf{R} = [\text{Cov } \mathbf{z}]$.

The expressions for the final (posterior) estimates for \mathbf{y} and its covariance are now exactly as for the nonsequential mode, except that the operation is carried out for one time step only:

$$\begin{aligned} \hat{\mathbf{y}}^n &= \hat{\mathbf{y}}^n + [\text{Cov } \hat{\mathbf{y}}^n]\mathbf{H}^T[\text{Cov } \hat{\mathbf{z}}^n]^{-1}(\mathbf{z}^n - \mathbf{h}(\hat{\mathbf{y}}^n)) \\ &= \hat{\mathbf{y}}^n + \mathbf{K}(\mathbf{z}^n - \mathbf{h}(\hat{\mathbf{y}}^n)), \end{aligned} \quad (9)$$

$$\begin{aligned} [\text{Cov } \hat{\mathbf{y}}^n]^{-1} &= [\text{Cov } \hat{\mathbf{y}}^n]^{-1} + \mathbf{H}^T\mathbf{R}\mathbf{H} \text{ or} \\ [\text{Cov } \hat{\mathbf{y}}^n] &= (\mathbf{I} - \mathbf{K}\mathbf{H})[\text{Cov } \hat{\mathbf{y}}^n], \end{aligned} \quad (10)$$

where $\mathbf{K} = [\text{Cov } \hat{\mathbf{y}}^n] \mathbf{H}^T [\text{Cov } \hat{\mathbf{z}}^n]^{-1}$ is the Kalman gain matrix. The two equalities in Eqn (10) are equivalent. Time step n is now complete, and we are ready for the next time step. We note that the ratio of the magnitudes of \mathbf{Q} and \mathbf{R} (model and data error covariances) is critical, since it largely determines how closely the evolution of \mathbf{y} follows that suggested by the dynamic model ($\mathbf{Q} \ll \mathbf{R}$) or the data. The tuning of \mathbf{Q} and \mathbf{R} is a crucial part of Kalman filter implementation; see Grewal & Andrews (1993) for an excellent extended discussion.

The concepts underlying the Kalman filter are now implemented in several different ways (see for example Grewal & Andrews, 1993; Evensen, 1994, 2003; Kasibhatla *et al.*, 2000; Reichle *et al.*, 2002; Drécourt, 2003), including the following:

1. The linear Kalman filter (LKF), in which both $\phi(\mathbf{y})$ and $\mathbf{h}(\mathbf{y})$ are linear in \mathbf{y} , can be shown to be an optimal solution for appropriate linear problems.

2. The extended Kalman filter (EKF) applies for nonlinear $\phi(\mathbf{y})$ and $\mathbf{h}(\mathbf{y})$, by linearizing the covariance propagation part of the analysis step (Eqn (7)), but not the prediction step, at each point. This is the algorithm sketched above.

3. The ensemble Kalman filter (EnKF) (Evensen, 1994, 2003) is appropriate for high-dimensional problems such as data assimilation into atmospheric and ocean models, where the error covariance matrix for \mathbf{y} is too large to store, let alone integrate forward. The EnKF uses stochastic methods based on multiple model runs to propagate the covariance matrix without storing it. Also, the EnKF does not explicitly require the Jacobian matrices $\phi(\mathbf{y})$ and $\mathbf{h}(\mathbf{y})$, which can be difficult to derive analytically and expensive to calculate numerically. Reichle *et al.*, (2002) summarize the differences between the EKF and the EnKF.

4. The Kalman smoother assimilates multitemporal information to constrain \mathbf{y}^n at each time point, by running both forward and backward in time (Todling, 2000). It produces an estimate of target variables at time step n based on the entire record, rather than only the record up to time step n . This gives the Kalman smoother the attributes of a nonsequential method, as data at all times are used together.

Adjoint methods. These form an additional group of methods applicable to sequential problems. The principle (le Dimet & Talagrand, 1986; Giering 2000) is to update the target variables (including the model state) by using measurements at nearby times such as the interval between steps n and $n + 1$, and an estimate of the gradient $\nabla_{\mathbf{y}} J$ obtained by backward integration of an 'adjoint model' over that interval. The target

variables are effectively the initial state variables for integration of the model from step n to $n + 1$. This approach underpins four-dimensional data assimilation (4DVAR) methods for assimilating data into atmospheric and oceanic circulation models on weather and climate time scales (Chen & Lamb, 2000; Park & Zupanski, 2003).

Discussion of model–data synthesis methods

Differences between nonsequential and sequential strategies. Parameter estimation and data assimilation problems tend to be amenable to solution by nonsequential and sequential search strategies, respectively. However, this is not an absolute correspondence: many problems can be solved using either nonsequential or sequential strategies.

The most important advantage for sequential methods is the ability of the optimal state to differ from that embodied in the model equations. This requires that the evolving model state \mathbf{x}^n be included among the target variables \mathbf{y} . In principle, \mathbf{y} can also include \mathbf{x}^n in nonsequential methods but, since all time steps are considered simultaneously, the size of the problem is usually intractable. Sequential methods also have the computational advantages that their size does not grow with the length of the model integration, and that they can easily handle incremental extensions to time series observations.

The advantages of nonsequential methods come, naturally, from their ability to treat all data at once. This is a direct advantage in itself. It is, for example, difficult for a sequential method to treat the impact of a datum on a state variable some time in the past, as can occur when, for example, signals are transported through the atmosphere so that the model state at some time is only observed later. This problem is often handled with the Kalman smoother.

Model and data error structures. The noise terms in the dynamic and observation models, Eqns (1) and (2), can in principle be quite general in form, including biases, drifts, temporal correlations, extreme outliers and so on. Many extant methods take these noise terms to be Gaussian with zero mean and no temporal correlation, as assumed in Eqns (4)–(10). However, more general error structures are very common, and the development of methods for dealing with such errors is an active area of current research. In the case of biases, drifts and temporal correlations, a promising approach is to introduce extra target variables to represent these features of the model or data error. Evensen (2003) showed how this approach can be used to treat both temporal correlation and bias in the model

error. Wang & Bentley (2002) introduced a target variable representing the temporally correlated part of the data error.

Nonsequential and sequential parameter estimation. Although parameter estimation is typically carried out with nonsequential strategies such as least-squares fitting, there can be advantages in using sequential methods such as Kalman filtering for parameter estimation. The approach is to treat parameters \mathbf{p} as components of the target vector \mathbf{y} (in addition to the state variables \mathbf{x}), with \mathbf{p} governed by the dynamic equation $d\mathbf{p}/dt = 0$ (+ noise) (Grewal & Andrews 1993). This means that the problem is almost always nonlinear and must be solved with the EKF or EnKF. Annan & Hargreaves (2004) show how this technique can be used to estimate parameters in the Lorenz system with chaotic dynamics. A potential advantage of this approach is that parameters can drift through time toward new values, in response to observations. This offers a means for model-data synthesis to respond to exogenous catastrophic events (such as fire, windthrow or clearing) which suddenly change the parameters in a terrestrial biosphere model, since exogenous changes in parameters are the usual way that catastrophic events are incorporated in the absence of a full dynamic model for the processes governing the catastrophe.

Model-data synthesis: examples

The methods outlined above are being applied in several fields relevant to terrestrial carbon observation. The first major example is parameter estimation. Most biogeochemical models contain parameters (\mathbf{p}) determining photosynthetic capacities, light use efficiencies, temperature and nutrient controls on photosynthesis and respiration, pool turnover times and so on. It is almost always necessary to choose \mathbf{p} to optimize the fit of the model to test data, usually obtained from multiple study sites. Techniques for doing this range from simple graphical fits ('chi-by-eye') to least-squares fitting procedures based on Eqn (3) or other cost functions.

A second example is provided by atmospheric inversion methods for inferring the surface-atmosphere fluxes of CO_2 and other trace gases from atmospheric composition observations. The data come from global flask networks and continuous *in situ* analysers, upper-air measurements from aircraft and tall towers, and potentially in the future from remote sensing of atmospheric composition. The observation model is a model of global atmospheric transport. The basic approach has been sketched in Eqns (4) and (5). There is now a significant literature on this technique (Enting

et al., 1995; Ciais & Meijer, 1998; Enting, 1999a,b; Rayner *et al.*, 1999; Rayner, 2001; Schimel *et al.*, 2001; Enting, 2002; Gurney *et al.*, 2002). In summary, atmospheric inversions at global scale provide good constraints on total global sources and sinks, but (presently) with very coarse spatial resolution (continental to hemispheric). In addition to global applications, atmospheric inversion methods have been applied regionally (Gloor *et al.*, 2001), in the atmospheric boundary layer (Lloyd *et al.*, 1996), and in vegetation canopies (Raupach, 2001).

A third example, combining the previous two, is the use of multiple constraints. This involves model-data synthesis with the simultaneous use of multiple kinds of observations (for example, atmospheric composition measurements, remote sensing, eddy-covariance fluxes, vegetation and soil stores, and hydrological data). This approach has two advantages: first, different kinds of observation constrain different processes. For example, atmospheric composition measurements and eddy fluxes directly determine net CO_2 exchanges (net ecosystem exchange, NEE) at large and small spatial scales, respectively, while remote sensing provides indirect constraints on gross exchanges (gross primary production, GPP) through indices such as the normalized difference vegetation index (NDVI). Second, different observations have different resolutions in space and time. Through assimilation into a terrestrial biosphere model, the high space-time resolution of environmental remote sensing can add space-time texture to estimates of NEE from methods such as atmospheric inversions or eddy-covariance fluxes. Some difficulties must also be noted: for example, handling data sources with quite different spatial and temporal scales of measurement (discussed further in 'Scale mismatches between measurements and models'), and also with very different sample numbers (remotely sensed data can swamp *in situ* data with realistic error specifications, because the former has a factor of 10^3 – 10^6 more data points).

Applications of the multiple-constraint concept include the combined use of atmospheric CO_2 concentrations and surface data at continental scales (Wang & Barrett, 2003; Wang & McGregor, 2003) and global scales (Kaminski *et al.*, 2001, 2002); use of genetic algorithms to constrain terrestrial ecosystem models of the global carbon cycle with multiple ecological data (Barrett, 2002); and discriminating vegetation and soil sources and sinks in forest canopies with concentration, isotopic and physiological data (Styles *et al.*, 2002).

A fourth example deserves more space than is available here: the use of data assimilation in atmospheric and ocean circulation models. This is now

well-developed and applied routinely in weather forecasting. A variety of techniques are employed, including ‘nudging’, three-dimensional and four-dimensional variational data assimilation (3DVAR and 4DVAR) based on adjoint methods, and use of the ensemble Kalman filter. Recent reviews are provided by Chen & Lamb (2000) and Park & Zupanski (2003).

Data characteristics: uncertainty in measurement and representation

We have emphasized that data uncertainties affect not only the predicted uncertainty of the eventual result of a model–data synthesis process, but also the predicted best estimate. This realization raises the challenge of evaluating the uncertainty properties of the main kinds of observation relevant to a TCOS, in forms directly usable for model–data synthesis. This is a very large goal, which embraces all categories of observation identified at the beginning of ‘Model–data synthesis: Methods’, and also a range of issues:

- The error magnitude σ_m for an observation z_m , inclusive of all error sources (in other words, the diagonal elements $[\text{Cov } \mathbf{z}]_{mm} = \sigma_m^2$ of the covariance matrix);
- The correlations $[\text{Cov } \mathbf{z}]_{mn}/(\sigma_m\sigma_n)$ among errors in different observations, quantified by the off-diagonal elements of the covariance matrix;
- The temporal structure of the errors: whether they are random in time or temporally correlated, and the possible presence of unknown long-term drifts or biases;
- The spatial structure of errors (random, slowly varying or bias as for temporal structure);
- The error distribution: normal (Gaussian), log-normal, skewed or the sum of multiple error sources with different distributions, such as a small Gaussian noise together with occasional large outliers because of measurement corruption events;
- Possible mismatches between the spatial and temporal averaging implicit in the model and the measurements (the ‘scaling problem’);
- The separate contributions to all the above error properties of measurement error (the distribution of the measurements \mathbf{z} around their true values) and representation error (the distribution of the error in the model representation of the measurement, $\mathbf{z} = \mathbf{h}(\mathbf{y})$).

This challenge is too large to meet fully here. To make a start, we consider (in the next four subsections) a

selection of observations from four categories of data: remote sensing of land surface properties, atmospheric composition measurements, direct flux measurements, and direct measurements of carbon stores. The aim is to make estimates of error properties for these categories of measurement. The discussion does not address all of the above issues, largely omitting questions of spatial and temporal error structure. We present tables indicating ranges for the diagonal elements $[\text{Cov } \mathbf{z}]_{mm} = \sigma_m^2$ of the error covariance matrix for measurement error, and the qualitative behaviour of the correlations which determine the off-diagonal elements. The entries in these tables are mostly ‘expert judgements’ by the authors and their colleagues, backed up by quantitative evidence where possible. There is, of course, no claim that our estimates are definitive; the intention is rather to indicate the kinds of uncertainty information required of observations for model–data synthesis purposes. The tables characterize measurement errors only; representation errors, which often exceed measurement errors, are discussed separately in qualitative terms only. The issue of scale mismatches between measurements and models, which arises in all cases as a significant contribution to representation error, is treated generically in a fifth subsection ‘Scale mismatches between measurements and models’.

Remote sensing of land surface properties

The main satellite-borne remotely sensed data on land surface properties come from two kinds of sensor, both polar-orbiting to provide frequent global coverage: moderate-spatial-resolution (~ 250 – 1000 m) and high-temporal-resolution (~ 1 day repeat interval) sensors such as AVHRR and MODIS; and high-spatial-resolution (~ 10 – 30 m) and moderate-temporal-resolution (~ 16 day repeat interval) sensors such as SPOT and LANDSAT. All these sensors provide multi-year records. One major application (among many) for the AVHRR-MODIS family is assessment of vegetation dynamics with indices such as NDVI (defined as $(\text{NIR} - \text{Red})/(\text{NIR} + \text{Red})$, where NIR and Red are radiances in the near-infrared and visible red spectral bands) and measures such as surface temperature. Applications for the SPOT-LANDSAT family include detection of land cover change and vegetation clearing and regrowth. In all cases, the measurements are at-sensor reflected radiances from the earth in several spectral bands (5 for AVHRR, 37 for MODIS).

In using these forms of remote sensing data for model–data synthesis applications, three kinds of error need to be considered: (1) errors associated with the measurement and spatial attribution of radiances at the sensor; (2) errors in relating radiances at sensor to

radiance at surface and (3) errors in relating radiances at the surface to biophysical quantities represented in a terrestrial biosphere model. The first of these is measurement error, the third is representation error, and the second can contribute to either depending on how the problem is formulated.

Measurement error. The primary measurements are radiances at the sensor, attributed to elements (pixels) on the earth's surface. Measurement errors arise from sensor noise, calibration drift, orbital decay, and incorrect geolocation. These errors are much more serious for AVHRR (designed in the 1970s) than for MODIS (designed in the 1990s with on-board calibration).

Representation error. This is the error associated with the model $\mathbf{z} = \mathbf{h}(\mathbf{y})$ relating the measurements \mathbf{z} (radiance) to the target variables \mathbf{y} (biophysical variables in the terrestrial biosphere model, such as leaf chlorophyll content, leaf or soil water status or leaf area index). In principle, the observation model involves two components. First, biophysical target variables (\mathbf{y}) are related to radiances at the earth's surface (say $\mathbf{z}_{\text{surface}}$), through a model $\mathbf{z}_{\text{surface}} = \mathbf{h}_{\text{surface}}(\mathbf{y})$. Examples are relationships between NDVI and leaf and soil properties (Tucker, 1979; Sellers, 1985; Sellers *et al.*, 1992; Myneni *et al.*, 1995a,b; Lu *et al.*, 2003). Representation errors associated with this component depend on the skill of the model $\mathbf{z}_{\text{surface}} = \mathbf{h}_{\text{surface}}(\mathbf{y})$. Second, the surface radiative properties are related to radiances at the satellite-borne sensor (say $\mathbf{z}_{\text{sensor}}$) through an atmospheric radiative transfer model (\mathbf{g}), which accounts for the effects of clouds, atmospheric absorption and scattering, and the bidirectional reflectance distribution function (BRDF) properties of the surface. Thus, we have $\mathbf{z}_{\text{sensor}} = \mathbf{g}(\mathbf{z}_{\text{surface}}, \text{ancillary data})$, where the ancillary data includes profiles of temperature and radiatively active constituents in the atmosphere. Errors in this component depend on the skill of the model \mathbf{g} and errors in the required ancillary data.

In model–data synthesis, there is a choice about whether to use sensor radiances ($\mathbf{z}_{\text{sensor}}$) or surface radiances ($\mathbf{z}_{\text{surface}}$) as the primary measurements. If $\mathbf{z}_{\text{surface}}$ is used, then the observation model is $\mathbf{z}_{\text{surface}} = \mathbf{h}_{\text{surface}}(\mathbf{y})$ and it is necessary to infer $\mathbf{z}_{\text{surface}} = \mathbf{g}^{-1}(\mathbf{z}_{\text{sensor}}, \text{ancillary data})$ from the at-sensor radiances and the atmospheric transfer model \mathbf{g} . In this case, \mathbf{g} is effectively part of the measurement system itself and its errors appear as measurement errors. On the other hand, if $\mathbf{z}_{\text{sensor}}$ is the primary set of measurements, then the observation model becomes $\mathbf{z}_{\text{sensor}} = \mathbf{h}_{\text{sensor}}(\mathbf{y}) = \mathbf{g}(\mathbf{h}_{\text{surface}}(\mathbf{y}), \text{ancillary data})$. The

model \mathbf{g} is then part of the observation model, and its errors appear as representation errors. New unknown target variables may also appear through the ancillary data, and it is necessary that these be estimated with available observations.

For a semi-quantitative illustration of properties of the error covariance matrix for some remote sensing observations, we consider radiances at the surface ($\mathbf{z}_{\text{surface}}$) to be the observations, so that errors in the radiative transfer model \mathbf{g} form part of the errors in $\mathbf{z}_{\text{surface}}$ and are treated as measurement error. As examples we consider NDVI and surface temperature. Table 1 shows some properties of measurement errors in these quantities. This table takes account of several considerations: first, errors in $\mathbf{z}_{\text{surface}}$ arise from sensor noise, sensor calibration and geolocation (contributing to error in $\mathbf{z}_{\text{sensor}}$), and from inaccurate cloud removal, atmospheric correction or BRDF correction (contributing to error in the radiative transfer model \mathbf{g}). Many of these errors are strongly positively correlated among different spectral bands, especially in the visible and NIR. NDVI is less sensitive to such errors than the band radiances themselves, because it is based on a normalized difference (an example of error cancellation through correlation, and a reason for the popularity of NDVI). Second, correlations between errors in NDVI and surface temperature tend to be positive. For example, unmasked sub-pixel cloud appears as low-NDVI and cold, causing correlated negative outliers and thus a positive error correlation between the two measurements. Finally, in some cases, experimental estimates of the errors are available. For example, comparison of several calibration procedures for the short-wave AVHRR channels on the NOAA-11 satellite revealed calibration discrepancies of around 5% (Mitchell *et al.*, 1996).

Table 1 describes measurement error only, and does not include representation error reflecting uncertainty in the relationship $\mathbf{z}_{\text{surface}} = \mathbf{h}_{\text{surface}}(\mathbf{y})$ (the third of the three error categories mentioned above). Estimates of representation error can be gained, for example, from the scatter in experimental tests of relationships between at-surface radiance properties and biophysical variables (for example, Lu *et al.*, 2003). Usually, these representation errors are comparable with or larger than measurement errors. For example, NDVI saturates at high leaf area index (3–4).

Atmospheric composition measurements

We consider (1) direct CO₂ concentration measurements from the global flask network and WMO Global Atmosphere Watch (http://www.wmo.ch/web/arep/gaw/gaw_home.html) stations with continuous CO₂ monitoring and (2) efforts to measure CO₂ with

Table 1 Indicative properties of the error covariance matrix for remote sensing observations of NDVI and surface temperature, describing measurement error only and omitting representation error

Observation (z_m)	Units	Typical range	Typical error (σ_m)	Error distribution	Error correlations
NDVI = (NIR–Red)/(NIR + Red)	–	0.1–1	0.1 (AVHRR) 0.05 (MODIS)	Normal with negative outliers because of undetected cloud	Errors in NDVI and surface temperature are probably highly positively correlated because of negative outliers in each associated with undetected cloud
Surface temperature	degree K	250–350	Land: 2 (AVHRR) 1 (MODIS) Ocean: 0.2 (AVHRR) 0.1 (MODIS)		

NDVI, normalized difference vegetation index.

spaceborne remote sensing. Our discussion of measurement and representation errors for these data sources is based on the atmospheric inversion approach for utilizing atmospheric composition data; see ‘Model–data synthesis: methods and examples’.

Direct measurements. The first global flask network developed from work by Keeling (1961) with several nations later commencing sampling in their respective regions. As of 2004, 23 laboratories from 15 nations contribute flask and continuous *in situ* sampling data from around 200 sites to international databases (<http://www.cmdl.noaa.gov/ccgg/globalview/co2/>). Large networks are maintained by the National Oceanic and Atmosphere Administration Climate Monitoring and Diagnostics Laboratory (NOAA-CMDL), Boulder, CO, USA, and CSIRO Atmospheric Research (CSIRO-GasLab), Melbourne, Australia. In the main, flask samples are collected 2–4 times per month from fixed sites, aircraft and ships, and opportunistically during intensive experimental campaigns. Flasks are returned to the central laboratories for automated analyses, which provide concentrations of a number of atmospheric constituents. The analytical procedure includes measurements against an international standard and frequent comparisons with air standards, and actual samples, exchanged between laboratories (Masarie *et al.*, 2001).

Continuous *in situ* analysers provide records at higher precisions and temporal densities than are available from the flask networks and are usually located at baseline stations in remote locations. With both flask and *in situ* monitoring, there has been ongoing improvement in precision and in the accuracy and propagation of the international standards (Francey *et al.*, 2001).

Spaceborne measurements. Remote sensing of atmospheric composition is gradually becoming a

reality. Simulations show that satellite observations improve atmospheric-inverse estimates of carbon fluxes by a factor of up to ten relative to the surface (flask and baseline) networks, because of vastly improved coverage in time and space, albeit with increased error for each measurement (Rayner & O’Brien, 2001; Rayner *et al.*, 2002). Efforts are underway on three fronts: first, proof-of-concept studies have already been undertaken with existing space-based radiometers such as TOVS and AVHRR (Chedin *et al.*, 2002, 2003a,b). Despite limited spectral resolution and unwanted absorption from other atmospheric constituents, signatures of seasonal cycles and trends in CO₂ and other greenhouse gases (N₂O and CO) have been extracted at coarse space-time resolution (15° × 15°, mid-troposphere, monthly). Second, current missions undertaken for other purposes, such as the Advanced Infra-red Sounder (AIRS) and Sciamachy instruments, will likely provide near-term improvements in measuring aspects of the atmospheric CO₂ distribution. Finally, future purpose-built instruments, such as the NASA Orbiting Carbon Observatory (OCO, Crisp *et al.*, 2004), CARBOSAT (European Space Agency) and GOSAT (Japan), should provide dramatic improvements in coverage and precision.

Measurement error. Table 2 gives estimates of measurement error for CO₂ measurements from flasks, continuous *in situ* analysers, and AVHRR (Chedin, 2003b). Errors in international CO₂ databases from near-surface measurements include errors in the assignment and propagation of CO₂-in-air standards on the WMO mole fraction scale, biases associated with different CO₂ measurement methods, flask storage effects, and other factors limiting the repeatability of measurements with the same system. These errors have different temporal characteristics. For example, both flask and continuous measurements from CSIRO share a current calibration bias of -0.1 ± 0.05 ppm

relative to NOAA-CMDL. For some laboratories with smaller networks, WMO round-robin intercomparisons in 1993, 1997 and 2001 indicate a larger, but improving range of calibration uncertainty. There are also differences between flask and *in situ* measurements from within one laboratory that exhibit slowly varying offsets. Such errors are likely to be highly correlated over long time periods. If not included in the inversion procedure they can invalidate aspects of the result, as demonstrated by Rayner *et al.* (2002) in their Observing System Simulation Experiment of remotely-sensed CO₂. Provided the structure of such errors can be predicted, the inversion can cope with them (Law *et al.*, 2003a,b).

Representation error. Current atmospheric inversion studies using atmospheric concentration data generally ascribe much larger errors to the measurements than the aggregate of the measurement errors shown in Table 2. This is partly because of the simple assumption of Gaussian uncorrelated noise used in most existing studies, but also because the contributions of the representation error are large. Representation errors arise in this context from the inability of the atmospheric transport model to simulate point observations in space, either because of systematic errors in model formulation or the implicit averaging in its grid representation. An analogous temporal representation error arises when flask measurements (actually grab samples in time) are interpreted as longer-term means; see 'Scale mismatches between measurements and models' for discussion. A further contribution to representation errors for most atmospheric inversion studies to date has been the projection of possible source distributions to a restricted subspace, usually by dividing the earth into a number of large regions. This is done both for computational reasons and to reduce the error amplification arising from under-determined problems. Errors in the prescription of flux distributions within these regions give rise to a so-called aggregation error, described and quantified by Kaminski *et al.* (2001). This error can be avoided by using adjoint representations of atmospheric transport that do not require aggregation (Rodenbeck *et al.*, 2003a,b).

There are few experiments where representation errors can be evaluated, since this requires simultaneous knowledge of sources and atmospheric transport. However, one can use the range of model simulations as a guide (e.g. Law *et al.*, 1996; Gurney *et al.*, 2003). Since representation errors are completely dependent on the inversion process (especially the atmospheric transport model) rather than the measurements themselves, we do not attempt to quantify them in Table 2.

Table 2 Indicative properties of the error covariance matrix for observations of atmospheric CO₂ concentrations, describing measurement error only and omitting representation error

Observation (z_m)	Units	Typical range	Typical error (σ_m)	Error distribution	Error correlations
[CO ₂] from flask	ppmv	350–500	0.01–1	Normal with slowly varying bias Normal?	Measurement errors: correlated offsets which vary over months to years Uncorrelated between different observations in space and time?
[CO ₂] from baseline station	ppmv	350–500	0.05 (CSIRO-GasLab, NOAA-CMDL) 0.5 (all laboratories)		
[CO ₂] from AVHRR (Chedin <i>et al.</i> 2003b) (15° × 15°, mid-troposphere)	ppmv	360–400	4		

Queries indicate speculation.

Direct flux measurements at intensive study sites

The global FluxNet network (Baldocchi *et al.*, 2001, Falge *et al.*, 2002a,b) includes over 200 sites where tower-based eddy covariance measurements are made of the land–air fluxes of sensible heat, latent heat, CO₂ and other entities, at sub-diurnal (typically half-hourly) temporal resolution. Many other meteorological variables are measured, including solar and net radiation, ground heat flux and precipitation. Many sites include measurements of other biogeochemical processes, such as soil respiration by chamber methods. Most measurements at these sites are at the patch scale, applying to a nominally homogeneous land unit. The extent of horizontal averaging varies with the measurement (for example, a narrow along-wind ellipse of order $1 \times 0.1 \text{ km}^2$ for eddy-covariance fluxes, a circle with radius of order 10 m for a net radiometer mounted 10 m above a forest, and a chamber footprint of order 1 m^2 for soil respiration). A great strength of intensive flux study sites is that independent checks on the uncertainty of the measurements are possible through checks on the closure of the energy and water balances.

Measurement error. Table 3 shows some estimated error properties for the above measurements. Several systematic errors are known to influence eddy-flux measurements, including high-frequency and low-frequency flux losses, storage in the air column below the measurement height, intermittency of nocturnal turbulence, and ‘rectification’ because of nocturnal cold-air drainage of air rich in respired CO₂ towards low-lying areas, causing systematic advection errors (Wofsy *et al.*, 1993). Most of these problems are much more severe at night than by day. Experimental techniques and analysis procedures (especially the use of longer averaging periods of up to several hours to include fluxes transported by large eddies of the scale of the entire atmospheric boundary layer) are progressively resolving these issues, with more success to date for daytime than nocturnal fluxes (Finnigan *et al.*, 2003). Relative hourly accuracies of order 10% represent the state of the art for daytime fluxes, attainable with excellent technique; however, errors of up to 30% are common. For nocturnal fluxes, eddy covariance measurements become unreliable in light wind conditions, that is, much of the time at night (Goulden *et al.*, 1996). Errors include biases associated mainly with the above systematic nocturnal effects.

Carbon flux measurements from intensive study sites are often aggregated to time-averaged (typically monthly to multi-annual) measures of primary production for the patch under study: $\text{GPP} = [\text{net assimilation}]$; $\text{net primary productivity (NPP)} =$

$[\text{GPP} - \text{autotrophic respiration}]$; $\text{net ecosystem productivity (NEP)} = [\text{NPP} - \text{heterotrophic respiration}]$; $\text{net biome productivity (NBP)} = [\text{NEP} - \text{disturbance flux}]$. All these have units $\text{gC m}^{-2} \text{yr}^{-1}$ and are positive for carbon uptake into the biosphere. The total exchange between the terrestrial biosphere and the atmosphere is NBP. The disturbance fluxes in NBP include grazing, harvest, and catastrophic events (fire, windthrow, clearing). If these processes do not occur on an intensive study site, an eddy-covariance CO₂ flux (when aggregated in time) yields an estimate of NEP. Table 3 also shows estimated properties of error covariances for these time-aggregated productivity estimates. The uncertainty in NEP estimates is much higher than for hourly CO₂ flux estimates, because of the difficulties mentioned above. Comparisons between eddy-covariance and other means of measuring NEP suggest typical uncertainties of order 20 to 50%, with the lower end of this range being attainable over actively growing ecosystems where NPP is much larger than heterotrophic respiration and NEP is relatively large. This range is also likely to be representative of GPP and NPP estimates. Uncertainties in NBP at patch scale are very much higher (in fact it is arguable that NBP cannot be defined at patch scale).

Representation error. Fluxes are almost always directly represented in terrestrial biosphere models, so representation errors are not an issue for direct flux measurements at the patch scale. However, the spatial aggregation issue (upsampling flux measurements from intensive study sites to yield estimates of fluxes or productivities over grid cells, large regions or continents) is a significant source of representation error. It is not usually possible to upscale by area weighting of patch-scale estimates without conditioning from other measurements and models, because of heterogeneity in landscapes and disturbance patterns; see ‘Scale mismatches between measurements and models’ for further discussion. A particular example of this issue is that patch-scale measurements (including fluxes) are often made at unmanaged or undisturbed sites, creating a bias with respect to the landscape as a whole.

Measurements of carbon stores in vegetation and soils

Terrestrial biospheric carbon pools include leaves, wood above and below ground, fine roots, coarse litter, fine litter and soil carbon; the soil carbon is often partitioned into fractions with different biochemical properties or turnover times, such as microbial, humic and inert. *In situ* measurements or estimates of these pools are available from vegetation and soil surveys,

Table 3 Indicative properties of the error covariance matrix for some direct flux measurements, describing measurement error only and omitting representation error

Observation (z_m)	Units	Typical range	Typical relative error (σ_m/z_m) (%)	Error distribution	Error correlations
Hourly net irradiance	$W m^{-2}$	0–600	5–15	Normal	Errors in radiation, eddy covariance and soil respiration measurements on hourly time scales are uncorrelated to first approximation. However, errors in hourly eddy covariance measurements of sensible heat, latent heat and CO_2 fluxes are highly correlated.
Hourly sensible heat flux (from eddy covariance)	$W m^{-2}$	0–600	10–30	Normal	
Hourly latent heat flux (from eddy covariance)	$W m^{-2}$	0–600	10–30	Normal	
Hourly CO_2 flux (from eddy covariance)	$gC m^{-2} d^{-1}$	0–10 (+ down)	20–50	Normal with nocturnal bias	Errors in annual estimates of GPP, NPP, NEP and NBP are positively correlated because of sequential estimation (GPP \rightarrow NPP \rightarrow NEP \rightarrow NBP)
Hourly soil-respired CO_2 flux	$gC m^{-2} d^{-1}$	0–5 (+ up)	20–50	Normal	
Annual GPP	$gC m^{-2} yr^{-1}$	0–4000 (+ down)	20–50	Normal with bias	
Annual NPP	$gC m^{-2} yr^{-1}$	0 to 2000 (+ down)	20–50	Normal with bias	
Annual NEE	$gC m^{-2} yr^{-1}$	–500–1500 (+ down)	20–50	Normal with bias	
Long-term average NBP	$gC m^{-2} yr^{-1}$	–100–100 (+ down)	$\gg 100$	Normal with bias	

GPP, gross primary production; NPP, net primary productivity; NEE, net ecosystem exchange; NBP, net biome productivity.

Table 4 Indicative properties of the error covariance matrix for some measurements of carbon stores in vegetation and soils, describing measurement error only and omitting representation error

Observation (\mathbf{z}_m)	Units	Typical range	Typical relative error (σ_m/\mathbf{z}_m) (%)	Error distribution	Error correlations
Leaf carbon	kgC m^{-2}	0–1	10–30	Log-normal (normal if many measurements are aggregated)	At an individual site, errors are uncorrelated because measurements are made with different techniques
Wood carbon (above and below ground)	kgC m^{-2}	0–50	15–50		
Fine root carbon	kgC m^{-2}	0–1	30–100		
Coarse litter (including standing dead)	kgC m^{-2}	0–10	30–100		
Fine litter	kgC m^{-2}	0–0.5	10–30		
Soil carbon (to 1 m)	kgC m^{-2}	0 to > 100	30–100		

long-term observations at ecological study sites, and national forest inventories.

Measurement errors. Table 4 shows some estimated error properties for these measurements. Large errors are associated with the below-ground pools. As for our error estimates for flux measurements at intensive study sites (Table 3), these estimates apply to measurement errors at the spatial scale of individual patches.

Representation errors. At first sight, representation errors for measurements of carbon stores are not a major issue at patch scales because carbon stores are directly represented in terrestrial biosphere models (as for flux measurements). However, representation errors arise in assigning measurements of carbon pools to their model counterparts (Barrett, 2002). For instance, litter is often measured in physically defined compartments such as fine litter and coarse woody debris, but represented in models in biogeochemical compartments such as metabolic and structural litter. Spatial aggregation of measurements for upscaling and incorporation into spatially coarse models is also a representation issue, as for direct flux measurements—for example, the bias introduced by the tendency for patch-scale measurements to be made at undisturbed sites.

Scale mismatches between measurements and models

The examples in the above four subsections all highlight the problem that measurements are almost always made at a different scale (that is, with a different spatial and temporal averaging operator) to that used in the dynamic model. In its spatial form, this problem arises when the measurements apply to particular points or homogeneous land elements (for example, eddy flux measurements or *in situ* measurements of carbon stores) while the model represents an aggregated set

of land elements by a single set of state variables (for example, for a grid cell in a coarse spatial model). A temporal version of the same problem arises when the measurements are intermittent (for example, satellite measurements at time of overpass, flask sampling of atmospheric composition, or occasional *in situ* measurements of carbon stores).

This ‘scaling problem’ is so prevalent that it is worth a brief discussion in the context of model–data synthesis. Suppose that the model is defined at ‘coarse’ space-time scales by physical parameterizations that demand some (usually implicit) space-time averaging of all variables, and by choices of space-time grids, which are compatible with these averaging requirements. Since it is necessary that all variables in both the dynamic and observation models be averaged in a consistent way throughout, the observation model can be written as $\mathbf{z}_{\text{coarse}} = \mathbf{h}(\mathbf{y}_{\text{coarse}}) + \text{noise}$, where the measurements (\mathbf{z}) and target variables (\mathbf{y}) are identified as applying at coarse (model) scale. However, the actual measurements are often available at some much finer scale \mathbf{z}_{fine} , and may also be sparsely sampled. The variability in an ensemble of measurements \mathbf{z}_{fine} (which cover the coarse scale) is defined by a covariance \mathbf{R}_{fine} , the integral of the cospectrum of \mathbf{z} between the fine and coarse scales. The problem is to find a relationship between \mathbf{z}_{fine} and $\mathbf{z}_{\text{coarse}}$ so that \mathbf{z}_{fine} can be used as an observable in the model–data synthesis process. There are several generic ways to do this.

A first, very simple option is to take the fine scale measurements (\mathbf{z}_{fine}) as a noisy sample of $\mathbf{z}_{\text{coarse}}$, with the variability in \mathbf{z}_{fine} treated as a contribution to the representation error. In this case, the observation model becomes $\mathbf{z}_{\text{fine}} = \mathbf{h}(\mathbf{y}_{\text{coarse}}) + [\text{measurement error in } \mathbf{z}_{\text{fine}}] + [\text{representation error in } \mathbf{h}] + [\text{noise with covariance } \mathbf{R}_{\text{fine}}]$, where the last noise term accounts for unresolved space-time variability in \mathbf{z}_{fine} . This is often the best option if there is no other information

available about the sources of variability in \mathbf{z}_{fine} , as for example with flask measurements of atmospheric composition or soil samples taken from imprecisely known locations.

A second option is that the fine scale measurements (\mathbf{z}_{fine}) can be aggregated directly to find $\mathbf{z}_{\text{coarse}} = \Sigma w(\mathbf{a}_{\text{fine}})\mathbf{z}_{\text{fine}}$, where w is a set of weights determined by ancillary fine-scale data \mathbf{a}_{fine} . For example, reflected radiances can be aggregated with area weighting, and *in situ* soil or biomass carbon measurements can be aggregated approximately using a weighting based on soil and microclimate attributes that can be inferred from fine-scale digital elevation maps. The inferred $\mathbf{z}_{\text{coarse}}$ becomes a pseudo-observation to be utilized through an observation model $\mathbf{z}_{\text{coarse}} = \mathbf{h}(\mathbf{y}_{\text{coarse}}) + \text{noise}$, where the noise term includes errors in \mathbf{z}_{fine} , errors in the weighting model for $\mathbf{z}_{\text{coarse}}$, and errors in \mathbf{a}_{fine} . This is a good option when plenty of fine-scale measurements are available within each coarse-scale element, for example from remote sensing.

A third option is to introduce a fine-scale process or empirical model $\mathbf{z}_{\text{fine}} = \mathbf{g}(\mathbf{x}_{\text{coarse}}, \mathbf{a}_{\text{fine}})$ which relates the fine-scale observations to coarse-scale state variables (which here act as boundary conditions) and additional fine-scale ancillary data such as topography and land surface attributes from remote sensing. The model \mathbf{g} then becomes the observation model for model–data synthesis. This is often the best option if the fine-scale data are sparse, expensive and of high information content. For example, for eddy flux measurements, \mathbf{g} may be a fine-scale model of land–air exchanges and microclimates in inhomogeneous or hilly terrain, or a neural-network-based model of flux distributions conditioned with available ancillary data \mathbf{a}_{fine} . It is likely that the development of \mathbf{g} involves substantial effort, so this is not an option to be taken lightly.

This brief discussion has focussed on the observational issue of scale mismatches between measurements and models, and their implications for the observation model $\mathbf{z} = \mathbf{h}(\mathbf{y}) + \text{noise}$, Eqn (2). Another form of the ‘scaling problem’ arises for the dynamic model $\mathrm{d}\mathbf{x}/\mathrm{d}t = \mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{p}) + \text{noise}$, Eqn (1), in translating process knowledge about the model function \mathbf{f} between scales. Typically \mathbf{f} is a set of phenomenological equations for fluxes contributing to changes in stores \mathbf{x} , which are valid only for certain (often implicit) spatial and temporal resolutions. For example, the Darcy law describes water movement in soil columns but not catchments or heterogeneous regions described by aggregated state variables. Translation of these phenomenological equations from fine to coarse scales is possible by treating the fine-scale variability statistically, but one result is that the fine-scale and coarse-scale equations are different (for instance, biased with

respect to each other) because of interactions between fine-scale variability and nonlinearity in the fine-scale function $\mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{p})$.

Summary and conclusions

In the context of terrestrial carbon observation, we have focussed on model–data synthesis and its implications for data, especially the specification of data uncertainty. Our analysis has been framed by an initial statement of the purposes and attributes of a TCOS. The purposes for a TCOS are congruent with those of a ‘Global Earth Observation System of Systems’, with specific contributions occurring in the areas of weather and climate prediction, water resource management, ecosystem management, agricultural sustainability, combating desertification and monitoring biodiversity. Among the major attributes of a TCOS are scientific credibility, consistency with global budgets, adequate spatial and temporal resolution, observation of sufficient ranges of entities and processes, and the requirement to quantify uncertainty. These attributes demand a model–data synthesis approach because of the need to combine a range of observations and models to determine the terrestrial stores and fluxes of carbon and related entities (water, nutrients, energy), and the ways that they are influenced by human management.

Data for model–data synthesis approaches come in two forms, observations and prior knowledge (for instance constraints on model parameters). For both, uncertainty estimates have an influence on the outcome of the synthesis process comparable with that of the data values themselves. Data uncertainties affect not only the predicted uncertainty of the eventual result, but also the predicted best estimate. Therefore, there is an urgent need for soundly based uncertainty specification, based initially on an error covariance matrix.

In this paper, we have made semi-quantitative estimates of some of the main properties of the covariance matrix for measurement error, for four kinds of data central to terrestrial carbon observation: remote sensing of land surface properties, atmospheric composition measurements, direct flux measurements, and measurements of carbon stores. Critical error properties include (1) the diagonal elements $[\text{Cov } \mathbf{z}]_{mm} = \sigma_m^2$ of the measurement error covariance matrix (where σ_m is the error magnitude for an observation z_m); (2) the correlations between different observations, quantified by the off-diagonal elements of the covariance matrix; (3) the temporal and (4) the spatial structure of errors, (5) the error distribution, (6) possible scale mismatches between measurements and models and (7) the representation of the observations in the model.

Finally, we note that there is a need not only to quantify uncertainty, but also to reduce uncertainty in our estimates. The approach outlined here contributes to this goal, by providing a framework to formalize the manner in which we constrain uncertainty. The critical step is to better understand the error structures of both the priors and the observations, leading to improved focus on the major sources of uncertainty.

Acknowledgments

This paper is the result of a workshop held in Sheffield, UK, 3–6 June 2003, organized jointly by the Terrestrial Carbon Observation (TCO) initiative of the Global Terrestrial Observing System (GTOS), and the Global Carbon Project (GCP; www.globalcarbonproject.org). We thank GTOS/TCO, NASA, IGBP, and the GCP Canberra Office (funded by Australian Greenhouse Science Program) for financial and organizational support, and all workshop participants for their contributions and enthusiasm. We thank several colleagues, particularly John Norton, John Finnigan and Roger Francey, for comments on a draft of the paper. This paper contributes to the implementation of the GCP research agenda on 'Development of model–data fusion techniques' (Activity 1.2), as GCP Publication No.8.

References

- Annan JD, Hargreaves JC (2004) Efficient parameter estimation for a highly chaotic system. *Tellus*, **56A**, 520–526.
- Baldocchi D, Falge E, Gu LH *et al.* (2001) FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, **82**, 2415–2434.
- Barrett DJ (2002) Steady state turnover time of carbon in the Australian terrestrial biosphere. *Global Biogeochemical Cycles*, **16**, 1108, doi:10.1029/2002GB001860.
- Chedin A, Serrar S, Armante R *et al.* (2002) Signatures of annual and seasonal variations of CO₂ and other greenhouse gases from comparisons between NOAA TOVS observations and radiation model simulations. *Journal of Climate*, **15**, 95–116.
- Chedin A, Saunders R, Hollingsworth A *et al.* (2003a) The feasibility of monitoring CO₂ from high-resolution infrared sounders. *Journal of Geophysical Research-Atmospheres*, **108**, 4064, doi:10.1029/2001JD001443.
- Chedin A, Serrar S, Scott NA *et al.* (2003b) First global measurement of midtropospheric CO₂ from NOAA polar satellites: TROPICAL zone. *Journal of Geophysical Research-Atmospheres*, **108**, 4581, doi:10.1029/2003JD003439.
- Chen CR, Lamb PJ (2000) Improved treatment of surface evapotranspiration in a mesoscale numerical model part II: via the assimilation of satellite measurements. *Terrestrial Atmospheric and Oceanic Sciences*, **11**, 789–832.
- Ciais P, Meijer HAJ (1998) The 18O/16O isotope ratio of atmospheric CO₂ and its role in global carbon cycle research. In: *Stable Isotopes: Integration of Biological, Ecological and Geochemical Processes* (ed. Griffiths H), pp. 409–431. Bios Scientific Publishers Ltd., Oxford.
- Ciais P, Moore BI, Steffen WL *et al.* (2004) *Integrated Global Carbon Observation Theme: A Strategy to Realise a Coordinated System of Integrated Global Carbon Cycle Observations*. Integrated Global Observing Strategy, Stockholm.
- Cihlar J, Denning S, eds (2002b) *Terrestrial Carbon Observation: The Frascati Report on In Situ Carbon Data and Information*. Environment and Natural Resources Series 5, Food and Agriculture Organization, Rome.
- Cihlar J, Denning S, eds (2002c) *Terrestrial Carbon Observation: The Rio de Janeiro Recommendations for Terrestrial And Atmospheric Measurements*. Environment and Natural Resources Series 3, Food and Agriculture Organization, Rome.
- Cihlar J, Denning AS, Gosz J, eds (2002a) *Terrestrial Carbon Observation: The Ottawa assessment of requirements, status and next steps*. Environment and Natural Resources Series 2, Food and Agriculture Organization, Rome.
- Crisp D, Atlas RM, Breon F-M *et al.* (2004) *The orbiting carbon observatory (OCO) mission*. http://oco.jpl.nasa.gov/pubs/OCO_Cospar_2002_final.pdf
- Drécourt J-P (2003) *Kalman filtering in hydrological modelling*. DAIHM Technical Report 2003-1, DHI Water & Environment, Hørsholm, Denmark.
- Enting IG (1999a) *Characterising the temporal validity of the global carbon cycle*. Technical Paper No. 40, CSIRO Atmospheric Research, Melbourne, Australia.
- Enting IG (1999b) Green's function methods of tracer inversions. In: *Inverse Methods in Global Biogeochemical Cycles* (ed. Kasibhatla P), American Geophysical Union. Washington, DC. pp. 19–31.
- Enting IG (2002) *Inverse Problems in Atmospheric Constituent Transport*. Cambridge University Press, Cambridge.
- Enting IG, Trudinger CM, Francey RJ (1995) A synthesis inversion of the concentration and delta-C-13 of atmospheric CO₂. *Tellus Series B-Chemical and Physical Meteorology*, **47**, 35–52.
- Evans SN, Stark PB (2002) Inverse problems as statistics. *Inverse Problems*, **18**, R55–R97.
- Evensen G (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. *Journal of Geophysical Research-Oceans*, **99**, 10143–10162.
- Evensen G (2003) The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, **53**, 343–367.
- Falge E, Baldocchi D, Tenhunen J *et al.* (2002a) Seasonality of ecosystem respiration and gross primary production as derived from FLUXNET measurements. *Agricultural and Forest Meteorology*, **113**, 53–74.
- Falge E, Tenhunen J, Baldocchi D *et al.* (2002b) Phase and amplitude of ecosystem carbon release and uptake potentials as derived from FLUXNET measurements. *Agricultural and Forest Meteorology*, **113**, 75–95.
- Field CB, Raupach MR, eds (2004) *The Global Carbon Cycle: Integrating Humans, Climate, and the Natural World*. Island Press, Washington.
- Finnigan JJ, Clement R, Malhi Y *et al.* (2003) A re-evaluation of long-term flux measurement techniques—Part I: averaging and coordinate rotation. *Boundary-Layer Meteorology*, **107**, 1–48.

- Francey RJ, Rayner PJ, Allison CE (2001) Constraining the global carbon budget from global to regional scales—the measurement challenge. In: *Global Biogeochemical Cycles in the Climate System* (eds Schulze E-D, Heimann M, Harrison S, Holland E, Lloyd J, Prentice IC, Schimel DS), pp. 245–252. Academic Press, San Diego.
- Gershenfeld NA (1999) *The Nature of Mathematical Modeling*. Cambridge University Press, Cambridge.
- Giering R (2000) Tangent linear and adjoint biogeochemical models. In: *Inverse Methods in Global Biogeochemical Cycles* (eds Kasibhatla P, Heimann M, Rayner P, Mahowald N, Prinn RG, Hartley DE), pp. 33–48. American Geophysical Union, Washington, DC.
- Global Carbon Project (2003) *Science Framework and Implementation*, Earth System Science Partnership (IGBP, IHDP, WCRP, Diversitas) Report No. 1; GCP Report No. 1, Global Carbon Project, Canberra.
- Gloor M, Bakwin P, Hurst D *et al.* (2001) What is the concentration footprint of a tall tower? *Journal of Geophysical Research—Atmospheres*, **106**, 17831–17840.
- Goulden ML, Munger JW, Fan SM *et al.* (1996) Exchange of carbon dioxide by a deciduous forest: response to interannual climate variability. *Science*, **271**, 1576–1578.
- Grewal MS, Andrews AP (1993) *Kalman Filtering: Theory and Practice*. Prentice-Hall, Englewood Cliffs, NJ.
- Gurney KR, Law RM, Denning AS *et al.* (2002) Towards robust regional estimates of CO₂ sources and sinks using atmospheric transport models. *Nature*, **415**, 626–630.
- Gurney KR, Law RM, Denning AS (2003) TransCom 3 CO₂ inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information. *Tellus Series B—Chemical and Physical Meteorology*, **55**, 555–579.
- Kalman RE (1960) A new approach to linear filtering and prediction problems. *ASME Journal of Basic Engineering, Series D*, **82**, 34–45.
- Kaminski T, Knorr W, Rayner PJ *et al.* (2002) Assimilating atmospheric data into a terrestrial biosphere model: a case study of the seasonal cycle. *Global Biogeochemical Cycles*, **16**, doi:10.1029/2001GB001463.
- Kaminski T, Rayner PJ, Heimann M *et al.* (2001) On aggregation errors in atmospheric transport inversions. *Journal of Geophysical Research—Atmospheres*, **106**, 4703–4715.
- Kasibhatla P, Heimann M, Rayner PJ *et al.*, eds (2000) *Inverse Methods in Global Biogeochemical Cycles*. American Geophysical Union, Washington DC.
- Keeling CD (1961) The concentration and isotopic abundance of carbon dioxide in rural and marine air. *Geochimica et Cosmochimica Acta*, **24**, 277–298.
- Law RM, Chen YH, Gurney KR (2003a) TransCom 3 CO₂ inversion intercomparison: 2. Sensitivity of annual mean results to data choices. *Tellus Series B—Chemical and Physical Meteorology*, **55**, 580–595.
- Law RM, Rayner PJ, Denning AS *et al.* (1996) Variations in modeled atmospheric transport of carbon dioxide and the consequences for CO₂ inversions. *Global Biogeochemical Cycles*, **10**, 783–796.
- Law RM, Rayner PJ, Steele LP *et al.* (2003b) Data and modelling requirements for CO₂ inversions using high-frequency data. *Tellus Series B—Chemical and Physical Meteorology*, **55**, 512–521.
- Le Dimet F-X, Talagrand O (1986) Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus Series A—Dynamic Meteorology and Oceanography*, **38**, 97–110.
- Lloyd J, Kruijt B, Hollinger DY *et al.* (1996) Vegetation effects on the isotopic composition of atmospheric CO₂ at local and regional scales: theoretical aspects and a comparison between rain forest in Amazonia and a Boreal Forest in Siberia. *Australian Journal of Plant Physiology*, **23**, 371–399.
- Lu H, Raupach MR, McVicar TR *et al.* (2003) Decomposition of vegetation cover into woody and herbaceous components using AVHRR NDVI time series. *Remote Sensing of Environment*, **86**, 1–18.
- Masarie KA, Langenfelds RL, Allison CE *et al.* (2001) NOAA/CSIRO flask air intercomparison experiment: A strategy for directly assessing consistency among atmospheric measurements made by independent laboratories. *Journal of Geophysical Research—Atmospheres*, **106**, 20445–20464.
- Mitchell RM, O'Brien DM, Forgan BW (1996) Calibration of the AVHRR shortwave channels. 2. Application to NOAA 11 during early 1991. *Remote Sensing of Environment*, **55**, 139–152.
- Myneni RB, Hall FG, Sellers PJ *et al.* (1995a) The interpretation of spectral vegetation indexes. *IEEE Transactions on Geoscience and Remote Sensing*, **33**, 481–486.
- Myneni RB, Maggion S, Iaquinto J *et al.* (1995b) Optical remote-sensing of vegetation—modeling, caveats, and algorithms. *Remote Sensing of Environment*, **51**, 169–188.
- Park SK, Zupanski A (2003) Four-dimensional variational data assimilation for mesoscale and storm-scale applications. *Meteorology and Atmospheric Physics*, **82**, 173–208.
- Press WH, Teukolsky SA, Vetterling WT *et al.* (1992) *Numerical Recipes in Fortran 77: The Art of Scientific Computing*. Cambridge University Press, Cambridge, UK.
- Raupach MR (2001) Inferring biogeochemical sources and sinks from atmospheric concentrations: general considerations and applications in vegetation canopies. In: *Global Biogeochemical Cycles in the Climate System* (eds Schulze E-D, Heimann M, Harrison S, Holland E, Lloyd J, Prentice IC, Schimel DS), pp. 41–60. Academic Press, San Diego.
- Raupach MR, Canadell JG, Bakker DCE *et al.* (2004) Interactions between CO₂ stabilization pathways and requirements for a sustainable earth system. In: *The Global Carbon Cycle: Integrating Humans, Climate, and the Natural World* (eds Field CB, Raupach MR), pp. 131–162. Island Press, Washington.
- Rayner PJ (2001) Atmospheric perspectives on the ocean carbon cycle. In: *Global Biogeochemical Cycles in the Climate System* (eds Schulze E-D, Heimann M, Harrison S, Holland E, Lloyd J, Prentice IC, Schimel DS), pp. 285–294. Academic Press, San Diego.
- Rayner PJ, Enting IG, Francey RJ *et al.* (1999) Reconstructing the recent carbon cycle from atmospheric CO₂, delta C-13 and O₂/N₂ observations. *Tellus Series B—Chemical and Physical Meteorology*, **51**, 213–232.

- Rayner PJ, Law RM, O'Brien DM *et al.* (2002) Global observations of the carbon budget-3. Initial assessment of the impact of satellite orbit, scan geometry, and cloud on measuring CO₂ from space. *Journal of Geophysical Research–Atmospheres*, **107**, 4557, doi:10.1029/2001JD000618.
- Rayner PJ, O'Brien DM (2001) The utility of remotely sensed CO₂ concentration data in surface source inversions. *Geophysical Research Letters*, **28**, 175–178.
- Reichle RH, McLaughlin DB, Entekhabi D (2002) Hydrologic data assimilation with the ensemble Kalman filter. *Monthly Weather Review*, **130**, 103–114.
- Rodenbeck C, Houweling S, Gloor M *et al.* (2003a) CO₂ flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport. *Atmospheric Chemistry and Physics*, **3**, 1919–1964.
- Rodenbeck C, Houweling S, Gloor M *et al.* (2003b) Time-dependent atmospheric CO₂ inversions based on interannually varying tracer transport. *Tellus Series B–Chemical and Physical Meteorology*, **55**, 488–497.
- Running SW, Baldocchi DD, Turner DP *et al.* (1999) A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modeling and EOS satellite data. *Remote Sensing of the Environment*, **70**, 108–127.
- Schimel DS, House JI, Hibbard KA *et al.* (2001) Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems. *Nature*, **414**, 169–172.
- Sellers PJ (1985) Canopy reflectance, photosynthesis and transpiration. *International Journal of Remote Sensing*, **6**, 1335–1372.
- Sellers PJ, Berry JA, Collatz GJ *et al.* (1992) Canopy reflectance, photosynthesis, and transpiration 3: a reanalysis using improved leaf models and a new canopy integration scheme. *Remote Sensing of the Environment*, **42**, 187–216.
- Styles JM, Raupach MR, Farquhar GD *et al.* (2002) Soil and canopy CO₂, (CO₂)-C-13, H₂O and sensible heat flux partitions in a forest canopy inferred from concentration measurements. *Tellus Series B–Chemical and Physical Meteorology*, **54**, 655–676.
- Tarantola T (1987) *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*. Elsevier, Amsterdam.
- Todling R (2000) Estimation theory and atmospheric data assimilation. In: *Inverse Methods in Global Biogeochemical Cycles* (eds Kasibhatla P, Heimann M, Rayner P, Mahowald N, Prinn RG, Hartley DE), pp. 49–66. American Geophysical Union, Washington, DC.
- Tucker CJ (1979) Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, **8**, 127–150.
- Wang YP, Barrett DJ (2003) Estimating regional terrestrial carbon fluxes for the Australian continent using a multiple-constraint approach I. Using remotely sensed data and ecological observations of net primary production. *Tellus Series B–Chemical and Physical Meteorology*, **55**, 270–289.
- Wang YP, Bentley ST (2002) Development of a spatially explicit inventory of methane emissions from Australia and its verification using atmospheric concentration data. *Atmospheric Environment*, **36**, 4965–4975.
- Wang YP, McGregor JL (2003) Estimating regional terrestrial carbon fluxes for the Australian continent using a multiple-constraint approach II. The Atmospheric constraint. *Tellus Series B–Chemical and Physical Meteorology*, **55**, 290–304.
- Wofsy SC, Goulden ML, Munger JW *et al.* (1993) Net exchange of CO₂ in a midlatitude forest. *Science*, **260**, 1314–1317.

Other Key Reference Papers for Dr. Dennis's Lecture:

- [1] Del Grosso SJ, Parton WJ, Mosier AR, Holland EA, Pendall E, Schimel DS, Ojima DS. 2005. Modeling soil CO₂ emissions from ecosystems. *Biogeochemistry* 73: 71-91
- [2] Janssen, PHM and PSC Heuberger. 1995. Calibration of process-oriented models. *Ecological Modeling* 83: 55-66

PART III: DATA ASSIMILATION METHODS

With theoretical and application perspectives, this part expatiates on data assimilation frameworks and presents the latest algorithms, including cost-function based methods and filtering-based methods. Their advantages and limitations are shown. An adaptive Kalman filter technique is proposed, and the uncertainty problems are further discussed. A powerful software tool on land data assimilation system will be demonstrated through a special lecture.

Assimilation of Observations, an Introduction

By Olivier Talagrand

Laboratoire de Météorologie Dynamique du CNRS, Paris, France

(Manuscript received 13 November 1995, in revised form 11 March 1997)

Abstract

Assimilation of meteorological or oceanographical observations can be described as the process through which all the available information is used in order to estimate as accurately as possible the state of the atmospheric or oceanic flow. The available information essentially consists of the *observations* proper, and of the *physical laws* which govern the evolution of the flow. The latter are available in practice under the form of a *numerical model*. The existing assimilation algorithms can be described as either *sequential* or *variational*. The links between these algorithms and the *theory of statistical estimation* are discussed. The performances of present algorithms, and the perspectives for future development, are also briefly discussed.

1. Introduction

Implementation of Numerical Weather Prediction (NWP) obviously requires the specification of appropriate initial conditions. In the early stages of Numerical Weather Prediction, forty years ago, observations were synchronous in time, and bore on the same physical variables (geopotential, horizontal winds) as those used in the prediction models for describing the state of the atmospheric flow. Long before the availability of computers, meteorologists had been performing what was called the *analysis* of the meteorological situation, which consisted in correcting every day, with the new observations, the forecast from the previous day, available under the form of meteorological maps. With the advent of NWP, this task was devoted to the computer. Two-dimensional interpolation algorithms were defined by Bergthorsson and Döös (1955) and Cressman (1959). These algorithms followed an approach which has remained essentially unchanged to this day. *Background* fields defined at the grid-points of the forecasting model are interpolated to the observation locations. The differences between observations and interpolated values are then interpolated back to grid-points in order to define *corrections* to be applied to the first-guess. Eliassen (1954) and Gandin (1963), giving a statistical basis to the analysis process, defined what has become known in meteorology as *optimal interpolation*, built on the statistical covariance functions of the meteorological fields. After many improvements, and under many variants, optimal interpolation is still to

this day at the heart of most procedures for defining initial conditions of numerical weather forecasts.

But specific problems also appeared very soon. It was observed that, if appropriate precautions were not taken, the early stages of forecasts produced by non-filtered models, which do not impose an *a priori*, time independent, relationship between the mass and velocity fields, exhibited unrealistic high-frequency motions, early identified as gravity wave oscillations (Hinkelmann, 1951). This led to the practice of *initialization*, performed after the analysis proper, and intended at producing “balanced” initial conditions, which would not result in unrealistic oscillations. It was also recognized that the requirement for balanced initial conditions could on occasions be in contradiction with the requirement that the initial conditions be close to the available observations. At about the same time, the celebrated work of Lorenz (1963) showed that the atmosphere possesses the remarkable property of extremely high *sensitivity to initial conditions*. This, in addition to marking the birth of the new discipline of *deterministic dissipative chaos*, showed that deterministic prediction of the atmospheric circulation was ineluctably limited in time. Meteorologists thus learnt that all the efforts they could make were in a sense doomed to failure, but also received an additional incentive for defining as accurately as possible the initial conditions of numerical weather forecasts. In the late sixties, the development of satellite observing systems, and the perspective that asynoptic observations, performed more or less continuously in time, would become more and more numerous in the future, led to the notion that the dynamical evo-



lution of the flow should be explicitly taken into account in the very definition of the initial conditions of the forecast. The word *assimilation* was coined at that time for denoting a process in which observations distributed in time are merged together with a dynamical numerical model of the flow in order to determine as accurately as possible the state of the atmosphere.

Since then, continuous progress in theory, in efficiency of numerical algorithms, as well as in available computing power, has led to a slow but steady progress in the methods for assimilation. This progress, together with improvements in the quality of the NWP models themselves (and, to a lesser extent, with improvements in the observing system), has significantly contributed to the continuous increase observed in the last decades in the quality of numerical weather forecasts. It is worth mentioning that the proportion of resources allocated to assimilation in the whole process of NWP has steadily increased over time. At the beginning of NWP, the computational cost of analysis was negligible in comparison to the cost of a 24-hour forecast. Now, the cost of the computations required by the assimilation (in addition to the cost of integrating the model over the assimilation period) is typically the cost of a 24-hour forecast. And major meteorological centres are considering to allocate daily, for a 24-hour assimilation, the equivalent of a ten-day forecast or more. This evolution has not resulted from a clearly stated voluntary choice. With hindsight, it would be more appropriately described as a progressive “natural selection” process, during which increase of the proportion of resources allocated to assimilation repeatedly and consistently proved to be beneficial.

If there is no doubt in the minds of knowledgeable people that improvements in assimilation methods have significantly contributed over the years to the improvements in the quality of numerical weather forecasts, purely objective proofs of that fact are not readily obtained, since all components of the whole NWP process have been improving simultaneously. An extremely interesting and instructive by-product of assimilation has been presented by Salstein and Rosen (1986) (see also, Oort, 1989). These authors have compared the rate of rotation of the Earth, as estimated from geodetic measurements, with the angular momentum of the atmosphere with respect to the Earth's axis of rotation, as obtained from the analyses produced by the US National Meteorological Center. After subtraction of well identified tidal components from the observed fluctuations of the rate of rotation, the latter and the atmospheric angular momentum are correlated, over periods of up to a few years, to a remarkable high degree of accuracy. This shows that non-tidal short-term fluctuations of the rotation of the Earth are essentially due to exchanges of angular momentum between

the atmosphere and the solid Earth (and that, at least over short ranges, oceans play only a minor role). Another extremely interesting and instructive by-product of assimilation has been the identification, by several authors (see, e.g., Vautard, 1990), of *weather regimes* in which the atmospheric circulation stabilizes over periods that can last for as long as a few weeks.

These examples show that, if assimilation of observations originated from the needs of NWP, and if the latter is still to this day the main incentive for research and improvement in that domain, assimilation has already proven to be useful for other purposes than weather prediction. Several meteorological centres are engaged in the *reassimilation* of past observations with present computing means and assimilation algorithms. These reassimilation projects, which bear on periods as long as several decades, will produce an homogeneous description of the atmospheric circulation over long periods of time (or at least as homogeneous a description as allowed by the evolving, even if slowly evolving, observing system). The long sequences of reassimilated states thus obtained will be extremely useful for climatological studies of many kinds. This is particularly important at the present time of strong concern about the possible impact of human activities on climate, making it particularly desirable to detect early and reliably any possible climate change. One can mention a major difference between assimilation intended at defining the initial conditions of a numerical weather forecast, and a *a posteriori* reassimilation of past observations. In the former case, one can of course use only observations performed before, or at the latest at, the time at which one wants to estimate the state of the flow. In the case of reassimilation, there is no reason to ignore observations performed after estimation time, and it is certainly desirable to use algorithms that are capable, in a way or another, to carry the information contained in the observations both forward and backward in time.

Assimilation of observations is also rapidly developing in the field of *dynamical oceanography*. The present state of development of numerical modelling of the oceanic circulation is described in this volume by Anderson, and the present oceanographical observing system, together with expected future developments, by Busalacchi. One basic difficulty is that the quantity of available observations is, relatively speaking, much smaller for oceanography than for meteorology (Ghil and Malanotte-Rizzoli, 1991, taking into account the appropriate characteristic spatial and temporal scales, have estimated that the temporal density of oceanographic observations has so far been four orders of magnitude less than the density of atmospheric observations). This makes validation of oceanic models particularly diffi-

cult, but also makes it particularly desirable to take the best possible advantage of the available observations. Assimilation of oceanographic observations is rather different from assimilation of atmospheric observations in that the primary purpose of assimilation of oceanographic observations is not (at least up to now) to define the initial conditions of a forecast, but more modestly to produce a reasonable description of the state of the oceanic circulation. The interest for assimilation of oceanographic observations has been strongly stimulated by the development, either already effective or anticipated, of new observing systems, in particular of satellite altimetry.

(Still another domain in which there exists growing interest for assimilation of observations is modelling of the *biosphere*. Numerical modelling of the oceanic or terrestrial biosphere, and of its interactions with the atmosphere or the oceans, is rapidly progressing. The full exploitation of models of the biosphere will require appropriate assimilation of the various observations, in particular satellite observations, bearing on the biosphere.)

Stated in general terms, the purpose of assimilation can be described as follows: using all the available information, determine as accurately as possible the state of the atmospheric or oceanic flow. Depending on one's particular eventual goal, one may wish to determine the state of the flow at a given time, or alternatively the history of the flow over a period of time. As for the available information, it consists first of the *observations proper*. As described in the contributions by Atlas and Busalacchi in this volume, the observations vary significantly in nature and accuracy, and have a highly irregular temporal and spatial distribution. In particular, the observations can be "direct", in that they bear on the same physical quantities to be used in the desired description of the flow (typically, velocity, temperature, plus humidity for the atmosphere or salinity for the ocean). Or they can be "indirect", *i.e.* bearing on quantities that are more or less "complicated" functions (usually some forms of integrals) of the quantities chosen for describing the flow. Satellite altimetric measurements of the ocean surface, and satellite measurements of the infra-red thermal flux emitted by the atmosphere, are examples of indirect measurements. Another example is provided by acoustic tomographic measurements performed in the ocean.

The second source of information to be used in the assimilation consists of the dynamical model, and more generally of the *physical laws* governing the flow. These physical laws are fundamentally the principles of conservation of mass, energy and momentum, and a numerical model is nothing else than a numerically usable (and approximate) state-

ment of these principles. In addition to the contribution by Anderson, relative to oceanic modelling, the present state of development of numerical models of the atmospheric circulation is described in this volume by Arakawa. It is clear that it must be possible to acquire some knowledge of the meteorological or oceanographical fields from appropriate combination of the observations and of the physical laws governing the flow. For instance, the time derivative of the surface pressure is, under the hydrostatic approximation, the vertical integral of the divergence of the horizontal wind. Observations of the surface pressure performed at a given point at successive times therefore contain information on the wind field. To mention another simple example, information on the wind field can also be obtained from the observed motion of tracers, such as humidity.

One could think that the observations on the one hand, and the physical laws that govern the flow on the other, together make up all the appropriate information, and that there is no need to add anything to these two basic sources of information. This is certainly true in principle, but the practical situation is somewhat different. It may be useful for instance to explicitly introduce in the assimilation climatological estimates of at least some of the quantities to be determined, even if it is known that climatological quantities are in the last instance determined by the physical laws governing the flow (and by the energy input to the system). A less obvious, but practically much more important example is given by *geostrophic balance*. It is known that, at least in middle latitudes, the atmospheric and oceanic flows are in approximate geostrophic balance. It has already been mentioned that starting a numerical weather forecast from initial conditions which are not in appropriate balance will result in the presence in the forecast, at least for some time, of unrealistic high frequency motions. Geostrophic balance must be a necessary consequence of the physical laws governing the flow. But it is only an asymptotic property, in the sense that any solution of the relevant equations will asymptotically tend to approximate geostrophic balance, but need not be in geostrophic balance at the initial time. And it is indeed observed that numerical models, provided they contain the nonlinearities associated with fluid advection and a reasonable form of dissipation (plus a reasonable form of energy forcing if they are to be integrated over long periods) produce solutions which tend to approximate geostrophic balance, even though the initial conditions may have been non-geostrophic. But this property is not sufficient by itself for ensuring geostrophic balance in the fields produced by assimilation of meteorological observations over a period of, say, 24 hours. Experience shows that it is necessary to explicitly introduce in the assimilation process the information that the

atmospheric flow is in approximate geostrophic balance.

Now, no information will ever be exact, whether it comes from observations, physical laws (especially if these physical laws are expressed in the form of a discretized numerical model) or from some other form of knowledge. And thus, there always will be some uncertainty on the result of the assimilation, originating from the uncertainty on the various sources of information used in the assimilation. An ideal assimilation system should therefore produce not only an estimate of the state of the flow, but also an estimate of the *associated uncertainty*.

A first difficulty one encounters in assimilation of meteorological observations is simply the numerical dimension of the problem. The number of individual scalar meteorological observations performed over a 24-hour period is at present typically on the order of 10^5 . As for the dimension of the largest NWP models (*i.e.* the number of independent parameters defining in the models the state of the flow at a given time), it is now in the range $10^6 - 10^7$. NWP models typically require one hour of elapsed computer time for 24 hours of simulated time. Assimilation of meteorological observations is most often performed over periods of 24 hours, which is of course the natural thing to do for meteorological services issuing forecasts on a daily basis. In addition to at least one 24-hour integration, the assimilation algorithms which are at present considered as most efficient essentially require the solution of one or several linear systems of equations whose dimension is either the number of observations or the dimension of the model. Fitting the corresponding computational load within the narrow limits of operational NWP imposes very strong constraints on assimilation. These constraints are critical for the choice and implementation of assimilation algorithms. Another difficulty arises from the nonlinear (actually chaotic) character of the atmospheric and oceanic flows. This imposes strong limits on assimilation, just as it imposes limits on predictability. But it must be mentioned that most of the work done so far on assimilation has in effect been done within the bounds of some appropriate local linear approximation (which will be described in some detail below), so that the full effects of nonlinearity have not been incorporated yet in assimilation.

(The goal of determining as accurately as possible the state of the atmospheric or oceanic flow, together with the associated uncertainty, may seem very ambitious.) We shall here attempt to convince the reader that it can be achieved, at least to a reasonable degree of precision, by solving an appropriate *generalized least-squares minimization problem*. A basic reference on assimilation of meteorological observations is a book by Daley (1991), which gives

a comprehensive description of existing assimilation methods. Bennett (1992), in the general context of assimilation of oceanographic observations, describes mathematical techniques, with emphasis on possibilities for the future. More specific aspects related to spatial interpolation have been studied in detail by Thiébaux and Pedder (1987) and Wahba (1990).

Before we come to the more technical aspects of assimilation of meteorological and oceanographical observations, it is worth mentioning that similar problems are encountered in many fields of science and engineering. Navigation of aircraft and spacecraft of various kinds, in which one wants to know at any time, as accurately as possible, the position and velocity of a vehicle "observed" through various instruments, is a form of assimilation of observations. All forms of "signal filtering" are also essentially of the same nature, and one basic tool of estimation theory, Kalman filtering, which is used in assimilation of meteorological and oceanographical observations, originated in electrical engineering. Many "inverse problems" also present similarities with assimilation of observations. In plasma physics, one often wants to know the internal state of a physical system from observations performed at its surface (and also to control the system through action at its surface). Solid Earth geophysics is another example. Most of what is known on the internal structure of the Earth comes from inversion of signals (mostly seismic signals) observed at its surface. All of these examples lead to estimation problems, in which one wants to infer the state of a physical system from information which may be extremely heterogeneous in origin, nature and accuracy, and which may be related only very "indirectly" to the quantities to be estimated. The interesting fact which we want to stress here is that, in spite of the diversity of the physical systems under consideration, the methods used for solving these different problems are basically very similar, even though they often have been developed independently. They are all stated, or at least can be stated, in a probabilistic setting, and aim in effect at determining some reasonable approximation to the *conditional probability distribution* function of the state of the system, given the available information (see, *e.g.*, Tarantola, 1987, or Lorenc, 1986). The numerical algorithms are often very similar, and are fundamentally independent of the "equations" governing the physical system under consideration. These equations are in effect data, introduced in the estimation process in a way that is not basically different from (and is in some cases identical with) the way the observations proper are themselves introduced. These basic similarities are extremely instructive in that they show that estimation procedures are essentially independent of the

particular properties of the physical system under consideration. But this does not mean of course that significant differences do not exist. In particular, assimilation of meteorological and oceanographical observations seems to be unique in its extremely large numerical dimensions.

2. Sequential and variational assimilation

From a purely algorithmic point of view (and independently of the underlying theory), assimilation exists at present under two forms, *sequential assimilation* and *variational assimilation*. In sequential assimilation, which is the only form to have been used so far in operational NWP, the assimilating model is integrated over the time interval over which the observations to be used are distributed. Whenever the model time reaches an instant at which observations are available, the state predicted by the model is used as a background which is “updated”, or “corrected”, with the new observations. The integration of the model is then restarted from the updated state, and the process is repeated until all the available observations have been used. In operational NWP, the state obtained at the end of the assimilation period is taken as the initial state for the ensuing forecast. As already mentioned, the operation which consists in correcting a background at a given time with new observations is called an analysis. Sequential assimilation is therefore an alternative sequence of *analyses performed at observation times*, and of *integrations of the model* between successive analyses. In all algorithms of sequential assimilation that have been developed so far, there is only one sweep of the model over the assimilation period, so that each individual observation is used once and only once.

One appealing feature of sequential assimilation is the constant updating it performs on the state predicted by the model: each new piece of observation is used for correcting the latest estimate of the state of the atmospheric flow. This feature makes sequential assimilation well adapted to NWP. Numerical forecasts are normally produced at the rate of one forecast every day. In order to define the initial conditions of a new forecast at time t_0 , it is very natural, starting from the initial conditions of the previous forecast at time $t_0 - 24$ hr, to perform a 24-hr sequential assimilation. This approach is implemented operationally in numerous NWP centres, and has now been running continuously in some of them for more than ten years without interruption (but with constant improvement of both the model and the assimilation algorithm itself).

However, sequential assimilation also possesses a serious drawback: because precisely of the sequential character of the assimilation, each individual piece of observation influences the estimated state of the flow only at later times, and not at previous

times. There is propagation of the information contained in the observations only from the past into the future, and not from the future into the past. As said in the Introduction, this is of no importance in the case of weather prediction, where one necessarily must at one stage run the model into the future. But it certainly is a disadvantage in the case of a *posteriori* reassimilation of past observations, where it seems preferable to use algorithms capable of carrying information both forward and backward in time.

Variational assimilation, on the other hand, aims at globally adjusting a model solution to all the observations available over the assimilation period.

The adjustment being simultaneous, the adjusted states at all times are influenced by all the observations over the assimilation period, thereby avoiding the difficulty mentioned above. In presently existing algorithms for variational assimilation, one first defines a scalar function which, for any model solution over the assimilation interval, measures the “distance”, or “misfit”, between that solution and the available observations. That so-called *objective function* (or *cost function*) will typically be a sum of squared differences between the observations and the corresponding model values, *e.g.*

$$J \equiv \sum_j \alpha_j (y_j - y_j^o)^2 \quad (2.1)$$

where the y_j^o 's are the observations, the y_j 's are the corresponding model values, and the α_j 's are numerical weights reflecting the accuracy of the various observations. One will then look for the model solution that minimizes the objective function. Since a model solution is uniquely defined by the corresponding initial conditions at the beginning of the assimilation period, these initial conditions are taken as *control variables*, *i.e.* as the variables with respect to which the minimization is effectively performed. The minimizing initial state is obtained through an iterative procedure, each step of which requires the explicit knowledge of the local values of the set of partial derivatives, or *gradient vector*, of the objective function with respect to the initial state. As will be explained below, this gradient can be determined, at a non-prohibitive numerical cost, through use of the *adjoint equations* of the assimilating model.

3. Least-squares statistical linear estimation. Generalities

Most (but not all) assimilation algorithms that have been used so far, either for research or for operational purposes, and either of the sequential or of the variational type, can be described as more or less simplified forms of *least-squares statistical linear estimation*. Many “nonlinear” applications correspond in fact to cases which are very close, in some sense, to linearity. Least-squares statistical linear

estimation is a classical tool, whose basic principles are very simple, even though practical implementation on large dimension systems can raise many problems. We will first describe the basic principles of least-squares statistical linear estimation on elementary examples, before going to present meteorological and oceanographical applications.

Let us consider the following simple estimation problem. We want to determine some unknown scalar quantity x^t from two known measurements z_1 and z_2 of the form

$$z_1 = x^t + \zeta_1 \quad (3.1a)$$

$$z_2 = x^t + \zeta_2 \quad (3.1b)$$

where ζ_1 and ζ_2 are “observational” errors. These errors are of course unknown, but we assume that the statistical performances of the instruments which have produced z_1 and z_2 are known. More precisely, we assume that these instruments are unbiased, *i.e.*

$$E(\zeta_1) = E(\zeta_2) = 0 \quad (3.2a)$$

where $E(\cdot)$ denotes the statistical mean, and that the statistical variances of ζ_1 and ζ_2 are known

$$E(\zeta_1^2) = \sigma_1^2 \quad E(\zeta_2^2) = \sigma_2^2 \quad (3.2b)$$

We assume in addition, for the sake of simplicity, that the observation errors are uncorrelated

$$E(\zeta_1 \zeta_2) = 0 \quad (3.2c)$$

This will be the case if, for instance, the two observations have been obtained with different instruments.

We now want to estimate x^t as a linear combination of the two observations z_1 and z_2 , *viz.*

$$x^a = a_1 z_1 + a_2 z_2 \quad (3.3)$$

where the weights a_1 and a_2 are to be determined. We first want the estimate x^a to be statistically unbiased, *i.e.* to verify the condition $E(x^a - x^t) = 0$. This will be verified if

$$a_1 + a_2 = 1 \quad (3.4)$$

We also want x^a , among all unbiased estimates, to minimize the statistical variance of the estimation error, *viz.*

$$\sigma^2 = E[(x^a - x^t)^2] \quad (3.5)$$

The solution to this simple constrained minimization problem (minimize 3.5 under constraint 3.4) is easily found to correspond to weights a_1 and a_2 which are inversely proportional to the variances of the corresponding observation errors, *i.e.*

$$a_1 = \sigma_2^2 / (\sigma_1^2 + \sigma_2^2) \quad a_2 = \sigma_1^2 / (\sigma_1^2 + \sigma_2^2) \quad (3.6)$$

In addition, the corresponding minimum of the estimation error variance σ^2 is given by the relationship

$$1/\sigma^2 = 1/\sigma_1^2 + 1/\sigma_2^2 \quad (3.7)$$

which has a simple interpretation : if one calls “precision” the inverse of an error variance, then the precision of the estimate x^a is the sum of the precisions of the observations.

The same estimate x^a can be found through a different approach : an acceptable estimate of the exact value x^t must be close to the observations, at least within the accuracy of the latter. For any value x , the “distance” between x and the observations can be measured by the following quadratic quantity

$$J(x) \equiv (x - z_1)^2 / \sigma_1^2 + (x - z_2)^2 / \sigma_2^2 \quad (3.8)$$

where the observational error variances σ_1^2 and σ_2^2 account for the accuracy of the observations. Now, the value of x which minimizes $J(x)$ is precisely the estimate x^a given by eqs (3.3–6). Minimization of (3.8) therefore provides another way of determining the estimate x^a , based on a (very elementary) variational principle.

Formulæ (3.3–6) generalize to the case of any number m of observations $z_j = x^t + \zeta_j$ ($j = 1, \dots, m$). The generalization is obvious if the observation errors are uncorrelated. It is slightly less obvious, but still elementary, when the errors are correlated. But it is more realistic to consider observations which do not necessarily bear on a quantity to be estimated. Observations are rarely performed at the times and spatial locations at which estimates are sought. In addition, as already mentioned, many observations are “indirect”, and do not bear on the physical quantities to be estimated. For instance, satellite-borne radiometers measure the radiative flux emitted by the Earth to outer space at different wavelengths, while what one basically wants are estimates of the atmospheric temperature and humidity fields. The measured fluxes are functions of these fields (and of other quantities, such as cloud amount and top level pressure, and surface emissivity) through the *radiative transfer equation*. Indirect measurements will become more and more numerous in the future, and assimilation methods must allow for measurements that are “complicated” functions of the physical parameters to be estimated. Statistical linear estimation can accommodate such indirect measurements (of course within the limits of linearity, which will be discussed later) as we will now proceed to show.

We assume that what we now want to estimate is a complete vector x^t , with dimension n and components x_i^t ($i = 1, \dots, n$). That vector can be thought of as consisting for instance of the values of one or several meteorological fields (temperature, wind components, humidity) at a given instant at the points of a two- or three-dimensional regular array. But the developments that follow are very general, and independent of the physical nature or significance of

the parameters to be estimated. The vector \mathbf{x}^t to be estimated will be called the *state vector*, since it will in general describe the state of a physical system, such as the atmosphere or the ocean. As for the available observations, they make up a vector \mathbf{z} , with dimension m and components $z_j (j = 1, \dots, m)$. We assume that the vector \mathbf{z} can be written under the form

$$\mathbf{z} = \Gamma \mathbf{x}^t + \zeta \quad (3.9)$$

where Γ is an $m \times n$ matrix which defines the link between the parameters to be estimated and the observed quantities. Γ will be called the *observation matrix*. If the observations bear on the physical field to be estimated, but are performed at points in space-time different from the points at which estimates are sought, Γ will represent some appropriate space-time interpolation. If the observations bear on “indirect” functions of the parameters to be estimated, Γ will represent an appropriate linearization of the physical and/or statistical relationship linking \mathbf{x}^t and \mathbf{z} . In the simple example (3.1) (where $n = 1$, $m = 2$), Γ is the matrix $(1 \ 1)^T$ (where the superscript T denotes transposition). As for the m -vector ζ in (3.9), its components are the errors affecting the observations. This error vector is of course unknown, but we will assume that it is statistically unbiased, *i.e.*

$$E(\zeta) = 0$$

and that the variances-covariances of its components, making up the matrix $E(\zeta \zeta^T)$, are known

$$E(\zeta \zeta^T) = \Sigma$$

To sum up, the exactly known quantities in eq. (3.9) are the observation vector \mathbf{z} and the observation matrix Γ , while the observation error vector is known only through its statistical properties, and the vector \mathbf{x}^t to be estimated is totally unknown. In order to estimate \mathbf{x}^t , we now proceed as in the example (3.1) above, and look for an estimate \mathbf{x}^a which is a linear function of \mathbf{z} , *i.e.* is of the form

$$\mathbf{x}^a = \mathbf{A} \mathbf{z} \quad (3.10)$$

where \mathbf{A} is an $n \times m$ matrix to be determined. We want the estimate \mathbf{x}^a to be unbiased, *i.e.* to be such that $E(\mathbf{x}^a - \mathbf{x}^t) = 0$. This condition is verified if

$$\mathbf{A} \Gamma = \mathbf{I}_n \quad (3.11)$$

where \mathbf{I}_n is the unit matrix of order n . In addition, among all the matrices verifying (3.11), we want to choose the one that minimizes the variance of the norm of the estimation error, *i.e.* the matrix that minimizes the trace (sum of the diagonal terms) of the covariance matrix $\mathbf{P}^a = E[(\mathbf{x}^a - \mathbf{x}^t)(\mathbf{x}^a - \mathbf{x}^t)^T]$ of the estimation error. The solution to that problem is

$$\mathbf{A} = [\Gamma^T \Sigma^{-1} \Gamma]^{-1} \Gamma^T \Sigma^{-1} \quad (3.12)$$

As for the corresponding matrix \mathbf{P}^a , it is equal to

$$\mathbf{P}^a = [\Gamma^T \Sigma^{-1} \Gamma]^{-1} \quad (3.13)$$

\mathbf{P}^a contains the variances and covariances of the estimation errors on all the components of \mathbf{x}^t . In particular, its diagonal terms are the variances of these estimation errors.

Formulæ (3.12–13) generalize formulæ (3.6–7). As in that previous example, there is a variational formulation to the estimation of \mathbf{x}^t , which now corresponds to minimizing the objective function

$$J(\mathbf{x}) \equiv [\Gamma \mathbf{x} - \mathbf{z}]^T \Sigma^{-1} [\Gamma \mathbf{x} - \mathbf{z}] \quad (3.14)$$

where \mathbf{x} is an n -vector. The meaning of expression (3.14), which generalizes (3.8), should be clear: for any \mathbf{x} , the vector $\Gamma \mathbf{x} - \mathbf{z}$ is the difference vector between what would be observed if the vector to be estimated was equal to \mathbf{x} (and if the observations were perfect) and the actual observations \mathbf{z} . $J(\mathbf{x})$ is the squared norm of that difference vector, weighted so as to take into account, through Σ^{-1} , the accuracies of the different observations (and also the possible correlations between the various observation errors). The estimate \mathbf{x}^a is the value of \mathbf{x} for which that squared norm is minimum, *i.e.* the value of \mathbf{x} that would produce, if exactly observed, the values closest to the actual observations.

It must be stressed that the problem of minimizing the statistical variance of the estimation error on \mathbf{x}^t , and of minimizing the objective function (3.14), are *a priori* distinctly different problems, even though the fact that they lead to identical results is algebraically obvious in simple cases. In the first problem, one minimizes a quantity defined on the n -dimensional space of *state vectors*, while in the second problem one minimizes a quantity defined on the m -dimensional space of *observations*. It is certainly not *a priori* obvious that the solutions to these two problems should always be identical.

The vector \mathbf{x}^a defined by Eqs. (3.10–12) is called the *Best Linear Unbiased Estimate*, or BLUE, of \mathbf{x}^t from \mathbf{z} . The theory leading to the BLUE is standard, and has been described here in order to stress its generality, and also as an introduction to the various methods for assimilation which, although they often do not explicitly refer to the theory of statistical linear estimation, can almost always be described, as we have already said, as more or less simplified applications of that theory. In addition to the estimate \mathbf{x}^a , statistical linear estimation produces the covariance matrix \mathbf{P}^a of the corresponding estimation error, thus fulfilling the goal assigned in the Introduction to an ideal assimilation algorithm. It can be noted that the matrix \mathbf{P}^a does not depend

on the observation vector \mathbf{z} , but only on the observation operator Γ and on the matrix Σ , *i.e.* on the nature and accuracy of the observations. In particular, the theory of statistical linear estimation can be used for evaluating the performance of an hypothetical observing system, defined by what that system would observe and with which accuracy, but independently of any actual observations.

Implementation of statistical linear estimation requires, in addition to the knowledge of the matrices Γ and Σ , the fact that the observation error ζ be unbiased. Actually, if that error was biased, but the bias was known, it would be sufficient to first subtract the bias from the observation vector in order to obtain a new, unbiased, observation vector. The requirement that the bias is zero is therefore in effect only a requirement that the statistical mean of the observation error is known.

To sum up the results obtained so far: before one can implement statistical linear estimation, one must know what has been observed, in terms of the parameters to be estimated (this is expressed by the observation matrix), and with which accuracy (this is expressed by the mean and the covariance matrix of the observation error vector). These requirements may seem extremely demanding, since there will always be instruments, especially newly developed instruments, for which it will certainly be very difficult to assign reliable values to the observational errors. But at the same time, it is obvious that a prerequisite for a rational use of a set of observations is to know what has been observed, and with which accuracy. It is therefore a good thing that the requirement for knowledge of the nature and accuracy of the observations comes out of the mathematics of estimation theory. If the required information is not available, one will then have to compensate for it by as reasonable as possible hypotheses on the observation matrix Γ and the corresponding error covariance matrix Σ . One advantage of studying assimilation in the perspective of general estimation theory is that it forces one to explicitly formulate hypotheses which are necessarily made in one way or another.

The fact that the required knowledge on the error vector is limited to the statistical moments of only the first two orders is due to the fact that the estimate \mathbf{x}^a has *a priori* been sought under the linear form (3.10). The determination of the most general least-variance estimate would require the knowledge of the entire probability distribution function of the observation error. However, in the case when the error vector is Gaussian, the associated conditional probability distribution function for \mathbf{x}^t is also Gaussian, with expectation defined by (3.10–12), and covariance defined by (3.13). In the Gaussian case, the BLUE therefore entirely solves the problem

of determining the conditional probability distribution function for the state vector \mathbf{x}^t .

4. The sequential form of statistical linear estimation. Kalman filtering

Formulae (3.10, 12 and 13) assume a particular form, extremely useful in many applications, when the “observation vector” \mathbf{z} defined by Eq. (3.9) can be decomposed into two components $\mathbf{z} = (\mathbf{x}^b, \mathbf{y}^o)^T$, where \mathbf{x}^b is a prior estimate of the vector \mathbf{x}^t , or *background* estimate, and \mathbf{y}^o is an additional set of observations, with dimension p . The background can be written as

$$\mathbf{x}^b = \mathbf{x}^t + \zeta^b \quad (4.1a)$$

where ζ^b is the corresponding error, while \mathbf{y}^o , assumed to be associated with a $p \times n$ observation matrix \mathbf{H} , can be written as

$$\mathbf{y}^o = \mathbf{H}\mathbf{x}^t + \epsilon \quad (4.1b)$$

where ϵ is the associated observation error. Formulae (4.1) are the analogue of (3.9), the dimension of the observation vector being now $m = n + p$, the corresponding observation matrix being $\Gamma = (\mathbf{I}_n, \mathbf{H}^T)^T$, and the corresponding observation error vector being $\zeta = (\zeta^b, \epsilon^T)^T$. As for the covariance matrix of the observation error, it is defined by

$$\Sigma = \begin{bmatrix} E(\zeta^b \zeta^{bT}) & E(\zeta^b \epsilon^T) \\ E(\epsilon \zeta^{bT}) & E(\epsilon \epsilon^T) \end{bmatrix} \quad (4.2)$$

It is important to stress that the background \mathbf{x}^b need not, and will normally not, consist of “observations” in the strict sense of the word. The background can for instance be an already known statistical or “climatological” average of the vector \mathbf{x}^t , or it can be any estimate of \mathbf{x}^t , obtained through whatever means may have been available: theoretical developments, or integration of a NWP model. The only important thing is that \mathbf{x}^b be numerically known, together with the corresponding statistical covariance of the error ζ^b .

Formulae (3.12–13) can be applied on the above quantities in order to obtain the corresponding BLUE \mathbf{x}^a and the covariance matrix \mathbf{P}^a of the associated estimation error. The results can be put into forms which are extremely useful from both the theoretical and the numerical points of view. For simplicity, we will assume that the errors ζ^b and ϵ are statistically uncorrelated, so that the off-diagonal terms in (4.2) are zero. The matrix $E(\zeta^b \zeta^{bT})$ will be denoted \mathbf{P}^b , and the matrix $E(\epsilon \epsilon^T)$ will be denoted \mathbf{R} . With these notations, Eqs. (3.10 and 12) and (3.13) can be put in the respective forms

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{P}^b \mathbf{H}^T [\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R}]^{-1} (\mathbf{y}^o - \mathbf{H} \mathbf{x}^b) \quad (4.3)$$

and

$$\mathbf{P}^a = \mathbf{P}^b - \mathbf{P}^b \mathbf{H}^T [\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R}]^{-1} \mathbf{H} \mathbf{P}^b \quad (4.4)$$

Considering first Eq. (4.3), we see that it defines the analysed state \mathbf{x}^a as the sum of the background \mathbf{x}^b and of a correction term. The latter is proportional to the vector $\mathbf{y}^o - \mathbf{H}\mathbf{x}^b$, *i.e.* to the difference between the additional observation vector \mathbf{y}^o and what the observation operator \mathbf{H} would produce if it was applied to the background \mathbf{x}^b . That difference is therefore essentially the lack of agreement between the background and the new observations. It is obvious that, if that difference happened to be exactly equal to zero, *i.e.* if the background happened to agree perfectly with the new observations, there would be no point in performing any correction on the background. And, in the linear approach followed here, the correction to be applied to the background naturally appears as a linear function of the difference vector $\mathbf{y}^o - \mathbf{H}\mathbf{x}^b$. The corresponding matrix $\mathbf{K} = \mathbf{P}^b \mathbf{H}^T [\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R}]^{-1}$, which is called the *gain matrix*, is simply the matrix which, taking into account the respective accuracies of the background and of the observations, as defined by the covariance matrices \mathbf{P}^b and \mathbf{R} , produces the best estimate, in the sense of the minimum of variance, of the state vector \mathbf{x}^t .

The vector $\mathbf{y}^o - \mathbf{H}\mathbf{x}^b$ is called the vector of *residuals*, or the *innovation* vector. That second denomination, which comes from general estimation theory, is extremely suggestive, because the vector $\mathbf{y}^o - \mathbf{H}\mathbf{x}^b$ effectively describes *all the new information* contained in the additional observation vector \mathbf{y}^o .

Another remark can be made about Eq. (4.3). Its numerical implementation requires the inversion of the matrix $\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R}$, which is of dimension p . The direct use of Eq. (3.12) requires (at least) the inversion of the matrix $\Gamma^T \Sigma^{-1} \Gamma$, which is of dimension n . If $p \ll n$ (which is usually the case in meteorological problems, where the number of observations available at a given time is normally much smaller than the dimension of the model state vector), use of formula (4.3) is much more economical. Also, formula (4.3) does not require the covariance matrix Σ to be invertible, contrary to what Eq. (3.12) does (at least apparently). If for instance, the additional observations \mathbf{y}^o are perfect ($\mathbf{R} = 0$), Eq. (4.3) can still be used. It is interesting to mention that, in this case of exact observations, the analysed state \mathbf{x}^a will be exactly compatible with the observations, in the sense that it will verify the equality $\mathbf{H}\mathbf{x}^a = \mathbf{y}^o$.

As for Eq. (4.4), it too has a clear significance. It defines the analysis error covariance matrix \mathbf{P}^a as the background error covariance matrix \mathbf{P}^b minus a correction matrix. The latter is symmetric, with non-negative eigenvalues, which implies that

the analysis error variance on any parameter is at most equal to the corresponding background error variance. The second term on the right-hand side of Eq. (4.4) therefore represents the gain brought about by the additional observations \mathbf{y}^o on the accuracy with which the state vector \mathbf{x}^t is known.

We mention another expression for \mathbf{P}^a , directly obtainable from (3.13)

$$(\mathbf{P}^a)^{-1} = (\mathbf{P}^b)^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}$$

This expression is generally not of much use for numerical applications, but its analogy with (3.7) is obvious: it also expresses that the “precision” of the analysis is the sum of the precisions of the background on the one hand, and of the additional observations on the other.

Finally, the variational problem associated with the estimation of \mathbf{x}^t from the background \mathbf{x}^b and the additional observation vector \mathbf{y}^o is easily seen from (3.14) to correspond to the minimization of the objective function

$$J(\mathbf{x}) \equiv [\mathbf{x} - \mathbf{x}^b]^T (\mathbf{P}^b)^{-1} [\mathbf{x} - \mathbf{x}^b] + [\mathbf{H}\mathbf{x} - \mathbf{y}^o]^T \times \mathbf{R}^{-1} [\mathbf{H}\mathbf{x} - \mathbf{y}^o] \quad (4.5)$$

where \mathbf{x} is, as before, an n -vector. The objective function is the sum of two terms, one measuring the distance to the background \mathbf{x}^b , the other measuring the distance to the additional observation vector \mathbf{y}^o . These two terms are weighted by the inverse covariance matrices of the corresponding errors.

Formulae (4.3–4) express the most general form of optimal interpolation, which, as already mentioned, is at the heart of most of the operational analysis techniques. It is most usually implemented in the following way: a background \mathbf{x}^b produced by the assimilating model for the analysis time is combined with a vector of observations \mathbf{y}^o at the same time through some approximate form of Eq. (4.3). The integration of the model is then restarted from the analysed state \mathbf{x}^a until the next observation time, at which a new analysis is performed. This forms the basis of sequential assimilation, as it is implemented at present, with many variants, in operational NWP. Implementation of Eq. (4.3) requires the preliminary knowledge of the covariance matrices \mathbf{P}^b and \mathbf{R} of the forecast and observation errors respectively. Reliable specification of those matrices, especially of \mathbf{P}^b , raises a number of problems, which will not be discussed here. Let us only mention that \mathbf{P}^b is usually modelled on the basis of a number of simple hypotheses on the shape and spatial extension of the corresponding covariance functions. It is also commonly assumed that the forecast errors on geopotential and wind are geostrophically related in middle and high latitudes, which allows unambiguous determination of all required covariances from the knowledge of the covariance of geopotential forecast error

only. A number of simplifications are also made in order to reduce the computational cost of Eqs. (4.3) and especially (4.4). Indeed, exact implementation of those formulæ would be much too costly for operational NWP. Concerning (4.3), its implementation requires at least solving one linear system of dimension p . In a typical meteorological situation, p can be of the order of $10^4 - 10^5$. This is much too large for the constraints of operational NWP. In order to reduce the corresponding cost, only observations located in the vicinity of a given model grid-point are used when performing the analysis at that point. That “selection” of observations is certainly legitimate in the sense that observations performed at a large distance of a given point must have a small influence on the analysed fields at that point. However, experience shows that it nevertheless introduces spatial noise in the analysed fields, which must be then filtered out by *ad hoc* smoothing. As for Eq. (4.4), it is not implemented in its totality, but only the diagonal terms of the matrix \mathbf{P}^a , i.e. the variances of the analysis error, are usually computed.

Experience also shows that, in spite of the geostrophic link assumed between the wind and geopotential forecast errors, the fields produced by optimal interpolation are still contaminated by unrealistic ageostrophic noise. This noise must be filtered out through additional appropriate initialization procedures, already alluded to in the Introduction (for more information on this aspect, see Temperton, 1988, and references therein).

Optimal interpolation, as implemented in present operational NWP, produces results which are on the whole quite satisfactory. Description of its performances can be found in a number of articles or reports. An article by Lorenc (1981), although it is not very recent, contains a clear description of the basic principles of operational optimal interpolation and of the main properties of the results it produces. Among recent developments, the National Meteorological Center (Washington, USA) has introduced in operations an implementation of optimal interpolation, based not on Eq. (4.3), but on an iterative minimization of the objective function (4.5) (Parrish and Derber, 1992). The results show improvement of the quality of the analysis for a lower computational cost. This is probably largely due to the suppression of the need for selecting observations in the vicinity of each analysis point. In addition, an appropriate definition of the matrix \mathbf{P}^b eliminates the need for initialization. It therefore appears that it is numerically more efficient to perform optimal interpolation, not by direct use of Eq. (4.3) and explicit exact solution of one (or several) linear system of equations, but by iterative (and approximate) minimization of the corresponding ob-

jective function (4.5). At the time of writing, the European Centre for Medium-range Weather Forecasts was also planning to introduce soon a similar “three-dimensional variational analysis”.

But there is much more to formulæ (4.3–4) than analysis at a given time, and those formulæ are at the basis of the technique of *Kalman filtering* which, in the linear context considered here, is the appropriate approach to sequential assimilation. Let us first consider the following situation. A vector \mathbf{z} of observations, of the form (3.9), has been processed through eqs (3.10–12) and (3.13) in order to produce the corresponding BLUE, which we will denote \mathbf{x}^a , together with the covariance matrix of the associated estimation error, which we will denote \mathbf{P}^a . At that stage, an additional vector of observations \mathbf{y}^o , of the form (4.1b), becomes available, and one wants to determine the BLUE \mathbf{x}^a of \mathbf{x}^t from the combined observation vector $\mathbf{v} = (\mathbf{z}^T, \mathbf{y}^{oT})^T$. Is it necessary to restart the computations from the beginning, or is it possible to take advantage of the computations that have already been performed and have led to \mathbf{x}^a and \mathbf{P}^a ? It must be clear from the foregoing developments, and it is easy to verify that, if the observation error vector ϵ associated with \mathbf{y}^o is uncorrelated with the estimation error vector $\mathbf{x}^a - \mathbf{x}^t$, then the BLUE \mathbf{x}^a and the associated covariance matrix of estimation error are given by formulæ (4.3) and (4.4), \mathbf{x}^b and \mathbf{P}^b being replaced by \mathbf{x}^a and \mathbf{P}^a respectively. There is no need therefore for restarting the computations from the original \mathbf{z} , and one can take advantage of the already performed computations. In particular, if the dimension p of the additional observation vector \mathbf{y}^o is small in comparison to the dimension n of \mathbf{x}^t , the numerical gain of using formulæ (4.3–4) rather than restarting the entire computations is obvious.

Formulæ (4.3–4) therefore provide a way for constantly updating with new observations the latest estimate of the state of the system under observation. If the errors associated with the successive observations are mutually uncorrelated, the estimate obtained at any stage of the process will always be the BLUE of \mathbf{x}^t from the observations already introduced into the process, and there will be no loss in the accuracy of the estimate resulting from the sequential character of the procedure.

In the case of assimilation of observations, an additional complication comes from the fact that the observations to be assimilated are distributed over a time period over which the state of the system is itself evolving. In order to take the temporal dimension into account, and in agreement with the linear approach followed so far, we consider a system whose state evolves in time according to the linear equation

$$\mathbf{x}_{k+1}^t = \mathbf{M}\mathbf{x}_k^t + \boldsymbol{\eta}_k \quad (4.6)$$

where \mathbf{x}_k^t denotes the state of the system at time k , while \mathbf{M} is a known *transition matrix* expressing the time variation of the state vector between times k and $k+1$. As for the term η_k , it represents contributions to the temporal variations of the state vector which are not represented by the transition matrix. One can consider that the transition matrix is the numerical model available for performing the assimilation, while the term η_k represents the accumulated effect, on the variation of the state vector between times k and $k+1$, of the various processes not represented by \mathbf{M} . The term η_k will accordingly be called the *model error*. It will be considered as an unbiased random vector, uncorrelated in time, with known covariance matrix \mathbf{Q}

$$E(\eta_k \eta_l^T) = \mathbf{Q} \delta_{kl} \quad (4.7)$$

where δ_{kl} is the classical Kronecker symbol (in order to keep notations simple, we do not allow for an explicit time dependence of \mathbf{M} and \mathbf{Q} ; introducing such a dependence is straightforward, and would not modify the essence of what is to follow).

We assume in addition that observations of the general form (4.1b) are available at the successive instants $k = 0, 1, \dots, N$:

$$\mathbf{y}_k^o = \mathbf{H} \mathbf{x}_k^t + \varepsilon_k \quad (4.8)$$

The corresponding observation errors ε_k are supposed as before to be unbiased, to be uncorrelated in time and have covariance matrix \mathbf{R} (with again no explicit time dependence for \mathbf{H} and \mathbf{R}). This leads to the expression

$$E(\varepsilon_k \varepsilon_l^T) = \mathbf{R} \delta_{kl}$$

In addition, the model and observation errors are supposed to be mutually uncorrelated

$$E(\varepsilon_k \eta_l^T) = 0$$

We now assume that the BLUE \mathbf{x}_k^a of the state \mathbf{x}_k^t of the system at time k from all observations up to time k has been determined, together with the covariance matrix \mathbf{P}_k^a of the corresponding estimation error. The BLUE of the state of the system at time $k+1$ from all observations up to time k can be shown to be equal to

$$\mathbf{x}_{k+1}^b = \mathbf{M} \mathbf{x}_k^a \quad (4.9)$$

As for the covariance matrix of the corresponding estimation error $\mathbf{x}_{k+1}^b - \mathbf{x}_{k+1}^t$, it is equal to

$$\begin{aligned} \mathbf{P}_{k+1}^b &\equiv E[(\mathbf{x}_{k+1}^b - \mathbf{x}_{k+1}^t)(\mathbf{x}_{k+1}^b - \mathbf{x}_{k+1}^t)^T] \\ &= E[(\mathbf{M} \mathbf{x}_k^a - \mathbf{M} \mathbf{x}_k^t - \boldsymbol{\eta}_k)(\mathbf{M} \mathbf{x}_k^a - \mathbf{M} \mathbf{x}_k^t - \boldsymbol{\eta}_k)^T] \\ &= E\{[\mathbf{M}(\mathbf{x}_k^a - \mathbf{x}_k^t) - \boldsymbol{\eta}_k][\mathbf{M}(\mathbf{x}_k^a - \mathbf{x}_k^t) - \boldsymbol{\eta}_k]^T\} \\ &= \mathbf{M} \mathbf{P}_k^a \mathbf{M}^T + \mathbf{Q} \end{aligned} \quad (4.10)$$

where the various non-correlation hypotheses have been used. The first term on the last line represents

the error at time $k+1$ resulting from the error at time k , while the second term is the contribution of the model error η_k between times k and $k+1$.

At this stage, it is easy to introduce the observation vector \mathbf{y}_{k+1}^o at time $k+1$: one simply has to use formulæ (4.3–4), \mathbf{x}^b and \mathbf{P}^b being replaced by \mathbf{x}_{k+1}^b and \mathbf{P}_{k+1}^b respectively, and \mathbf{y}^o being replaced by \mathbf{y}_{k+1}^o . This leads to the BLUE \mathbf{x}_{k+1}^a of the state of the system at time $k+1$, from all the observations up to, and including, time $k+1$

$$\begin{aligned} \mathbf{x}_{k+1}^a &= \mathbf{x}_{k+1}^b + \mathbf{P}_{k+1}^b \\ &\times \mathbf{H}^T [\mathbf{H} \mathbf{P}_{k+1}^b \mathbf{H}^T + \mathbf{R}]^{-1} (\mathbf{y}_{k+1}^o - \mathbf{H} \mathbf{x}_{k+1}^b) \end{aligned} \quad (4.11)$$

and to the covariance matrix of the corresponding estimation error, *viz.*

$$\begin{aligned} \mathbf{P}_{k+1}^a &= \mathbf{P}_{k+1}^b - \mathbf{P}_{k+1}^b \\ &\times \mathbf{H}^T [\mathbf{H} \mathbf{P}_{k+1}^b \mathbf{H}^T + \mathbf{R}]^{-1} \mathbf{H} \mathbf{P}_{k+1}^b \end{aligned} \quad (4.12)$$

The sequential process defined by Eqs. (4.9) to (4.12) is called *Kalman filtering* (Kalman, 1960). At any stage, Kalman filtering produces the BLUE of the state of the system under observation, using all observations up to estimation time. It also produces the covariance matrix of the corresponding estimation error.

Kalman filtering has been applied to many different problems. A general description of the theory of Kalman filtering and of its properties can be found in, *e.g.*, Jazwinski (1970). In the case of assimilation of meteorological or oceanographical observations, one can see that, if one accepts the linear hypotheses which underlie Eqs. (4.6) and (4.8), Kalman filtering fulfills the goal assigned in the Introduction to an ideal assimilation system : namely, to use all the available information in order to produce the most accurate possible description of the state of the flow, together with the uncertainty resulting from the uncertainties on the various sources of information. In the present case, the available information consists, not only of the observations \mathbf{y}_k^o (Eq. 4.8), but also of the model (4.9) (and of the initial estimate \mathbf{x}_0^b from which the whole process must be started). As for the associated uncertainties, they are defined by the covariance matrices \mathbf{R} and \mathbf{Q} (and the initial covariance matrix \mathbf{P}_0^b). Kalman filtering consistently combines all these elements in order to produce the BLUE (4.11) and the associated covariance matrix (4.12).

The application of Kalman filtering to assimilation of meteorological and oceanographical observations has been studied by a number of authors, in particular Ghil and collaborators (see, *e.g.*, Ghil, 1989, or Ghil and Malanotte-Rizzoli, 1991). Experiments performed with various linear systems have produced convincing results as to the capability of

the method for effectively extracting the information contained in the observations and the model. One major difficulty with Kalman filtering in the context of assimilation of meteorological and oceanographical observations is its numerical cost. Writing the first term on the last line of Eq. (4.10) under the form $\mathbf{M}(\mathbf{M}\mathbf{P}^a_k)^T$ shows that the corresponding computations require two successive matrix multiplications by \mathbf{M} . Now, the multiplication of one vector by \mathbf{M} corresponds to one integration of the model between times k and $k+1$ (Eq. 4.9). Implementation of Eq. (4.10) therefore requires $2n$ integrations of the model, where n is, as above, the dimension of the state vector of the model. With values of n on the order of $10^6 - 10^7$, this goes largely beyond the possibilities of assimilation for operational NWP, or even of *a posteriori* assimilation. In operational NWP, the computation (4.10) is replaced by a simple multiplication of the variances of the analysis errors by an *a priori* specified coefficient (typically, 1.5 for a 6hr-forecast), the associated correlations being modelled independently, as already mentioned above. The corresponding numerical cost is negligible, but that procedure amounts to ignoring the influence of the particular meteorological situation under consideration, and especially of the particular instabilities that may develop, on the evolution of the forecast error. This certainly is one of the major weaknesses of present operational assimilation methods, which, in the perspective taken here, can be described as degraded but economical forms of Kalman filtering. Comparisons of the results produced by variational assimilation and by algorithms similar to operational algorithms (Rabier *et al.*, 1993) suggest that a more accurate description of the evolution of forecast error might substantially improve the quality of assimilations.

Now, the correlation between forecast errors at points located a large distance apart must be negligible, and a large proportion of the entries of covariance matrices such as \mathbf{P}^a_k must have zero or negligible values. This should allow to reduce the cost of computation (4.10). In addition, it is known that the most rapidly amplifying modes in the evolution of the forecast error are geostrophic modes (see, *e.g.*, Lacarra and Talagrand, 1988), so that it should be possible to restrict computation (4.10) to a subset of all the model modes. These ideas have been exploited by several authors (see, *e.g.*, Cohn and Parrish, 1991, Dee, 1991, or Bouttier, 1994) in order to reduce the cost of computation (4.10). Much active research is now being done on the problem of defining algorithms for describing the temporal evolution of the forecast error that are both economical enough for practical implementation, and accurate enough for improving on present operational methods of sequential assimilation.

5. The variational form of statistical linear estimation

We will now restrict ourselves to the case where the model is supposed to be perfect, *i.e.* $\eta_k = 0$ in Eq. (4.6), so that the exact evolution of the flow reduces to

$$\mathbf{x}^t_{k+1} = \mathbf{M}\mathbf{x}^t_k \quad (5.1)$$

The variational form (3.14) of the estimation problem defined by Eqs. (4.8) and (5.1) leads to the objective function

$$J(\mathbf{x}) \equiv \sum_{0 \leq k \leq N} [\mathbf{H}\mathbf{x}_k - \mathbf{y}^o_k]^T \mathbf{R}^{-1} [\mathbf{H}\mathbf{x}_k - \mathbf{y}^o_k] \quad (5.2)$$

where $\mathbf{x} = (\mathbf{x}_k^T)^T$ is a sequence of model states at successive times, linked by the model equation (5.1). $J(\mathbf{x})$ is the sum of the model-minus-observations squared differences, weighted by the inverse of the observation error covariance matrices. Minimizing the objective function (5.2) under the constraint (5.1) will produce at any time k the BLUE of the real state \mathbf{x}^t_k of the system at time k , from *all the available observations*, *i.e.* from observations performed before, at, and after time k . In particular, the state at the end of the assimilation period will be the same as the state produced by Kalman filtering (under the assumption of an exact model, *i.e.* under the condition that $\mathbf{Q} = 0$ in Eq. 4.10). The variational form of statistical estimation therefore provides a way to globally adjust a model to observations distributed in time.

But we can also note that the assumption of linearity, necessary to establish the link with the theory of statistical linear estimation and with Kalman filtering, is by no means necessary for a variational problem of type (5.1–2). One can very well consider the problem of minimizing an objective function of the form (5.2), under a constraint of the form (5.1), where the matrices \mathbf{M} and \mathbf{H} are replaced by nonlinear operators. Indeed, numerical models of the atmospheric or oceanic flows are nonlinear, and many observations are nonlinearly related to the atmospheric or oceanic variables one wants to estimate. For an already mentioned example, infrared radiances measured by satellites are related to the temperature and humidity profiles of the emitting atmospheric columns through the radiative transfer equation, which is strongly nonlinear. We will therefore drop for the time being the hypothesis of linearity (only to come back to it later in order to show that it is often justified in some sense) and consider the problem of minimizing a nonlinear objective function (5.2) (*i.e.* an objective function with nonlinear observation operators) under a nonlinear constraint of the form (5.1). In order to stress that we are now dealing with nonlinear operators, we will

use the notations M and H instead of \mathbf{M} and \mathbf{H} respectively.

There basically exist two methods for solving a constrained minimization problem. The principle of the first method is obvious, and consists in *reducing the constraint*, i.e. in eliminating some of the constrained variables so as to transform the problem into an unconstrained problem. In the present case, one can note that a model solution (5.1) is uniquely defined by the specification of the corresponding initial condition \mathbf{x}_0 . The objective function J can therefore be considered as a function of \mathbf{x}_0 only, upon which no constraint is imposed, so that one is led to a problem of unconstrained minimization with respect to \mathbf{x}_0 . The second method, whose principle is much less obvious, consists in associating unknown coefficients, called *Lagrange multipliers*, with the constraints of the problem, and to form the corresponding *Lagrangian*. In the present case, there are N constraints (5.1), each of which of dimension n , and the set Λ of Lagrange multipliers consists of N vectors $\Lambda_k (k = 0, 1, \dots, N-1)$, each of dimension n . The associated Lagrangian reads

$$L(\mathbf{x}, \Lambda) = J(\mathbf{x}) + \sum_{1 \leq k < N} \Lambda_k^T [\mathbf{x}_{k+1} - M \mathbf{x}_k]$$

A well-known theorem then says that the minima of the constrained minimization problem (5.1–2) correspond to the stationary points of the Lagrangian $L(\mathbf{x}, \Lambda)$, considered as a function of the independent variables \mathbf{x} and Λ .

The method of *adjoint equations*, which is a classical tool of control theory (Lions, 1971), seems to be by far the most efficient way for numerically solving the minimization problem (5.1–2). Interestingly enough, the method of adjoint equations can be derived by either reducing the constraint (5.1) so as to use only the initial state \mathbf{x}_0 as independent variable (see, e.g., Le Dimet and Talagrand, 1986, or Talagrand and Courtier, 1987), or alternatively by using the technique of Lagrange multipliers (see, e.g., Thacker and Long, 1988). Assuming for instance that we want to solve problem (5.1–2) as a problem of unconstrained minimization with respect to the initial state \mathbf{x}_0 , it is necessary, in order to even start solving the problem, to be able to relate the variations of the initial state \mathbf{x}_0 to the corresponding variations of the objective function J . For a given initial state, these variations are related through the local vector of partial derivatives, or *gradient vector*, of the objective function with respect to the components of the initial state. In particular, if one is able to numerically compute the gradient for a given initial state, it will be possible to feed that gradient into a standard minimization algorithm which will determine the minimizing initial state through successive iterations. In most situations, it will of

course be impossible to establish explicit analytical expressions for the gradient. It is possible to numerically (and approximately) determine the gradient through explicit finite perturbations of the initial state, but this would be much too costly for practical implementation: it would require to compute the objective function, i.e. to effectively integrate the model over the assimilation period, as many times as there are independent components in the initial state. The method of adjoint equations provides a way for computing the gradient at a numerical cost which is at most a few times the cost of one computation of the distance function. The principle of the method is extremely simple. Let us consider a computer code (or part of a code) which, starting from some input vector \mathbf{u} with components $u_i (i = 1, \dots, q)$, produces an output vector \mathbf{v} with components $v_j (j = 1, \dots, r)$. The process can be described by the equation

$$\mathbf{v} = G(\mathbf{u}) \quad (5.3)$$

where G stands for all the computations that lead from \mathbf{u} to \mathbf{v} . For a given perturbation du on the input, the resulting perturbation $\delta \mathbf{v}$ on the output is equal to first order to

$$\delta \mathbf{v} = G' \delta \mathbf{u} \quad (5.4)$$

where G' is the matrix of local partial derivatives, or *jacobian matrix*, of the components of \mathbf{v} with respect to the components of \mathbf{u} . Eq. (5.4) is called the *tangent linear equation* to (5.3). Let now $J(\mathbf{v})$ be a scalar function of the output \mathbf{v} . The gradient of J with respect to \mathbf{u} is given by the chain rule

$$\frac{\partial J}{\partial u_i} = \sum_{j=1}^r \frac{\partial v_j}{\partial u_i} \frac{\partial J}{\partial v_j} \quad i = 1, \dots, q$$

or, in transparent matrix notation

$$\nabla_{\mathbf{u}} J = G'^T \nabla_{\mathbf{v}} J \quad (5.5)$$

where, as before, the superscript T denotes transposition.

The adjoint method is based on a systematic use of formula (5.5). More precisely, let us suppose that the process G is the composition of a number of more elementary processes, namely

$$G = G_c \circ \dots \circ G_2 \circ G_1$$

the jacobian G' will be product of the elementary jacobians

$$G' = G'_M \dots G'_2 G'_1$$

and the transpose G'^T will be the product of the elementary transposes, taken in reversed order

$$G'^T = G_1'^T G_2'^T \dots G_M'^T$$

This shows that, in order to numerically determine the gradient $\nabla_{\mathbf{u}}J$ with respect to the input \mathbf{u} , it is sufficient to proceed backwards through the direct computations and, at every step, to perform the corresponding transpose, or *adjoint* computations. The total cost of one adjoint computation (5.5) will generally be of the same order of magnitude as the cost of one direct computation (5.3). (It can be shown, see *e.g.*, Morgenstern, 1984, that the total operation count of one adjoint computation can be reduced to at most 4 times the total operation count of the corresponding direct computation, this ratio being reduced to 2 if one considers only multiplications and divisions). This is of course much more economical than direct perturbations of the input vector.

In the case of the determination of the gradient of the (nonlinear) objective function (5.2) with respect to the initial state \mathbf{x}_0 of the (nonlinear) assimilating model (5.1), the adjoint computations reduce to integrating the equation (see, *e.g.*, Talagrand and Courtier, 1987)

$$\delta' \mathbf{x}_k = M'^T \delta' \mathbf{x}_{k+1} + H'^T \mathbf{R}^{-1} [H \mathbf{x}_k - \mathbf{y}^o_k] \quad (5.6)$$

backwards in time, starting from the “final” state $\delta' \mathbf{x}_{N+1} = 0$. In this equation, M' and H' are the jacobians of the respective nonlinear model and observation operators M and H . The gradient of the objective function with respect to the initial state \mathbf{x}_0 is equal to $2\delta' \mathbf{x}_0$.

It is seen that the basic model solution \mathbf{x}_k under consideration explicitly appears in the adjoint equation (5.6) in the quantity $H \mathbf{x}_k - \mathbf{y}^o_k$, which, except for its sign, is the innovation vector of Eqs. (4.3) and (4.11). This means that the basic solution will have to be computed, and kept in memory, before the adjoint integration can be performed. In the general case of nonlinear operators M and H , the basic solution will also be necessary for determining the jacobians M' and H' . Saving the basic solution in memory may entail large core requirements, which constitute one important feature of the adjoint method.

The method of adjoint equations for performing variational assimilation of meteorological observations seems to have been first suggested by Penenko and Obraztsov (1976), who applied it to a simple, small-dimensional linear problem. Since then, a large number of experiments have been performed on (usually nonlinear) models of increasing complexity and dimensions, and with various types of observations. Experiments have been performed on both meteorological and oceanographical examples (for meteorological applications see, *e.g.*, Lewis and Derber, 1985, Talagrand and Courtier, 1987, Derber, 1987, Courtier and Talagrand, 1987, 1990, Lorenc, 1988, Thépaut and Courtier, 1991, Navon *et al.*, 1991, Rabier and Courtier, 1992; for oceanographical applications see, *e.g.*, Thacker and Long, 1988,

Sheinbaum and Anderson, 1990a and b, Greiner and Perigaud, 1994). The first general conclusion that can be drawn from these experiments is that variational assimilation works in that it is capable of minimizing the objective function. Also, and contrary to what happens in sequential assimilation, there is propagation of information, as should be, both forward and backward in time. However, when the objective function contains only terms measuring the misfit between individual observations and model values, the minimization solution tends to contain unrealistic “noise”, often under the form of small-scale oscillations and/or of ageostrophic gravity waves. The minimizing solution is physically realistic only if appropriate terms, measuring the energy contained in the small scales of the flow, or its ageostrophy, are added to the objective function (see, *e.g.*, Courtier and Talagrand, 1990, Thépaut and Courtier, 1991). Indeed, the need for adding terms intended at avoiding unrealistic oscillations in the estimated fields is by no means restricted to meteorological or oceanographical applications, but is almost universal in problems where fields are estimated through a variational method. Such terms are often called “smoothing”, “penalizing” or still “regularizing” terms. But it must be stressed that the need for appropriate smoothing is not restricted to variational methods. It is also present in statistical estimation when implemented through Kalman filtering which, as already said, must lead to the same final result as variational algorithms. In present operational optimal interpolation, the requirement for appropriate smoothing is satisfied on the one hand through the presence of the background \mathbf{x}^b , which defines what the analyzed field must be in data-void areas, and on the other hand through the “initialization” process, which filters out unrealistic gravity wave oscillations.

The most recent experiments of variational assimilation of meteorological observations have been performed with multilevel primitive equation models similar, but not identical yet, to the models used in NWP (Thépaut *et al.*, 1993). The remaining differences lie in the resolution, which is still coarser in variational assimilation experiments (typically one order of magnitude less points in the horizontal than in operational models), and in the representation of many “physical” processes, such as convection and water phase changes, which are still absent from variational assimilation. These recent experiments confirm the results previously obtained, and show in particular that variational assimilation, because it explicitly uses the evolution equations of the system, is able to propagate the information contained in the observations much more accurately than operational optimal interpolation.

Variational assimilation, like Kalman filtering, is therefore able to assimilate observations in a way

that is exactly consistent with the dynamics of the system, as described by the model equations. But, as in the case of Kalman filtering, and in spite of the fact that the adjoint equations are by far the most efficient way for computing the gradient of the objective function, the computational price to be paid is heavy : in addition to the necessity of storing in memory the model solution produced by the direct integration which must be performed before each adjoint integration, minimization of the objective function typically requires from 10 to 30 iterations of the minimization algorithm. Each iteration itself requires one integration of the model over the assimilation period, followed by one adjoint integration. The cost of one adjoint integration is about twice the cost of one direct integration, so that one minimization typically requires the equivalent of between 30 and 100 integrations of the model over the assimilation period. This is of course very costly,

but it now seems it will be possible to operationally implement in the coming years simplified forms of variational assimilation, in which the assimilation will be performed at a somewhat lower resolution than the full NWP model (Courtier *et al.*, 1994).

The present situation as concerns assimilation methods is therefore rather clear. In addition to the relatively simple, rather *ad hoc*, but economical and basically satisfactory operational algorithms, there exist two broad classes of algorithms that are capable, in the general framework of statistical linear estimation, of consistently extracting the information contained in the observations on the one hand and in the physical laws expressed by the assimilating model on the other : Kalman filtering and variational assimilation. In order to implement these algorithms, one must express the observations under the general form (3.9), *i.e.* one must know what has been measured, in terms of the parameters to be estimated, and with which accuracy. Now, exact implementation of either of these two classes of algorithms is numerically costly, and a large part of the research being done at present on assimilation is in effect directed at determining the most cost-efficient simplifications that can be made on them. This task may indeed be with us for a long time : no end is foreseen to the increase in the power of computers and to the deep modifications in their structures, nor to the development of new observing systems and of more realistic models. Changes in any of these aspects may radically modify any conclusion one may have reached as to the most efficient way to perform assimilation.

We will now briefly comment on the relative advantages and disadvantages of Kalman filtering and variational assimilation. As already mentioned, both algorithms will lead to the same final state at the end of the assimilation period in the case of lin-

ear observations (*i.e.* of a linear observation operator \mathbf{H}) and of a perfect (*i.e.* $\eta_k = 0$) linear model (5.1). A basic difference between the two algorithms is that Kalman filtering carries information only from the past into the future, while variational assimilation carries information in both time directions. On the other hand, variational assimilation, contrary to Kalman filtering, does not take into account the fact that the assimilating model, like the observations, will never be perfect and will always contain errors. But it must also be said that these differences are only true of these algorithms as they have been described here and as they have been most usually implemented so far in meteorological and oceanographical applications. As concerns Kalman filtering, there exists a procedure, called Kalman *smoothing* (see, *e.g.*, Anderson, 1979), which allows, once a first pass has been performed over the assimilation period, to proceed backward in time so as to obtain, at any intermediate time k , the BLUE of the state of the system at time k from all available observations, performed before, at or after time k . We will not describe here the theory of Kalman smoothing, which is related to the theory of adjoint equations, and will only refer to Gaspar and Wunsch (1989) as a simple but instructive example of an application of Kalman smoothing to an oceanographical problem. And, as concerns variational assimilation, it can incorporate model errors: it suffices to impose the model equation (5.1) not as a constraint to be exactly satisfied by the sequence \mathbf{x}_k of assimilated states, but (to use the vocabulary introduced by Sasaki, 1970) as a “weak constraint” to be satisfied only approximately. This can be done by modifying the objective function (5.2) to

$$J(\mathbf{x}) \equiv \sum_{0 \leq k < N} [H\mathbf{x}_k - \mathbf{y}_k^o]^T \mathbf{R}^{-1} [H\mathbf{x}_k - \mathbf{y}_k^o] + \sum_{0 \leq k < N} [\mathbf{x}_{k+1} - M\mathbf{x}_k]^T \mathbf{Q}^{-1} [\mathbf{x}_{k+1} - M\mathbf{x}_k] \quad (5.7)$$

where \mathbf{Q} is, as in (4.7) the covariance matrix of the model error. The meaning of the second sum on the right-hand-side of (5.7) must be clear : it simply expresses that the difference $\mathbf{x}_{k+1} - M\mathbf{x}_k$ must not be considered as exactly zero, as it would be if the model was exact, but equal to zero only within the uncertainty defined by the matrix \mathbf{Q} . Accordingly, the sequence of states must be considered as unconstrained, and the minimization of J must be performed with respect to the entire sequence $\mathbf{x} = (\mathbf{x}_k^T)^T$. It is not difficult to see from the variational form (3.14) of statistical linear estimation, that the sequence of states minimizing (5.7) is made up of the BLUEs, at all times k , of the state of the system, from all observations over the entire assimilation period. It results in particular that minimization of (5.7) must lead to the same sequence

of assimilated states as Kalman filtering, followed by Kalman smoothing. For additional information on weak constraint variational assimilation, see Bennett (1992) and Bennett *et al.* (1993).

Another difference between Kalman filtering and variational assimilation is that the former produces, in addition to the BLUE of the state of the system, the covariance matrix of the corresponding estimation error, while the latter produces only the BLUE. In this sense, only Kalman filtering fulfills the goal assigned to an ideal assimilation system in the Introduction. This of course is obtained at the already mentioned much higher cost of Kalman filtering. Now, it is easy to verify from eq. (3.13) and (3.14) that the covariance matrix of the estimation error is the inverse of the hessian (matrix of second derivatives) of the objective function. The question therefore arises whether the inverse hessian can be computed in variational assimilation, at least to a sufficient degree of accuracy, at a lower cost than in Kalman filtering. Indeed, some minimization algorithms, of the so-called quasi-Newton type (see, *e.g.*, Gill *et al.*, 1982), do compute an approximate inverse hessian in the course of the minimization. The problem of the determination of the estimation error in variational assimilation is the subject of active research (Fisher, *pers. com.*).

To conclude with the theoretical and methodological aspects of assimilation, we will discuss a point which we have so far left in some obscurity, namely the validity of the linear hypothesis which underlies the theory of Kalman filtering and which, although not necessary for variational assimilation, gives it a special significance and facilitates the understanding and analysis of the results it produces. Not only are the equations governing the atmospheric and oceanic flows strongly nonlinear, but their nonlinearity is at the origin of one of the most important properties of these flows, namely their chaotic character. This character imposes stringent limits on the predictability of these flows, and one can legitimately wonder whether a linear hypothesis is legitimate in the context of assimilation. Considering first the equations for Kalman filtering, let us assume that the model and observation operators are nonlinear, and accordingly denoted M and H respectively. If the difference $\mathbf{x}^a_k - \mathbf{x}^t_k$ is small enough, the quantity $M\mathbf{x}^a_k - M\mathbf{x}^t_k$ (second line of Eq. 4.10) can be approximated by $M'(\mathbf{x}^a_k - \mathbf{x}^t_k)$, where M' is the jacobian matrix of the operator M , taken at point \mathbf{x}^a_k . Eq. (4.10) accordingly becomes

$$\mathbf{P}^b_{k+1} = M'\mathbf{P}^a_k M'^T + \mathbf{Q} \quad (5.8)$$

Similarly, if the difference $\mathbf{x}^t_{k+1} - \mathbf{x}^b_{k+1}$ is small enough so that the innovation vector $\mathbf{y}^o_{k+1} - H\mathbf{x}^b_{k+1} = H\mathbf{x}^t_{k+1} - H\mathbf{x}^b_{k+1} + \varepsilon_{k+1}$ can be approximated by $H'(\mathbf{x}^t_{k+1} - \mathbf{x}^b_{k+1}) + \varepsilon_{k+1}$, calculations

show that Eqs. (4.11) and (4.12) can be respectively replaced by

$$\begin{aligned} \mathbf{x}^a_{k+1} &= \mathbf{x}^b_{k+1} \\ &+ \mathbf{P}^b_{k+1} H'^T [H' \mathbf{P}^b_{k+1} H'^T + \mathbf{R}]^{-1} \\ &(\mathbf{y}^o_{k+1} - H\mathbf{x}^b_{k+1}) \end{aligned} \quad (5.9)$$

and

$$\begin{aligned} \mathbf{P}^a_{k+1} &= \mathbf{P}^b_{k+1} \\ &- \mathbf{P}^b_{k+1} H'^T [H' \mathbf{P}^b_{k+1} H'^T + \mathbf{R}]^{-1} \\ &H' \mathbf{P}^b_{k+1} \end{aligned} \quad (5.10)$$

In these equations, H has been replaced by the jacobian H' , except in the expression for the innovation vector. The algorithm defined by eqs (5.8 to 10), to which the nonlinear analogue to (4.9) must be added, is called *extended Kalman filtering* (see, *e.g.*, Jazwinski, 1970). It is valid whenever the differences between the real and estimated states of the system are small enough to allow local linearizations as just described.

A similar argument holds for variational assimilation. For a linear model and linear observation operators, the objective function (5.2) will be a quadratic function of the initial state \mathbf{x}_0 . For a nonlinear model or a nonlinear observation operator, the objective function will not be quadratic, but will remain approximately quadratic in a neighbourhood of its minimum. If the initial uncertainty on the state of the system (defined for example by the point from which the minimization process is initiated) is small enough to ensure that, at any stage of the assimilation, the estimated state of the system will always lie within that neighbourhood, the theoretical nonlinearity of the objective function will have no practical effect. In particular, the minimizing solution will be the BLUE of the state of the flow.

The linear hypothesis made previously, and the associated logic of statistical linear estimation, including in particular the equivalence between sequential and variational assimilation, will therefore be valid if the differences between the real and estimated states of the system are always small enough to allow the local linearizations described above. The validity of this so-called *tangent linear approximation* has been checked systematically in a number of situations. For instance, Lacarra and Talagrand (1988) have shown on a barotropic model that, for realistic amplitudes of the error on the initial state of the flow, a linear approximation for the evolution of the forecast error is valid up to about 48 or 72 hours. Similarly, Thépaut and Moll (1990) have shown that, within the uncertainty existing in practice on the atmospheric profiles of temperature and humidity, the tangent linear hypothesis is valid for the radiance observations performed by the TIROS Operational Vertical Sounder (TOVS) carried by the

satellites of the NOAA series. In addition, many results have indirectly confirmed the validity of the tangent linear hypothesis. This means that sequential or variational assimilation can confidently be expected to produce reasonable estimates of the state of the flow. But it certainly does not mean that there do not remain problems. There must exist limitations to the validity of the tangent linear hypothesis, due for instance to the presence, in the physically most realistic models, of processes capable of inducing sharp variations in the model fields. These limitations have so far not been clearly identified, and research on fully nonlinear assimilation is only starting (see, *e.g.*, Miller *et al.*, 1994, or Pires *et al.*, 1996). But there is also no doubt that much development work remains to be done within the context of statistical linear estimation.

These considerations apply primarily to the atmosphere. Concerning the ocean, Evensen (1992, 1994) has shown strong evidence that the tangent linear approximation may not always be valid, essentially because the temporal density of observations is too low. In the context of sequential assimilation, Evensen suggests the use of an "Ensemble Kalman filtering", in which the temporal evolution of the estimation error covariance matrix is computed, not through a formula of form (4.10), but through an ensemble of forecasts performed with the fully nonlinear model.

6. Assimilation of "indirect" observations

It has been shown above that, in order to implement statistical estimation, it is necessary to know, for each individual observation, what has been measured, and with which accuracy. A rather general practice so far, when dealing with "indirect" satellite observations, has been to first "invert" them to "geophysical variables", such as for instance temperatures and humidities, and then to use the inverted fields as observations in the assimilation algorithm. This is commonly done, for instance, for radiance observations, which are inverted to produce estimates of the atmospheric temperature and humidity profiles. Now, such a preliminary inversion is by no means necessary. It does not avoid the basic need for the definition of an appropriate observation operator and for the specification of the associated observation error. And it usually requires a background which itself depends on the other available observations. This leads to interdependence of the errors associated with the various "observations" used in the assimilation. The problems raised by such an interdependence can be solved in the context of statistical estimation, but it certainly seems preferable to avoid them in the first place. For these reasons, and also in order to define a systematic approach to be followed for any type of observations, the tendency is now to avoid as much as possible preliminary processing

of the observations before the assimilation, and to introduce raw observations in the assimilation algorithm with an appropriate observation operator. In the case of radiance measurements, the associated problems have been studied, among others, by Eyre (1989a and b) and Thépaut and Moll (1990). Several meteorological services are taking steps to directly incorporate radiance measurements in their operational assimilation algorithms. The same general trend is followed for all types of measurements: for instance, radial winds measured by Doppler effect either from ground-based radars or from satellite-borne lidars can be assimilated through an observation operator which reduces to the computation of the wind component along the appropriate direction (see respectively, *e.g.*, Sun *et al.*, 1991, and Courtier *et al.*, 1992). It is presumably the same approach which will be followed for the assimilation of observations made by future observing systems.

7. Conclusions

Our primary purpose in these notes was to describe the principles that lie at the basis of assimilation of observations. These principles are those of statistical linear estimation, and it has been shown that they lead to a generalised least-squares approach, amounting to minimizing a measure of the difference between the available observations and the state to be estimated. The word "observations" must be taken here in a very broad sense, to include all information available in quantitative form, and in particular the equations governing the assimilating model. Two classes of algorithms, sequential and variational assimilation, can be used for actually performing the required computations. Both algorithms are costly, and neither of them can be considered at the present stage as intrinsically superior.

Whatever the algorithm used, it is necessary to specify, for each individual piece of information used in the assimilation, the relationship of that particular piece of information with the variables to be estimated, and the accuracy of that relationship. This amounts to expressing the available information under the general form (3.9), where the matrix Γ must in the most general case be replaced by a nonlinear operator Γ . The latter expresses the relationship between the available information and the variables to be estimated. As for the corresponding accuracy, it is defined, in the basically linear approach described here, by the first and second order statistical moments of the error ζ (mean $E(\zeta)$ and covariances $S = E(\zeta\zeta^T)$). Any assimilation algorithm requires hypotheses, either explicit or implicit, on what Γ and the statistical moments of ζ are.

This leads us to our final remark. Estimating Γ and the statistical moments of ζ is by itself an estimation problem, *a priori* no easier than estimat-

ing the state of the atmospheric flow at a given time. One difference is that those quantities can be estimated by appropriate statistical accumulation. And, as concerns Γ , the physics of the measurement process is of course also fundamental. It has been repeatedly mentioned in these notes that statistical estimation, in addition to the BLUE of the state of the system, produces the covariances of the associated estimation errors. These covariances depend on Γ and Σ (Eq. 3.13). Any disagreement between the predicted analysis-minus-observations differences and the *a posteriori* effectively observed differences must therefore be due to inaccurate estimation of $E(\zeta)$ and/or Σ , and must be usable for improving the corresponding estimates. As for the observation operator Γ , it is in principle possible to determine it as the statistical minimizer of the innovation vector. Work along these lines of *adaptive filtering* has been done by Daley (1992), Dee (1993), Hoang Hong *et al.* (1995) and Blanchet (pers. com.).

References

- Anderson, B.D.O., 1979, *Optimal Filtering*, Prentice Hall, Englewood Cliffs, 357 pp.
- Bennett, A.F., 1992, *Inverse Methods in Physical Oceanography*, Cambridge University Press, Cambridge, United Kingdom, 346 pp.
- Bennett, A.F., L.M. Leslie, C.R. Hagelberg and P.E. Powers, 1993, Tropical Cyclone Prediction Using a Barotropic Model Initialized by a Generalized Inverse Method, *Mon. Wea. Rev.*, **121**, 1714–1729.
- Bergthorsson, P. and B. Döös, 1955, Numerical weather map analysis, *Tellus*, **5**, 329–340.
- Bouttier, F., 1994, A Dynamical Estimation of Forecast Error Covariances in an Assimilation System, *Mon. Wea. Rev.*, **122**, 2376–2390.
- Cohn, S.E. and D.F. Parrish, 1991, The Behavior of Forecast Error Covariances for a Kalman Filter in Two Dimensions, *Mon. Wea. Rev.*, **119**, 1757–1785.
- Courtier, P., P. Gauthier, F. Rabier, P. Flamant, A. Dabas, F. Lieutaud and H. Renault, 1992, *Study of preparation of the use of Doppler wind lidar information in meteorological assimilation systems*, Final report, ESA Contract 8850/90/HGE-I, ESA, Paris.
- Courtier, P. and O. Talagrand, 1987, Variational assimilation of meteorological observations with the adjoint vorticity equation. I : Numerical results, *Q. J. R. Meteorol. Soc.*, **113**, 1329–1347.
- Courtier, P. and O. Talagrand, 1990, Variational assimilation of meteorological observations with the direct and adjoint shallow-water equations, *Tellus*, **42A**, 531–549.
- Courtier, P., J.-N. Thépaut and A. Hollingsworth, 1994, A strategy for operational implementation of 4D-Var, using an incremental approach, *Q. J. R. Meteorol. Soc.*, **120**, 1367–1387.
- Cressman, G.P. 1959, An operational objective analysis scheme, *Mon. Wea. Rev.*, **87**, 367–374.
- Daley, R., 1991, *Atmospheric Data Analysis*, Cambridge Atmospheric and Space Science Series, Cambridge University Press, Cambridge, 457 pp.
- Daley, R., 1992, The Lagged Innovation Covariance : A Performance Diagnostic for Atmospheric Data Assimilation, *Mon. Wea. Rev.*, **120**, 178–196.
- Dee, D.P., 1991, Simplification of the Kalman filter for meteorological data assimilation, *Q. J. R. Meteorol. Soc.*, **117**, 365–384.
- Dee, D.P., 1993, A simple scheme for tuning forecast error covariance parameters, in *Variational assimilation, with special emphasis on three-dimensional aspects*, ECMWF, Reading, England, 191–205.
- Derber, J.C., 1987, Variational four dimensional analysis using quasi-geostrophic constraints, *Mon. Wea. Rev.*, **115**, 998–1008.
- Eliassen, A., 1954, *Provisional report on calculation of spatial covariance and autocorrelation of the pressure field*, Report no 5, Videnskaps-Akademiets Institutt for Vaer-Og Klimaforskning, Oslo, Norway, 12 pp.. Reprinted in Bengtsson, L., M. Ghil and E. Källén, (editors), 1981, *Dynamic Meteorology. Data Assimilation Methods*, Springer Verlag, New York, USA, 330 pp., 319–328.
- Evensen, G., 1992, Using the Extended Kalman Filter with a Multilayer Quasigeostrophic Ocean Model, *J. Geophys. Res.*, **97** (C11), 17,905–17,924.
- Evensen, G., 1994, Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, **99** (C5), 10,143–10,162.
- Eyre, J., 1989a, Inversion of cloudy satellite sounding radiances by nonlinear optimal estimation. I: Theory and simulation for TOVS, *Q. J. R. Meteorol. Soc.*, **115**, 1001–1026.
- Eyre, J., 1989b, Inversion of cloudy satellite sounding radiances by nonlinear optimal estimation. II: Application to TOVS data, *Q. J. R. Meteorol. Soc.*, **115**, 1027–1037.
- Gandin, L.S., 1963, *Objective analysis of meteorological fields*, Gidrometeor. Izd., Leningrad (in Russian), (English Translation by Israel Program for Scientific Translations, Jerusalem, 1965).
- Gaspar, P. and C. Wunsch, 1989, Estimates from Altimeter Data of Barotropic Rossby Waves in the Northwestern Atlantic Ocean, *J. Phys. Oceanogr.*, **19**, 1821–1844.
- Ghil, M., 1989, Meteorological data assimilation for oceanographers. Part I. Description and theoretical framework, *Dyn. Atmos. Oceans*, **13**, 171–218.
- Ghil, M. and P. Malanotte-Rizzoli, 1991, Data assimilation in meteorology and oceanography, *Adv. in Geophys.*, **33**, 141–266.
- Gill, P.E., W. Murray and M.H. Wright, 1982, *Practical Optimization*, Academic Press, London.
- Greiner, E. and C. Perigaud, 1994, Assimilation of Geosat Altimetric Data in a Nonlinear Reduced-Gravity Model of the Indian Ocean. Part 1: Adjoint Approach and Model-Data Consistency, *J. Phys. Oceanogr.*, **24**, 1783–1804.
- Hinkelmann, K., 1951, Der Mechanismus des meteorologischen Lärmes, *Tellus*, **3**, 285–296.
- Hoang Hong S., P. De Mey, O. Talagrand and R. Baraille, *Assimilation of Altimeter Data in a Multilayer Quasi-Geostrophic Ocean Model by Simple*

- Nonlinear Adaptive Filter, Proceedings, International Symposium on Assimilation of Observations in Meteorology and Oceanography*, World Meteorological Organization, Tokyo, Japan, March 1995.
- Jazwinski, A.H., 1970, *Stochastic Processes and Filtering Theory*, Academic Press, New-York, 376 pp.
- Kalman, R.E., 1960, A new approach to linear filtering and prediction problems, *J. Basic Eng.*, **82D**, 35–45.
- Lacarra, J.F. and O. Talagrand, 1988, Short-range evolution of small perturbations in a barotropic model, *Tellus*, **40A**, 81–95.
- Le Dimet, F.X. and O. Talagrand, 1986, Variational algorithms for analysis and assimilation of meteorological observations : theoretical aspects, *Tellus*, **38A**, 97–110.
- Lewis, J.M. and J.C. Derber, 1985, The use of adjoint equations to solve a variational adjustment problem with advective constraints, *Tellus*, **37A**, 309–322.
- Lions, P.L., 1971, *Optimal control of systems governed by partial differential equations*, Springer-Verlag, Berlin, 396 pp.
- Lorenc, A., 1981, A global three-dimensional multivariate statistical interpolation scheme, *Mon. Wea. Rev.*, **109**, 701–721.
- Lorenc, A., 1986, Analysis methods for numerical weather prediction, *Q. J. R. Meteorol. Soc.*, **112**, 1177–1194.
- Lorenc, A., 1988, Optimal nonlinear objective analysis, *Q. J. R. Meteorol. Soc.*, **114**, 205–240.
- Lorenz, E.N., 1963, Deterministic Nonperiodic Flow, *J. Atmos. Sci.*, **20**, 130–141.
- Miller, R.N., M. Ghil and F. Gauthiez, 1994, Advanced Data Assimilation in Strongly Nonlinear Dynamical Systems, *J. Atmos. Sci.*, **51**, 1037–1056.
- Morgenstern, J., 1984, *How to compute fast a function and all its derivatives. A variation on the theorem of Baur-Stressen*, Report No 49, Laboratoire CNRS 168, Université de Nice, Nice, France, 5 pp.
- Navon, I.M., X. Zou, K. Johnson, J. Derber and J. Sela, 1991, *Variational data assimilation with an adiabatic version of the NMC spectral model*, Report FSU-SCRI-91-13, The Florida State University, Tallahassee, Florida, USA, 43pp.
- Oort, A.H., 1989, Angular Momentum Cycle in the Atmosphere-Ocean-Solid Earth System, *Bull. Amer. Meteor. Soc.*, **70**, 1231–1242.
- Parrish, D. and D. Derber, 1992, The National Meteorological Center's Spectral Statistical-Interpolation Analysis System, *Mon. Wea. Rev.*, **120**, 1747–1763.
- Penenko, V.V. and N.N. Obraztsov, 1976, A variational initialization method for the fields of the meteorological elements (English translation), *Soviet Meteorol. Hydrol.*, no 11, 1–11.
- Pires, C., R. Vautard and O. Talagrand, 1995, On extending the limits of variational assimilation in nonlinear chaotic systems, *Tellus*, **48A**, 96–121.
- Rabier, F. and P. Courtier, 1992, Four-dimensional assimilation in the presence of baroclinic instability, *Q. J. R. Meteorol. Soc.*, **118**, 649–672.
- Rabier, F., P. Courtier, J. Pailleux, O. Talagrand and D. Vasiljevic, 1993, A comparison between four-dimensional variational assimilation and simplified sequential assimilation relying on three-dimensional variational analysis, *Q. J. R. Meteorol. Soc.*, **119**, 845–880.
- Salstein, D.A. and R.D. Rosen, 1986, Earth Rotation as a Proxy for Interannual Variability in Atmospheric Circulation, 1860-Present, *J. Climate App. Meteor.*, **25**, 1870–1877.
- Sasaki, Y., 1970, Some basic formalisms in numerical variational analysis, *Mon. Wea. Rev.*, **98**, 875–883.
- Sheinbaum, J. and D.L.T. Anderson, 1990a, Variational Assimilation of XBT Data. Part I, *J. Phys. Oceanogr.*, **20**, 672–688.
- Sheinbaum, J. and D.L.T. Anderson, 1990b, Variational Assimilation of XBT Data. Part II : Sensitivity Studies and Use of Smoothing Constraints, *J. Phys. Oceanogr.*, **20**, 689–704.
- Sun, J., D.W. Flicker and D.K. Lilly, 1991, Recovery of Three-Dimensional Wind and Temperature Fields from Simulated Single-Doppler Radar Data, *J. Atmos. Sci.*, **48**, 876–890.
- Talagrand, O. and P. Courtier, 1987, Variational assimilation of meteorological observations with the adjoint vorticity equation. I : Theory, *Q. J. R. Meteorol. Soc.*, **113**, 1311–1328.
- Tarantola, A., 1987, *Inverse Problem Theory*, Elsevier, Amsterdam, The Netherlands, 613 pp.
- Temperton, C., 1988, Implicit normal mode initialization, *Mon. Wea. Rev.*, **116**, 1013–1031.
- Thacker, W.C. and R.B. Long, 1988, Fitting dynamics to data, *J. Geophys. Res.*, **93**, 1227–1240.
- Thépaut, J.N. and P. Courtier, 1991, Four-dimensional variational data assimilation using the adjoint of a multilevel primitive-equation model, *Q. J. R. Meteorol. Soc.*, **117**, 1225–1254.
- Thépaut, J.-N., R.N. Hoffman and P. Courtier, 1993, Interactions of Dynamics and Observations in a Four-Dimensional Variational Assimilation, *Mon. Wea. Rev.*, **121**, 3393–3414.
- Thépaut, J. N., and P. Moll, 1990, Variational inversion of simulated TOVS radiances using the adjoint technique, *Q. J. R. Meteorol. Soc.*, **116**, 1425–1448.
- Thiébaux, H.J. and M.A. Pedder, 1987, *Spatial Objective Analysis, with Applications in Atmospheric Sciences*, Academic Press, New York, United States.
- Vautard, R., 1990, Multiple Weather Regimes over the North Atlantic: Analysis of Precursors and Successors, *Mon. Wea. Rev.*, **118**, 2056–2081.
- Wahba, G., 1990, *Spline Models for Observational Data*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, USA, 169 pp.

Other Key Reference Papers for Prof. Li's Lecture:

- [1] Han, X., X. Li (2008), An evaluation of the nonlinear/non-Gaussian filters for the sequential data assimilation, *Remote Sensing of Environment* 112:1434–1449
- [2] Donald r. Johnson et al. (1991), *Four-Dimensional Model Assimilation of Data: A Strategy for the Earth System Sciences*, ISBN: 0-309-58348-9, 1991
- [3] Evensen, G. (2003), The Ensemble Kalman Filter: theoretical formulation and practical implementation, *Ocean Dynamics*, 53: 343–367
- [4] Ide, K., P. Courtier, et al. (1997), Unified Notation for data assimilation: operational, sequential and variational, *Journal of the Meteorological Society of Japan*, 75 (1B): 181–189
- [5] 李新, 摆玉龙. 顺序数据同化的Bayes滤波框架. *地球科学进展*, 2010, 25(5): 515–522
- [6] 李新, 黄春林等. 中国陆面数据同化系统研究的进展与前瞻. *自然科学进展*, 2007, 17(2) 163–173

Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal

Jun Qin,¹ Shunlin Liang,² Kun Yang,¹ Ichiro Kaihotsu,³ Ronggao Liu,⁴ and Toshio Koike⁵

Received 25 October 2008; revised 10 March 2009; accepted 3 June 2009; published 5 August 2009.

[1] Soil moisture is a very important variable in land surface processes. Both field moisture measurements and estimates from modeling have their limitations when being used to estimate soil moisture on a large spatial scale. Remote sensing is becoming a practical method to estimate soil moisture globally; however, the quality of current soil surface moisture products needs to be improved in order to meet practical requirements. Data assimilation (DA) is a promising approach to merge model dynamics and remote sensing observations, thus having the potential to estimate soil moisture more accurately. In this study, a data assimilation algorithm, which couples the particle filter and the kernel smoothing technique, is presented to estimate soil moisture and soil parameters from microwave signals. A simple hydrological model with a daily time step is utilized to reduce the computational burden in the process of data assimilation. An observation operator based on the ratio of two microwave brightness temperatures at different frequencies is designed to link surface soil moisture with remote sensing measurements, and a sensitivity analysis of this operator is also conducted. Additionally, a variant of particle filtering method is developed for the joint estimation of soil moisture and soil parameters such as texture and porosity. This assimilation scheme is validated against field moisture measurements at the CEOP/Mongolia experiment site and is found to estimate near-surface soil moisture very well. The retrieved soil texture still contains large uncertainties as the retrieved values cannot converge to fixed points or narrow ranges when using different initial soil texture values, but the retrieved soil porosity has relatively small uncertainties.

Citation: Qin, J., S. Liang, K. Yang, I. Kaihotsu, R. Liu, and T. Koike (2009), Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal, *J. Geophys. Res.*, *114*, D15103, doi:10.1029/2008JD011358.

1. Introduction

[2] Soil moisture plays a significant role in the terrestrial water cycle [Daly and Porporato, 2005; Hirabayashi et al., 2005; Reichle et al., 2007; Sheffield and Wood, 2007]. It is very important to obtain information about soil moisture due to its profound impacts on practical water resource applications such as flood forecasting, weather and climate prediction, crop growth monitoring, and water resource

management [Claussen, 1998; Davies and Allen, 1973; Drusch, 2007; Drusch and Viterbo, 2007; Foley, 1994; Schmugge et al., 2002; Texier et al., 1997]. There are two common methods to obtain the soil moisture status [Moradkhani, 2008]. One is to measure it in the field with instruments. These measurements are merely representative over a small spatial scale since the soil moisture has large spatial heterogeneity. It is not practical to densely install many instruments on a large scale. The other is to simulate soil moisture by running land surface models (LSMs) with meteorological data and other parameters as inputs. The simulated soil moisture performs well when both the model parameters and meteorological forcing are known with a high degree of precision and accuracy. This can be realized at only a very limited number of sites, where a variety of measurement instruments are installed. When running the model on a large scale, it is very difficult to accurately obtain model inputs and parameter values.

[3] Microwave remote sensing data has offered another means to map land surface soil moisture on a large scale [Kerr et al., 2001; Njoku et al., 2003; Wagner et al., 2003]. However, it also has many limitations and thus mapping

¹Key Laboratory of Tibetan Environment Changes and Land Surface Processes, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China.

²Department of Geography, University of Maryland, College Park, Maryland, USA.

³Department of Natural and Environmental Sciences, Faculty of Integrated Arts and Sciences, Hiroshima University, Hiroshima, Japan.

⁴Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China.

⁵Department of Civil Engineering, School of Engineering, University of Tokyo, Tokyo, Japan.

results cannot satisfy the practical requirements. Data assimilation (DA) methods can consistently couple both modeling and observations and thus yield superior soil moisture retrievals [Entekhabi *et al.*, 1994; Galantowicz *et al.*, 1999; Houser *et al.*, 1998; Margulis *et al.*, 2002; McLaughlin, 2002; Reichle *et al.*, 2001; Walker and Houser, 2001]. Thus it has attracted much attention from researchers in many fields.

[4] Data assimilation techniques can generally be divided into two categories: sequential-based and cost-function-based methods. Sequentially based methods [Bertino *et al.*, 2003], especially those based on a Monte Carlo approach such as Ensemble Kalman Filtering (EnKF [Evensen, 2003]), are currently popular in land data assimilation research and applications since they can be applied to nonlinear and discontinuous models and be realized easily [Huang *et al.*, 2008a, 2008b]. This is especially important in land surface data assimilation because there are many land surface parameterizations which are not continuous or not differentiable and this makes it difficult or inefficient to use cost-function-based methods as assimilation algorithms. In addition, the cost-function-based method cannot directly consider uncertainties in atmospheric inputs which are used to drive a LSM. It just treats uncertainties included in inputs as one part of model noise. These uncertainties can, however, be handled readily in the sequential methods [Liang and Qin, 2008].

[5] Reichle *et al.* [2002] used EnKF to assimilate L-band (1.4 GHz) microwave brightness temperature observations into a LSM. Their research indicated that the EnKF is a flexible and robust DA option that gives satisfactory estimates even for moderate ensemble sizes although the updating process is suboptimal. Crow [2003] and Crow and Wood [2003] applied EnKF to assimilate L-band microwave data to correct for the impact of poorly sampled rainfall on land surface modeling of root-zone soil moisture and surface energy fluxes. The results suggested that the EnKF-based assimilation system is capable of correcting a substantial fraction of model errors in root-zone soil moisture and latent heat flux predictions associated with the use of temporally sparse rainfall measurements as the forcing data. Ni-Meister *et al.* [2006] assimilated retrieved soil surface moisture from Scanning Multichannel Microwave Radiometer (SMMR) data using EnKF. Reichle *et al.* [2007] applied EnKF to assimilate retrieved soil surface moisture from the Advanced Microwave Sounding Radiometer-Earth Observing System (EOS) (AMSR-E) as observations into a LSM. Comparisons were also performed between EnKF and other Monte Carlo-based filtering methods [Zhou *et al.*, 2006].

[6] Most of the studies mentioned above assimilated microwave brightness into a LSM with an hourly or sub-hourly time step rather than with a daily time step. It is because the microwave radiative transfer equation (RTE) as the observation operator requires instantaneous soil surface and canopy temperatures as inputs, which have the apparent diurnal variations, but the daily-based model lacks such a temporal resolution. It is obvious that significant computational cost could be saved if a daily-based model is used in the process of assimilation. Furthermore, most of previous land surface assimilation studies focus on either retrieving state variables such as soil moisture or estimating some model

parameters independently. Moreover, some aforementioned studies assimilated L-band brightness temperature, which has not been available for large spatial regions. Other studies assimilated retrieved soil surface moisture from AMSR-E and/or SMMR data at a continental scale. In addition, all these studies assumed that soil texture data or hydraulic properties are available, although it is rather difficult to obtain their accurate values at a large scale. Few investigations [Moradkhani *et al.*, 2005a; Yang *et al.*, 2007] put forward the idea of jointly retrieving state variables and model parameters, and perform assimilation experiments using a conceptual rainfall-runoff model and a complex LSM, respectively.

[7] In this study, a simple model is used to characterize the water movement in soils with a daily-based time step, an observation operator is designed to link the AMSR-E microwave signal and soil surface moisture, and a variant of particle filtering method is used to simultaneously estimate soil surface moisture and soil parameters such as texture, and porosity, and surface parameters. Then, the whole DA scheme is validated against the field measurements. In this paper, the DA scheme is first described. Validation results are then presented and finally followed by discussions and conclusions.

2. Data Assimilation Scheme

[8] A DA system consists of four parts: model dynamics, observation operator, assimilation algorithm, and error models [Lermusiaux and Robinson, 2001]. In the following subsections, details of these four parts are presented.

2.1. Land Surface Water Balance Model

[9] A land surface scheme to model the water balance on a daily basis is simplified from the Simple Biosphere Model 2 (SiB2 [Sellers *et al.*, 1996]). We aim to develop the data assimilation system mainly for arid or semiarid areas; thus the interception storage of the canopy can be ignored, since the leaf area index (LAI) normally peaks around 1.5. The soil column is vertically divided into three layers: surface layer, root zone layer, and recharge zone. The governing equations characterizing the water movement in the soil are as follows:

$$\frac{\partial \theta_1}{\partial t} = \frac{1}{D_1} \left[P_t - Q_{1,2} - \frac{1}{\rho_w} E_g \right], \quad (1)$$

$$\frac{\partial \theta_2}{\partial t} = \frac{1}{D_2} \left[Q_{1,2} - Q_{2,3} - \frac{1}{\rho_w} E_{tr} \right], \quad (2)$$

$$\frac{\partial \theta_3}{\partial t} = \frac{1}{D_3} [Q_{2,3} - Q_3], \quad (3)$$

where θ_i is volumetric soil moisture content of each layer, D_i the soil thickness of each layer, P_t is the precipitation, $Q_{1,2}$, $Q_{2,3}$, and Q_3 are soil water fluxes between layers and out of the bottom layer, E_g and E_{tr} are evaporation from the soil surface and transpiration from the vegetation canopy, respectively, and ρ_w is the water density. Equations (1)–(3)

are discretized on a daily basis using an implicit difference scheme. The formulas for soil water flux are as follows:

$$Q_{i,i+1} = \left(\frac{\psi_i - \psi_{i+1}}{0.5(D_i + D_{i+1})} + 1 \right) \left(\frac{K_i \psi_i - K_{i+1} \psi_{i+1}}{\psi_{i+1} - \psi_i} \right) \left(\frac{B}{B+3} \right), \quad (4)$$

$$\psi_i = \psi_{sat} \left(\frac{\theta_i}{\theta_{sat}} \right)^{-B}, \quad (5)$$

$$K_i = K_{sat} \left(\frac{\theta_i}{\theta_{sat}} \right)^{2B+3}, \quad (6)$$

$$K_{sat} = 7.0556 \cdot 10^{-6.884+0.0153 \cdot \%sand} \quad (7)$$

$$B = 2.91 + 0.159 \cdot \%clay \quad (8)$$

where K_i is the hydraulic conductivity of each layer, ψ_i the matrix potential of each layer, θ_{sat} the soil porosity, K_{sat} the hydraulic conductivity at saturation, and B the empirical parameter related to soil texture, %sand the sand content, and %clay the clay content. The drainage out of the bottom layer is assumed to be K_3 and the surface runoff occurs when surface soil water content θ_1 exceeds the porosity θ_{sat} .

[10] Both the evaporation E_g from the soil surface and transpiration E_{tr} from the vegetation canopy are the important components in equations (1)–(3). There exist many methods to compute the potential evapotranspiration on a daily basis, including Penman-Monteith, Priestley-Taylor, and so on. The actual evapotranspiration and its partition into evaporation and transpiration are needed in the calculation of the water balance. In this study, a variant of the Priestley-Taylor equation [Davies and Allen, 1973] is taken to estimate the daily actual evapotranspiration and the vegetation coverage is used to separate it. The daily evapotranspiration process is parameterized as follows [Sau et al., 2004]:

$$ET_a = \alpha \frac{\Delta}{\Delta + \gamma} R_n \cdot \left\{ 1 - \exp \left[\beta \left(\frac{\theta_1}{\theta_{sat}} \right)^3 \right] \right\}, \quad (9)$$

$$E_g = ET_a \cdot (1 - f_v), \quad (10)$$

$$E_{tr} = ET_a \cdot f_v, \quad (11)$$

$$f_v = 1 - \exp(-0.5LAI), \quad (12)$$

where ET_a denotes the actual evapotranspiration, $\alpha = 1.26$ the Priestly-Taylor constant, β the constraint coefficient, Δ the slope of the saturated vapor pressure with respect to the air temperature, γ the psychrometric constant, and f_v the vegetation coverage. The power 3 does not exist in its

original form of equation (9) and it is found that equation (9) performs better after adding an exponent of 3 [Nakayama et al., 1993].

[11] An implicit scheme is used for the computation with daily time step and is stable, but big errors may occur immediately after a rainfall event. Nevertheless, such errors may be compensated to some degree by information from satellite signals.

2.2. Observation Operator and Microwave Data

[12] In this study, a microwave RTE is implemented to link the surface soil moisture to satellite measurements. It is a Q-h model with minor revisions to include vegetation effect. The concrete form of this RTE is as follows:

$$T_{bp} = T_g (1 - \Gamma_p) \exp(-\tau_c) + T_c (1 - \omega) \cdot [1 - \exp(-\tau_c)] [1 + \Gamma_p \exp(-\tau_c)], \quad (13)$$

where the subscript p represents the vertical or horizontal polarization, Γ_p the soil reflectivity, τ_c the vegetation optical depth, and ω the vegetation single scattering albedo. A Q-h model is used to calculate the soil reflectivity as follows:

$$\Gamma_p = [(1 - Q)R_p + QR_q] \exp(-h), \quad (14)$$

where the subscripts p denotes the vertical or horizontal polarization, respectively, Q and h are empirical surface roughness parameters, and R the Fresnel power reflectivity with a smooth soil surface. The R is determined using the following equations:

$$R_p = \left| \frac{\cos \gamma - \sqrt{\varepsilon_r - \sin^2 \gamma}}{\cos \gamma + \sqrt{\varepsilon_r - \sin^2 \gamma}} \right|^2, \quad (15)$$

$$R_q = \left| \frac{\varepsilon_r \cos \gamma - \sqrt{\varepsilon_r - \sin^2 \gamma}}{\varepsilon_r \cos \gamma + \sqrt{\varepsilon_r - \sin^2 \gamma}} \right|^2, \quad (16)$$

where p and q denote the horizontal and vertical polarization, γ the incident angle, and ε_r the soil dielectric constant. The soil dielectric constant is computed as follows:

$$\varepsilon_r = \left[1 + (1 - \theta_{sat})(\varepsilon_s^a - 1) + \theta_1^b \varepsilon_{fw}^a - \theta_1 \right]^{1/a}, \quad (17)$$

where $\varepsilon_s = 4.7 + 0.0j$ denotes the dielectric constant for mineral soil, ε_{fw} the dielectric constant of free water, $a = 0.65$, and b the coefficient dependent upon the soil texture. The parameters in equations (13) and (14) are dependent on wave frequency and can be parameterized as follows:

$$h = (k \cdot s) \sqrt{0.1 \cos \gamma}, \quad (18)$$

$$Q = Q_0 (k \cdot s)^{0.795}, \quad (19)$$

$$\tau_c = \frac{b'(100\lambda)^\chi w_c}{\cos \theta}, \quad (20)$$

$$w_c = \exp(LAI/3.3) - 1 \quad (21)$$

$$\omega = \frac{0.00083}{\lambda}, \quad (22)$$

where λ [m] is the wavelength, k the wave number defined as $2\pi/\lambda$, s the standard deviation of surface roughness, w_c [kg m⁻²] the vegetation water content, and Q_0 , b' , and χ the empirical coefficients. Equations (19) and (22) are empirical formulas fitted from limited microwave experimental data [Fujii, 2005] and were firstly used by Yang *et al.* [2007] to reduce the parameter number of the radiative transfer equation.

[13] As shown in equation (13), both the soil surface temperature and the canopy temperature are important factors to determine the value of the simulated microwave brightness temperature. Since the time step is one day in this model, the brightness temperature cannot be simulated at the satellite overpassing times. So it is challenging to directly assimilate the satellite-observed brightness temperature into the dynamics. However, this dilemma can be removed by defining a ratio of two brightness temperatures at different frequencies, which can reflect the soil wetness, based on the assumption that the soil surface temperature and the canopy temperature have equal values. When AMSR-E data are used as the information source to be assimilated, a new index called the soil water ratio (SWR) is defined in this study in accordance with equation (13) by canceling out temperatures as follows:

$$SWR = \frac{T_{bq}^{18.7}}{T_{bq}^{6.9}} = \frac{\{(1 - \Gamma_q) \exp(-\tau_c) + (1 - \omega)[1 - \exp(-\tau_c)][1 + \Gamma_q \exp(-\tau_c)]\}^{18.7}}{\{(1 - \Gamma_q) \exp(-\tau_c) + (1 - \omega)[1 - \exp(-\tau_c)][1 + \Gamma_q \exp(-\tau_c)]\}^{6.9}} \quad (23)$$

where the superscript denotes the frequency and $T_{bq}^{18.7}$ and $T_{bq}^{6.9}$ the vertical polarization brightness temperatures at 18.7 GHz and 6.9 GHz. The reason for choosing the vertical polarization temperatures in (23) is that they are less sensitive to vegetation heterogeneity than the horizontal polarization temperatures. Up to now, there are still two problems not to be resolved. One is that SWR does not completely eliminate the influence of temperature since Γ_q in equation (23) is a function of the dielectric constant of free water which is in turn dependent on the soil temperature. It is, however, found that SWR is not sensitive to the soil surface temperature and thus this temperature can be replaced with the daily averaged air temperature. The other is the issue of whether the defined SWR really has the capacity to reflect the soil wetness. The solutions to these two problems will be shown in section 4.

2.3. Assimilation Algorithm

[14] Mainstream sequential-based methods include Kalman Filtering (KF) and Particle Filtering (PF), and their variants. PF is also called sequential Monte Carlo filtering.

It has been applied in many engineering fields and attracted some data assimilation practitioners since the posterior distribution of the state vector can be represented with Monte Carlo samples and the Gaussian assumption can be avoided. KF and its variants, however, just evaluate the mean and covariance of the posterior distribution. Thus PF can better grasp the filtering density evolution of the nonlinear system in time than KF and its variants do. PF itself also has many variants such as the sampling importance resampling filter (SIR). Han and Li [2008] make a detailed evaluation of PF, KF, and their variants and conclude that PF is suited to applications in land surface data assimilation according to both effectiveness and efficiency. Before moving on, some notations are introduced to facilitate the following discussions.

$$x_{t+1} = f(x_t, u_t, \xi) + v_t \quad (24)$$

where x denotes the model state vector, u the external forcing data, ξ the model parameter vector, v the model noise, t the subscript for time step, and $f(\cdot)$ the model operator mapping the previous state x_t to the next state x_{t+1} .

$$y_t = h(x_t, \xi) + e_t \quad (25)$$

where y denotes the observation vector, e the observation noise, and $h(\cdot)$ the observation operator. In this study, the state vector $x = [\theta_1, \theta_2, \theta_3]^T$, external forcing $u = [Pt, Rn, Ta, LAI]^T$, $\xi = [\%sand, \%clay, \theta_{sat}, \beta, s, Q_0, b', \chi]^T$, $f(\cdot)$ is the discrete form from equations (1)–(3), and $h(\cdot)$ is the equation (23).

[15] In the framework of sequential filtering techniques, the joint estimation of state variables and model parameters can be performed through the state augmentation method

[Chen *et al.*, 2005]. This approach regards model parameters to be estimated as part of the state vector. The new augmented state vector becomes $[x^T, \xi^T]^T$. Conventionally, the random walk model is assumed for the time evolution of ξ [Moradkhani *et al.*, 2005b]. However, this model results in much larger variances of ξ than actual ones in the estimation process. The kernel smoothing technique is currently introduced to remove this feature.

[16] In this study, both the SIR method and kernel smoothing technique are combined [Chen *et al.*, 2005] to merge the dynamics and the observations for simultaneous estimation of the model states and parameters. Main steps for the entire assimilation algorithm [Thomas, 2006] are as follows:

Step 1: for $t = 0$, sample $\{\tilde{\xi}_0^{(i)}\}_{i=1}^N : p(\xi_0)$ and $\{\tilde{x}_0^{(i)}\}_{i=1}^N : p(x_0)$;

Step 2: draw $\{\xi_{t+1}^{(i)}\}_{i=1}^N : N(\xi_{t+1}|m_t^{(i)}, h^2 \cdot V_t)$, where $m_t^{(i)} = (\sqrt{1 - \bar{r}^2})\tilde{\xi}_t^{(i)} - (1 - \sqrt{1 - \bar{r}^2})\tilde{\xi}_t^{(i)}$, V_t denotes the covariance matrix of ξ_t , and s an adjustable parameter;

Step 3: draw $\{x_{t+1}^{(i)}\}_{i=1}^N : p(x_{t+1}|\tilde{x}_t^{(i)}, \xi_{t+1}^{(i)})$, where $x_{t+1}^{(i)} = f(\tilde{x}_t^{(i)}, u_t^{(i)}, \xi_{t+1}^{(i)}) + v_t^{(i)}$, $v_t^{(i)} : p(v_t)$;

Step 4: compute weights $w_{t+1}^{(i)} = \frac{p(y_{t+1}|x_{t+1}^{(i)}, \xi_{t+1}^{(i)})}{\sum_{i=1}^N p(y_{t+1}|x_{t+1}^{(i)}, \xi_{t+1}^{(i)})}$, where $p(y_{t+1}|x_{t+1}^{(i)}, \xi_{t+1}^{(i)})$ denotes the value of $p(y_{t+1} | (x_{t+1}^{(i)}, \xi_{t+1}^{(i)}))$;

Step 5: resample $\{x_{t+1}^{(i)}, \xi_{t+1}^{(i)}\}_{i=1}^N$ with replacement according to weights $w_{t+1}^{(i)}$ in order to get $\{\tilde{x}_{t+1}^{(i)}, \tilde{\xi}_{t+1}^{(i)}\}_{i=1}^N$ with weights $\{1/N\}_{i=1}^N$;

Step 6: set $t = t + 1$ and go to step 2.

[17] The parameter l is an adjustable constant for considering that the model parameters to be estimated are to change quickly or slowly. l is set to $0 < l < 0.2$ for slowly varying parameters and $0.8 < l < 1.0$ for rapidly varying parameters. In this work, only variances of each model parameter are computed and nondiagonal elements are set to zero for the covariance matrix V_t in Step 2 above. Thus different values of l can be easily set independently for determination of different parameters.

2.4. Error Models

[18] One significant advantage of sequential-based data assimilation methods is that uncertainties, which are from model structure, model parameters, and inputs, etc., can be handled explicitly in their own framework. As shown in above subsections, there are four error sources, including errors in the model dynamics, observations, input forcing, and model parameters, respectively. All of these errors can be divided into biased and unbiased noises. Only the unbiased part in these errors is taken into account in this study. The error model must be specified for each error source.

[19] As summarized by Hamill [2006], four methods can be applied to parameterize uncertainties in model dynamics. The so-called covariance inflation approach is used and coupled with the assimilation algorithm. Before assimilating observation information into dynamics, deviations of particles around their mean are inflated by a factor r , which is a little bit greater than 1.0, as follows:

$$x_t^{(i)} \leftarrow r \left(x_t^{(i)} - \overline{x_t^{(i)}} \right) + \overline{x_t^{(i)}} \quad (26)$$

where the operation \leftarrow denotes the replacement of the previous value of $x_t^{(i)}$. It is found that a moderate inflation improves the assimilation accuracy. More details can be obtained in the work of Hamill [2006]. One advantage of sequential data assimilation methods over variational ones is that errors in input forcing can be explicitly considered by adding perturbations to them according to some error parameterizations. In this work, error models for inputs are taken as the following uniform probability distribution function (PDF):

$$u_t^{(i)} = u_t(1.0 + \zeta) \quad \zeta : U[-\delta, +\delta] \quad (27)$$

where ζ obeys the uniform distribution and δ reflects the knowledge of inputs. Different values of δ can be assigned to each component in $u = [Pt, Rn, Ta, LAI]^T$. Similarly, uniform distributions are also assumed for the parameter error models. The observation error model is assumed to be Gaussian.

3. Sensitivity Study of Soil Water Ratio

[20] As mentioned in section 2.2, two issues with SWR need to be addressed. As for the first problem, the variance-based sensitivity analysis method is applied to determine the global sensitivity of SWR to each input parameter in equation (23). The variance-based sensitivity method is briefly introduced below [Helton et al., 2006].

[21] The entire SWR formula can be denoted by $Y = h(X)$ where Y means SWR, X the input parameters, and $h(\cdot)$ the equation (23), as given in the previous section. If each component of vector $X = [x_1, x_2, \dots, x_{nX}]^T$ is considered to be an independent random variable, then the variance V_Y of Y can be decomposed and expressed as:

$$V_Y = \sum_i V_i + \sum_{i < j} V_{ij} + \sum_{i < j < k} V_{ijk} + \dots + V_{12\dots nX} \quad (28)$$

where V_i is the contribution of x_i to V_Y , V_{ij} the contribution of the interaction of x_i and x_j to V_Y , and so on up to $V_{12\dots nX}$, which is the contribution of the interaction of x_1, x_2, \dots, x_{nX} to V_Y . Two types of sensitivity indices can be defined as:

$$S_i = \frac{V_i}{V_Y} \quad (29)$$

and

$$S_{Ti} = \frac{V_Y - V_{\sim i}}{V_Y}$$

where $V_{\sim i}$ is the sum of all variance terms which do not include the index i . S_i is the first-order sensitivity index for the i th parameter. This index characterizes the main influence of parameter x_i on the output variable Y and measures the variance reduction that would be achieved by fixing that parameter. S_{Ti} is the total sensitivity index for the i th parameter and measures the sum of all effects related to this parameter, considering the interaction between the i th parameter and other ones.

[22] The Monte Carlo based method can be used to evaluate these indices. The computation steps are as follows:

Step 1: generate two sets of random samples according to given distributions to inputs X

$$X^{(a)} = [x_{i1}^{(a)}, x_{i2}^{(a)}, \dots, x_{inX}^{(a)}], \quad i = 1, 2, \dots, nS \quad (30)$$

and

$$X^{(b)} = [x_{i1}^{(b)}, x_{i2}^{(b)}, \dots, x_{inX}^{(b)}], \quad i = 1, 2, \dots, nS$$

Table 1. Distributions of Input Parameters to RTE and Their Sensitivity Indices to SWI

	T_a	θ_1	%sand	%clay	θ_{sat}	s	Q_0	b'	χ	LAI
Distribution	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform
Units	[K]	[m ³ /m ³]	[–]	[–]	[m ³ /m ³]	[m]	[–]	[–]	[–]	[m ² /m ²]
Min	280.0	0.05	30	10	0.2	2.80E–4	0.6	4.0	–1.6	0.0
Max	310.0	0.55	65	35	0.7	3.50E–4	0.9	6.0	–1.2	2.0
First-order sensitivity indices S_i	0.005	0.32414	0.005	0.005	0.005	0.005	0.005	0.005	0.1098	0.41488
Total sensitivity indices S_{Ti}	0.005	0.36938	0.005	0.005	0.005	0.005	0.005	0.058	0.15385	0.53896

in which nS is the number of samples.

Step 2: estimate the sample mean and variance of Y as

$$\hat{E}_Y = \frac{\sum_{i=1}^{nS} \hat{h}(X_i^{(a)})}{nS} \quad (31)$$

and

$$\hat{V}_Y = \frac{\sum_{i=1}^{nS} \hat{h}^2(X_i^{(a)})}{nS} - \hat{E}_Y^2$$

Step 3: calculate some intermediate parameters as

$$\hat{V}_i = \frac{1}{nS} \sum_{p=1}^{nS} \hat{h}(x_{p(\sim i)}^{(a)}, x_{pi}^{(a)}) \hat{h}(x_{p(\sim i)}^{(b)}, x_{pi}^{(b)}) - \hat{E}_Y^2 \quad (32)$$

and

$$\hat{V}_{\sim i} = \frac{1}{nS} \sum_{p=1}^{nS} \hat{h}(x_{p(\sim i)}^{(a)}, x_{pi}^{(a)}) \hat{h}(x_{p(\sim i)}^{(a)}, x_{pi}^{(b)}) - \hat{E}_Y^2$$

Step 4: evaluate sensitivity indices as follows:

$$S_i = \hat{V}_i / \hat{V}_Y \quad (33)$$

and

$$S_{Ti} = (1 - \hat{V}_{\sim i} / \hat{V}_Y) \hat{V}_Y$$

The sensitivity of all input parameters in equation (23) to SWR can be evaluated according to the above global sensitivity analysis algorithm. The distributions of these input parameters are listed in Table 1. Since there is no information on these parameters, the uniform distribution is assumed for them. The sensitivity analysis results are shown in Table 1. This sensitivity analysis answers two questions raised in section 2.2. It is found that three of the most sensitive input parameters in the calculation of SWI are leaf area index, soil surface moisture, and χ in equation (20), respectively. Other parameters merely have very small effects on SWI, including the soil surface temperature. Thus it is reasonable to construct SWI as the assimilated data.

4. Determination of Model Parameters

[23] There are a total of eight model parameters $\xi = [\%sand, \%clay, \theta_{sat}, \beta, s, Q_0, b', \chi]^T$ in the model operator and observation operator. Not all of the parameters can be estimated in terms of the sensitivity analysis performed above through the data assimilation algorithm presented in section 2.3. A careful analysis is needed to determine which parameters should be retrieved together with model states.

[24] Soil parameters %sand, %clay, and θ_{sat} have very small influences on SWI, but they highly affect the soil water movement as shown in equations (1)–(8) and then influence the value of SWI at the subsequent instant through surface soil moisture θ_1 . Thus it is possible to retrieve %sand, %clay, and θ_{sat} step-by-step in time. In fact, there exist parameterization schemes to estimate θ_{sat} from the soil texture. However, they are not used in this study since these schemes lack sufficient accuracy and therefore θ_{sat} is independently estimated.

[25] Parameters s , Q_0 , and b' have no apparent impacts on SWI in accordance with sensitivity analysis results. At the same time, they are also not similar to %sand, %clay, and θ_{sat} which can influence the subsequent soil surface moisture and in turn the SWI. So their values can be fixed to median ones between maximum and minimum values as listed in Table 1. Since χ can affect the SWI to some degree, it needs to be estimated.

[26] Many investigations [Castellvi *et al.*, 2001; Kustas *et al.*, 1996] indicate that the Priestly-Taylor constant varies in different situations and does not always keep the value of 1.26. Some researchers present modifications to the original Priestly-Taylor formula. Equation (9) is one of them, which introduces the soil water content as a limiting factor. The original value of the parameter β is estimated to be -10.563 by fitting equation (9) to some data sets. However, it is not guaranteed that this value of β can be applicable to other cases. Thus it may improve the application scope by taking β as a parameter to be estimated.

5. Experiment Site and Data

[27] The CEOP/Mongolia experiment site located at Mandalgobi of Mongolia covers a flat area of $120 \text{ km} \times 160 \text{ km}$, where 12 long-term Automatic Stations for Soil Hydrology (ASSH) and 6 Automatic Weather Stations (AWS) are deployed. Their geographic locations are shown in Figure 1. The detailed description of this experiment can be found in the work of Kaihotsu [2005]. In this study, the experimental period (a total of 153 days), was chosen from 1 May 2003 to 30 September 2003 because the soil freezing and thawing processes are not parameterized in the system.

[28] At ASSH, soil temperature and water content were measured at 3 cm and 10 cm depth. Meteorological parameters were observed at AWS, including wind, temperature, humidity, pressure, precipitation, net radiation, and soil temperature and moisture profiles. However, two AWS (TDS and CRS) data sets are not archived in the CEOP project. In this article, near-surface water content data comes from 12 ASSH and 4 AWS (BTS, DGS, DRS, and MGS). The meteorological forcing data with a temporal resolution of 3 hours, including precipitation, net radiation,

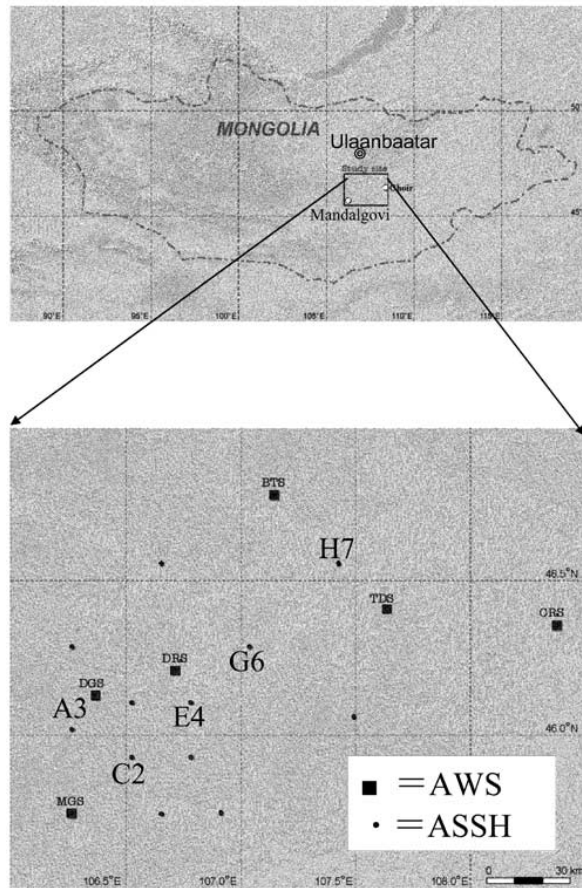


Figure 1. Schematic of the experiment site and locations of 6 AWS and 12 ASSH at CEOP/Mongolia.

and near-surface air temperature, were extracted from the Global Land Data Assimilation System (GLDAS [Rodell *et al.*, 2004]). The microwave data used in this study are the brightness temperatures of AMSR-E 6.9 GHz and 18.7 GHz vertical polarization. The MODIS LAI product was chosen as input to the dynamic model in this paper.

[29] Since GLDAS meteorological data has a spatial resolution of $1^\circ \times 1^\circ$, roughly matching the experiment site area, the mean of 16 soil moisture measurements was used to represent the integral soil water status and compared to the retrieved values through the data assimilation system built in section 3. In addition, the area-average values of brightness temperatures and LAIs are also calculated for the same region. As discussed above, the time step of the dynamics is daily. Consequently, daily averaging is performed on the *in situ* soil moisture observations and GLDAS meteorological data.

6. Results and Discussions

[30] Numerical experiments are performed to validate results from this assimilation system against moisture estimates from a reference run and from other methods. Furthermore, issues about the parameter estimation are also addressed. Before moving on, concrete settings for uncertainties used in the following numerical experiments are

AWS

No.	Site code	Lat. (N)	Lon. (E)	Alt. (m)
1	MGS	45° 44'34.9"	106° 15'52.2"	1393
2	DRS	46° 12'31.2"	106° 42'53.0"	1297
3	DGS	46° 07'38.3"	106° 22'06.8"	1409
4	BTS	46° 46'35.4"	107° 08'32.2"	1371
5	TDS	46° 24'22.4"	107° 38'03.5"	1365
6	CRS	46° 21'08.0"	108° 22'30.5"	1287

ASSH

No.	Site code	Lat. (N)	Lon. (E)	Alt. (m)
1	E2	45° 55'22.5"	106° 46'47.2"	1422
2	A3	46° 00'46.2"	106° 15'52.1"	1502
3	E4	46° 06'10.0"	106° 46'47.2"	1318
4	G6	46° 16'57.6"	107° 02'13.1"	1350
5	GUS	46° 03'14.2"	107° 29'20.3"	1472
6	H7	46° 33'08.9"	107° 25'22.0"	1383
7	D0	45° 44'23.4"	106° 39'05.5"	1342
8	F0	45° 44'34.9"	106° 54'30.2"	1332
9	A6	46° 16'57.6"	106° 15'52.1"	1407
10	C2	45° 55'22.5"	106° 31'21.2"	1422
11	C4	46° 06'10.0"	106° 31'21.2"	1383
12	D7	46° 33'08.9"	106° 39'04.2"	1357

There were 6 Automatic Weather Stations (AWS), 12 Automatic Stations for Soil Hydrology (ASSH). Data at two AWS (TDS, CRS) were not archived in CEOP.

presented, although their parameterization forms are introduced in the previous sections. PDFs for initial moistures $[\theta_1^0, \theta_2^0, \theta_3^0]$ in three soil layers are all subject to the same uniform distribution $U[0.05, 0.55]$. Initial PDFs for parameters to be estimated are listed in Table 1. The parameter δ is set to 0.4 in this study. Controlling constants l for slowing varying parameter and rapidly one are set to 0.12 and 0.9, respectively. The parameter r controlling inflation of particles is set to 1.01. The observation noise e conforms to the normal distribution $N(0, 0.008)$. In the reference run, all parameters are set as median values in their respective ranges.

6.1. Comparison With Other Moisture Retrievals

[31] As many investigations have shown, the size of ensembles or particles has a significant impact on the final assimilation results when Monte Carlo-based assimilation methods are applied. When relatively small particle size is used in the assimilation computation, the results display instability and contain large sampling errors although the computation burden can be reduced. Conversely, the assimilation results keep stable and their errors decline when large particle size is used. The computational intensity is acceptable for assimilation experiments at some sites when large particle size is selected, but this can be unaffordable for applications at a large scale. Thus the balance point needs to

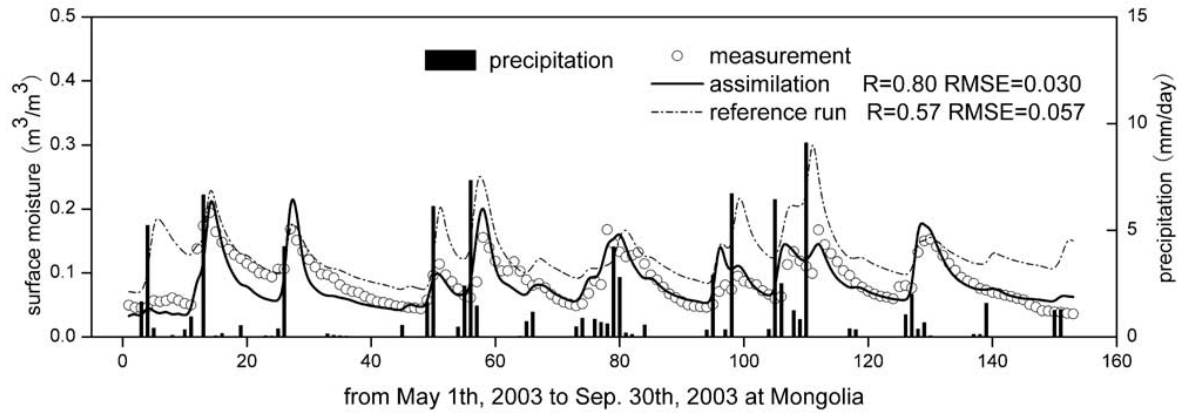


Figure 2. Comparison of computing results from assimilation run and reference run with field measurements. Correlation coefficients denoted by R for assimilation run and reference run are 0.80 and 0.57, respectively. RMSE are 0.030 and 0.057 for assimilation run and reference run, respectively.

be found for the particle size in the assimilation practice, around which both the computational cost and assimilation precision are all acceptable. For this purpose, a series of trials are performed using different particle sizes in accordance with assimilation settings in preceding sections. The analysis indicates that a particle size of 500 is large enough to achieve stable results. In the following, all experiments and analyses are conducted based upon a 500-particle size.

[32] In order to verify the effectiveness of the data assimilation algorithm presented above, the reference run is performed and then resultant soil surface moistures are compared with ones from the assimilation system in Figure 2. GLDAS precipitation data is also depicted and represented by black bars in Figure 2. As have been shown, the assimilation algorithm obviously improves the simulation results in comparison with ones from the reference run. RMSE drops from 0.057 to 0.033 and correlation coefficient R rises from 0.57 to 0.72, respectively. The assimilation results do better agree with field measurements than the reference ones do. This can be explained in accordance with

the following uncertainty sources. The first is the uncertainty in the meteorological forcing, especially in precipitation as shown in Figure 3. GLDAS data grasps main precipitation events during the data assimilation period through the analysis of observed soil surface water content denoted by blank circles in Figure 2. However, the amount of precipitation in each precipitation event could contain large errors, causing the reference run to poorly model the surface soil moisture, but the response of the soil moisture to the precipitation event looks reasonable. Other forcing data also contain more or less uncertainties. The second source is the model structure itself and the third is uncertainties included in model parameters. Effects of these three sources can be alleviated in the framework of the current assimilation algorithm due to the constraint of microwave remote sensing information. Thus assimilation results are superior to those in the reference run.

[33] Figure 4 shows the assimilation results compared with NASA AMSR-E standard surface soil moisture products. Both NASA AMSR-E daytime and nighttime products

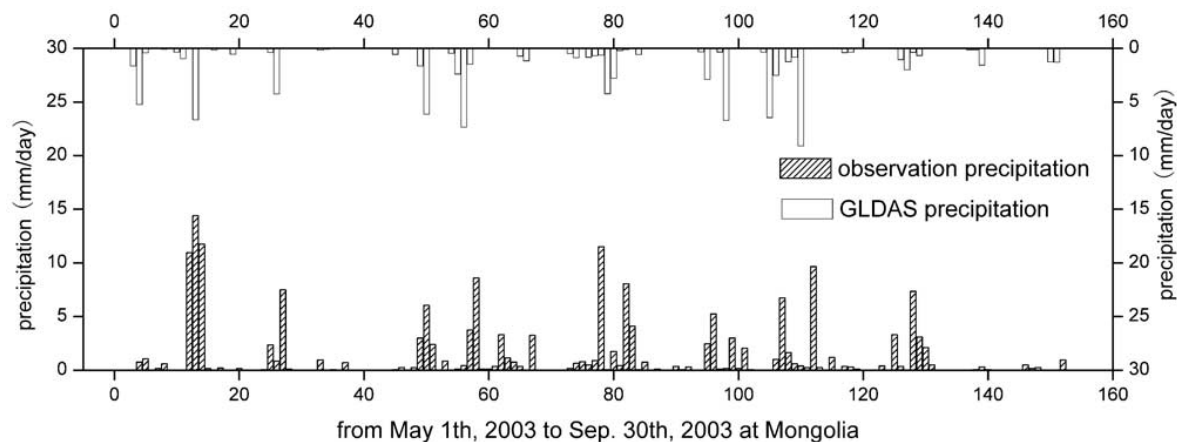


Figure 3. Comparison between observed precipitation and GLDAS precipitation at the experiment site. Large uncertainties are included in GLDAS precipitation data.

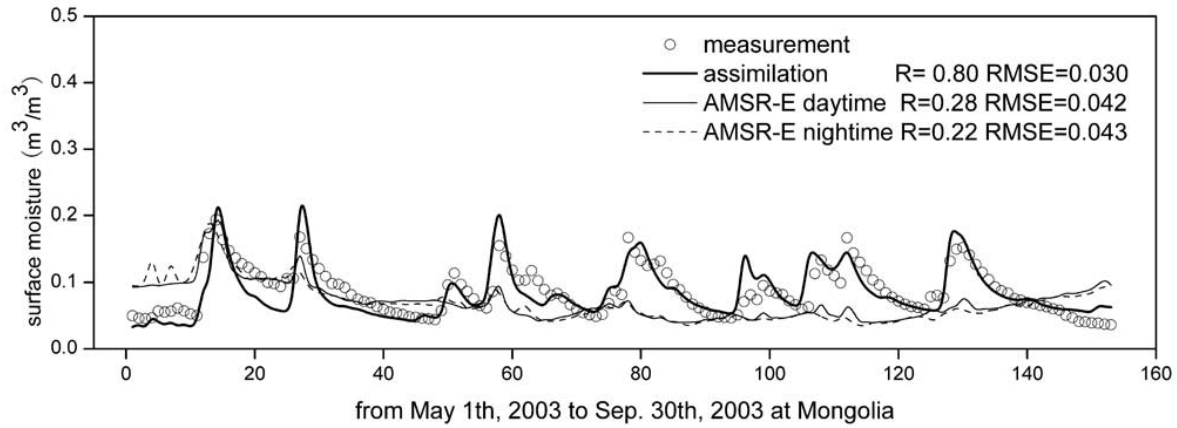


Figure 4. Comparison of soil moisture between the assimilation and NASA AMSR-E retrievals at ascending pass and descending pass.

are shown in Figure 4 and they have almost the same values. Surface soil moisture from the assimilation algorithm clearly outperforms NASA AMSR-E moisture products according to two error metrics, RMSE (Root Mean Square Error) and R (Correlation Coefficient). It is also found that NASA AMSR-E soil moisture products cannot reflect the absolute soil moisture values over the temperate and semiarid regions [Gruhier *et al.*, 2008]. Some studies [Draper *et al.*, 2009] indicate that the poor performance of NASA AMSR-E soil moisture is due to its retrieval algorithm. In addition, inversion algorithms retrieves soil surface moisture merely according to the instantaneous information recorded by remote sensors and other ancillary ground data, which are hard to be acquired accurately at a regional or global scale. Here the merit of the data assimilation algorithm presented in this study is obviously exhibited, though the hydrological model and the water index (SWI) are simple.

[34] The dynamic model used in the assimilation algorithm is constructed at the daily time step and the thermal

process is not taken into account. This differs from the assimilation work by Yang *et al.* [2007]. In their work, the model SiB2 is applied as dynamics, which characterizes both hydrological and thermal processes of the land surface and integrates forward with hourly meteorological forcing data. Consequently, microwave brightness temperature is directly assimilated through a dual-pass method that first estimates optimal model parameters and then estimates soil moisture. Figure 5 shows comparisons of soil moisture between Yang's and our algorithms at the same site. The algorithm presented in this work slightly overestimates the water content relative to Yang's method during conditions of intensive precipitation, but the two estimates are quite comparable according to the error metrics or visually.

6.2. Retrievals of Model Parameters

[35] The retrieval of soil texture properties from soil surface information, especially the remote sensing signal, attract much attention from researchers in different fields since they play an important role in the determination of soil

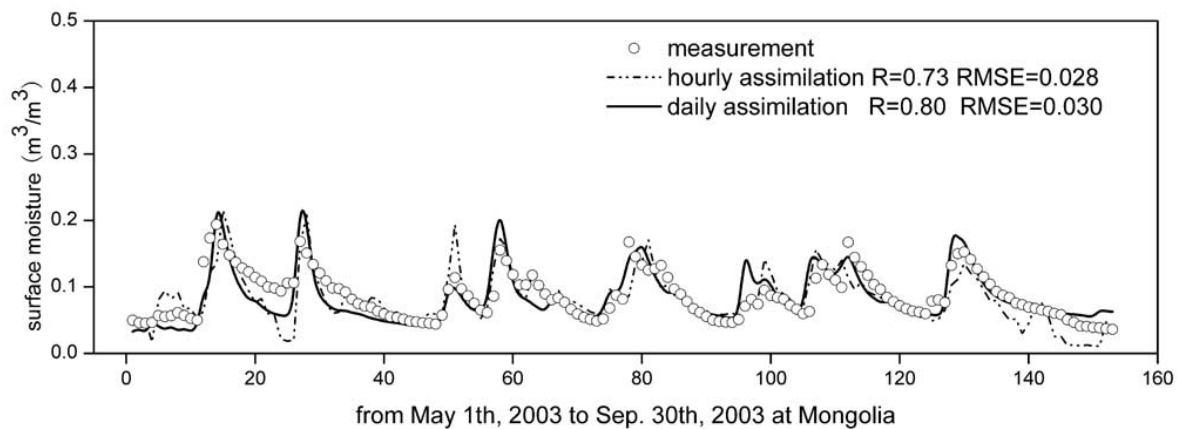


Figure 5. Comparison of daily mean soil moisture between the algorithm presented in this study and the dual-pass assimilation by Yang *et al.* [2007]. Note that this study uses a daily time step for integration of the LSM while Yang *et al.* uses hourly time step.

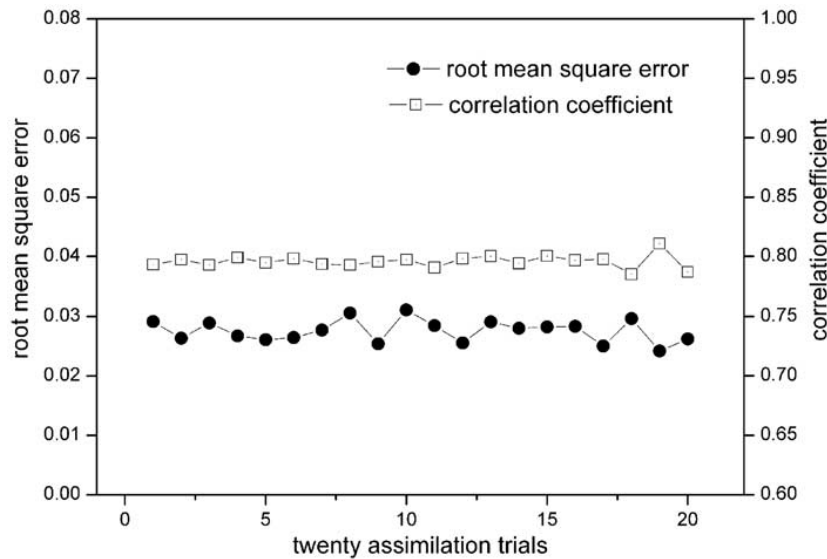


Figure 6. Retrieval results of surface soil moisture according to RMSE and R in 20 assimilation trials with different initial values.

hydraulic and thermal properties, which greatly affect the soil water and heat movement. However, barely effective values could be retrieved for each pixel with remote sensing information as constraints, because the variability of soil properties in both horizontal and vertical directions is rather large and there is not enough information to retrieve these heterogeneous properties.

[36] In this work, model parameters are also estimated in addition to soil moisture, which are [%sand, %clay, θ_{sat} , β , χ]. In the following, the main focus is on three soil parameters %sand, %clay, and θ_{sat} since they are physical properties of soil, from which soil parameters that control water and heat movement are estimated. These three parameters should keep stable or change slightly around some values after going through an intense adjustment at the initial stage of the whole assimilation. Moreover, twenty trials are performed in order to investigate the sensitivity of both the retrieved soil moisture and the retrieved values of %sand, %clay, and θ_{sat} to initial values of state variables and all parameters. As shown in Figure 6, initial settings almost have no influence on retrieved surface soil moisture in accordance with two error metrics. However, the values of retrieved parameters %sand, %clay, and θ_{sat} do not converge to certain fixed points although they are stable after the initial stage, as indicated in Figure 7. The retrieved %sand and %clay have a large variability, but retrieved soil saturation θ_{sat} shows a small variability and stays within a reasonable range according to measured soil saturation at three different sites, MGS, DRS, and BTS. There are no soil texture data available. Thus some quantitative comparisons cannot be given to %sand and %clay retrievals. Many investigators have found that it is very difficult to retrieve soil properties by using soil surface moisture information. This could be explained by the concept of equifinality [Beven and Freer, 2001] which essentially recognizes that different initial states can lead to similar end states. According to Beven's point of

view, only a certain *a posteriori* joint probability distribution could be obtained for parameters to be estimated. The more information, which is closely related to variables and parameters, is included in observations, the more uncertainties are removed in retrieved results. Different combinations of values of parameters could lead to similar observations. Each line in Figure 7 can be regarded as a realization of the posterior distribution.

7. Conclusions

[37] Accurate estimation of soil moisture is very important since it is a key parameter in the terrestrial water cycle. Traditional methods for collecting soil moisture information cannot meet the requirements in many applications. Remote sensing has become a feasible approach to map surface soil moisture on a global scale. However, it cannot obtain soil moisture status in the root zone since the remote sensing signal generally reflects superficial land surface information. The data assimilation method can couple land surface models and remote sensing observations so that it opens up prospects for accurate estimation of the soil moisture. Research on this topic has been widely performed in a range of fields such as hydrology, agriculture, and meteorology.

[38] In this study, a relatively simple data assimilation system is developed. This DA scheme takes the daily-based model as dynamic constraints, a new water index derived from the microwave radiative transfer as the observation operator, and a particle filter as merging scheme, and then assimilates microwave AMSR-E data to jointly estimate soil moisture and model parameters. Retrieval results from the assimilation algorithm presented in this study are compared with field measurements, retrieved soil moisture from the standard AMSR-E inversion algorithm, and a dual-pass assimilation scheme with an hourly-based LSM. Comparisons indicate that this assimilation algorithm can estimate

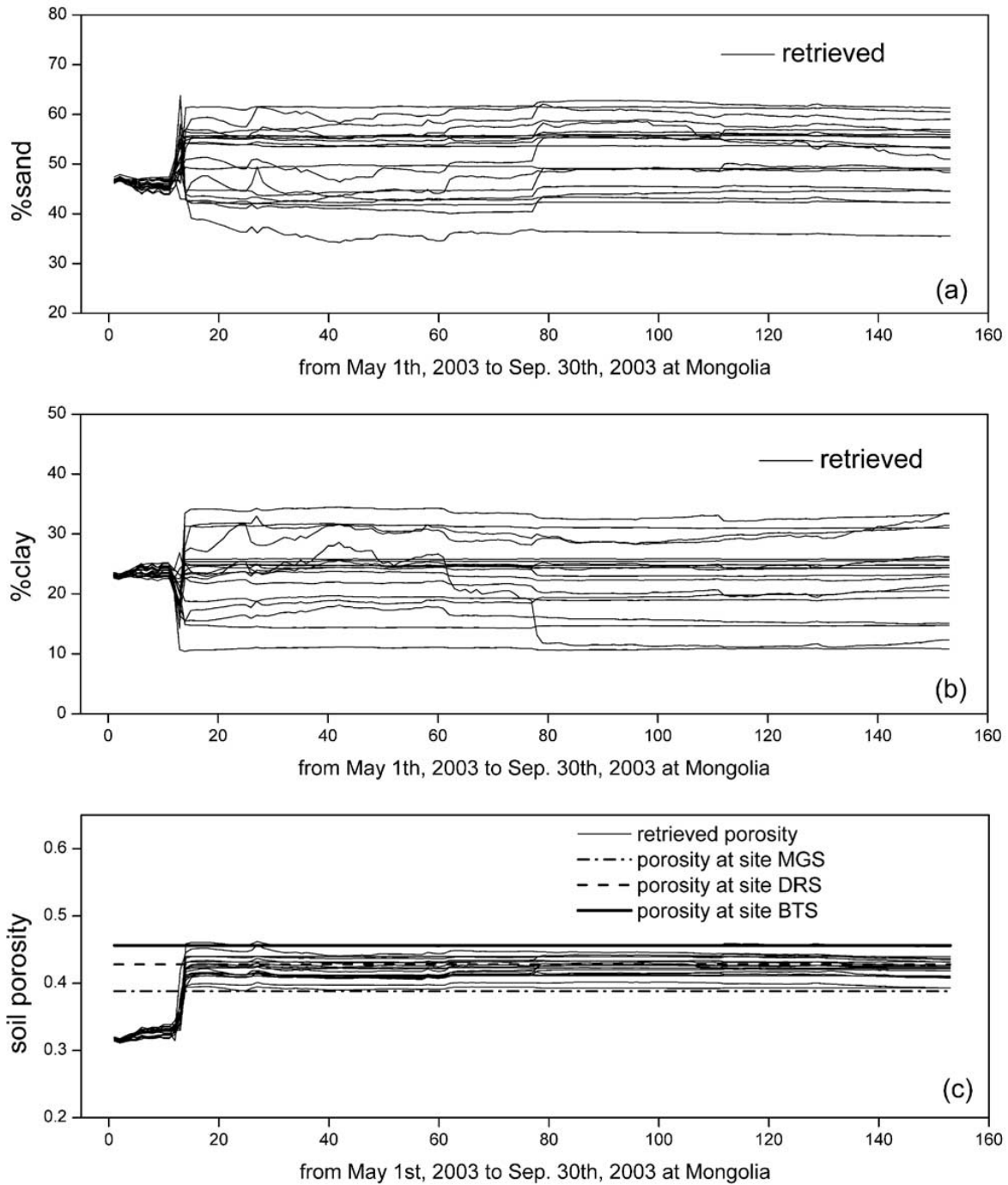


Figure 7. Variation of retrieved soil texture and soil porosity with time, given different initial values in 20 trials.

soil surface moisture with satisfactory speed and precision at a daily time resolution. Particularly, the temporal variability of AMR-E product is too small compared to the observed one in our studied area. At the same time, soil hydraulic properties are also estimated. Results show that the retrieved soil texture cannot converge to a fixed value or

a narrow range in our experiments while the retrieved soil porosity is confined to a relatively narrow range. However, retrieved soil surface moisture agrees well with station-averaged moisture measurements. Soil moisture in the root zone or deep zone could not be retrieved with a high accuracy since there is no sufficient information to derive

relevant soil parameters, which possibly change with the soil depth. Some researchers also find this problem that the water content in the root zone cannot be easily retrieved, just relying on the soil surface moisture information without accurate soil hydraulic properties.

[39] **Acknowledgments.** This work was supported by the Chinese Natural Science Foundation under contract 40701129, the “100-Talent” Project of the Chinese Academy of Sciences, Chinese 863 project under contract 2006AA01A120, and Data Sharing Network of Earth System Science under contract 2005DKA32300. We thank Dr. Carol Russell for English editing. AMPEX was implemented in the framework of NASDA-JRA “Ground Truth for Evaluation of Soil Moisture and Geophysical/Vegetation parameters Related to Ground Surface Conditions with AMSR and GLI in the Mongolian Plateau” (PI: Prof. Ichiro Kaihitsu, University of Hiroshima). The Mongolian partnership is among the Institute of Meteorology and Hydrology, the National Agency for Meteorology, Hydrology, and Environment Monitoring of Mongolia.

References

- Bertino, L., et al. (2003), Sequential data assimilation techniques in oceanography, *Int. Stat. Rev.*, **71**, 223–241.
- Beven, K., and J. Freer (2001), Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, *J. Hydrol.*, **249**, 11–29.
- Castellvi, F., et al. (2001), Comparison of methods for applying the Priestley-Taylor equation at a regional scale, *Hydrol. Processes*, **15**, 1609–1620.
- Chen, T., et al. (2005), Particle filters for state and parameter estimation in batch processes, *J. Process Control*, **15**, 665–673.
- Claussen, M. (1998), On multiple solutions of the atmosphere-vegetation system in present-day climate, *Global Change Biol.*, **4**, 549–559.
- Crow, W. T. (2003), Correcting land surface model predictions for the impact of temporally sparse rainfall rate measurements using an ensemble Kalman filter and surface brightness temperature observations, *J. Hydrometeorol.*, **4**, 960–973.
- Crow, W. T., and E. F. Wood (2003), The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using Ensemble Kalman filtering: A case study based on ESTAR measurements during SGP97, *Adv. Water Resour.*, **26**, 137–149.
- Daly, E., and A. Porporato (2005), A review of soil moisture dynamics: From rainfall infiltration to ecosystem response, *Environ. Eng. Sci.*, **22**, 9–24.
- Davies, J. A., and C. D. Allen (1973), Equilibrium, potential and actual evaporation from cropped surfaces in southern Ontario, *J. Appl. Meteorol.*, **12**, 649–657.
- Draper, C. S., et al. (2009), An evaluation of AMSR-E derived soil moisture over Australia, *Remote Sens. Environ.*, **113**, 703–710.
- Drusch, M. (2007), Initializing numerical weather prediction models with satellite derived surface soil moisture: Data assimilation experiments with ECMWF's Integrated Forecast System and the TMI soil moisture data set, *J. Geophys. Res.*, **112**, D03102, doi:10.1029/2006JD007478.
- Drusch, M., and P. Viterbo (2007), Assimilation of screen-level variables in ECMWF's integrated forecast system: A study on the impact on the forecast quality and analyzed soil moisture, *Mon. Weather Rev.*, **135**, 300–314.
- Entekhabi, D., et al. (1994), Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely sensed observations, *Geosci. Remote Sens., IEEE Trans.*, **32**, 438–448.
- Evensen, G. (2003), The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean Dyn.*, **53**, 343–367.
- Foley, J. A. (1994), The sensitivity of the terrestrial biosphere to climatic change: A simulation of the middle Holocene, *Global Biogeochem. Cycles*, **8**, 505–525.
- Fujii, H. (2005), Development of a microwave radiative transfer model for vegetated land surface based on comprehensive in-situ observations, Ph.D. thesis, Univ. of Tokyo, Japan.
- Galanowicz, J. F., et al. (1999), Tests of sequential data assimilation for retrieving profile soil moisture and temperature from observed L-band radiobrightness, *Geosci. Remote Sens., IEEE Trans.*, **37**, 1860–1870.
- Gruhier, C., et al. (2008), Evaluation of AMSR-E soil moisture product based on ground measurements over temperate and semi-arid regions, *Geophys. Res. Lett.*, **35**, L10405, doi:10.1029/2008GL033330.
- Hamill, T. M. (2006), Ensemble-based atmospheric data assimilation, in *Predictability of Weather and Climate*, edited by T. Palmer and R. Hagedorn, pp. 124–156, Cambridge Univ. Press, New York.
- Han, X., and X. Li (2008), An evaluation of the nonlinear/non-Gaussian filters for the sequential data assimilation, *Remote Sens. Environ.*, **112**, 1434–1449.
- Helton, J. C., et al. (2006), Survey of sampling-based methods for uncertainty and sensitivity analysis, *Reliab. Eng. Syst. Saf.*, **91**, 1175–1209.
- Hirabayashi, Y., et al. (2005), A 100-year (1901–2000) global retrospective estimation of the terrestrial water cycle, *J. Geophys. Res.*, **110**, D19101, doi:10.1029/2004JD005492.
- Houser, P. R., et al. (1998), Integration of soil moisture remote sensing and hydrologic modeling using data assimilation, *Water Resour. Res.*, **34**, 3405–3420.
- Huang, C., et al. (2008a), Retrieving soil temperature profile by assimilating MODIS LST products with ensemble Kalman filter, *Remote Sens. Environ.*, **112**, 1320–1336.
- Huang, C., et al. (2008b), Experiments of one-dimensional soil moisture assimilation system based on ensemble Kalman filter, *Remote Sens. Environ.*, **112**, 888–900.
- Kaihitsu, I., T. Yamanaka, T. Koike, D. Oyumbaatar, and G. Davaa (Eds.) (2005), Ground truth for evaluation of soil moisture and geophysical/vegetation parameters related to ground surface conditions with AMSR and GLI in the Mongolian Plateau, in *Ground-based Observations for the ADEOS II/AQUA Validation in the Mongolian Plateau*, pp. 5–21, Japan Aerosp. Explor. Agency.
- Kerr, Y. H., et al. (2001), Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission, *Geosci. Remote Sens., IEEE Trans.*, **39**, 1729–1735.
- Kustas, W. P., et al. (1996), Variability in surface energy flux partitioning during Washita'92: Resulting effects on Penman-Monteith and Priestley-Taylor parameters, *Agric. For. Meteorol.*, **82**, 171–193.
- Lermusiaux, P. F. J., and A. R. Robinson (2001), Data assimilation in models, in *Encyclopedia of Ocean Sciences*, edited by J. Steel et al., pp. 623–634, Academic, London.
- Liang, S., and J. Qin (2008), Data assimilation methods for land surface variable estimation, in *Advances in Land Remote Sensing*, edited by S. Liang, pp. 313–339, Springer, Netherlands.
- Margulis, S. A., D. McLaughlin, D. Entekhabi, and S. Dunne (2002), Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment, *Water Resour. Res.*, **38**(12), 1299, doi:10.1029/2001WR001114.
- McLaughlin, D. (2002), An integrated approach to hydrologic data assimilation: Interpolation, smoothing, and filtering, *Adv. Water Resour.*, **25**, 1275–1286.
- Moradkhani, H. (2008), Hydrologic remote sensing and land surface data assimilation, *Sensors*, **8**, 2986–3004.
- Moradkhani, H., et al. (2005a), Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, *Water Resour. Res.*, **41**, W05012, doi:10.1029/2004WR003604.
- Moradkhani, H., et al. (2005b), Dual state-parameter estimation of hydrological models using ensemble Kalman filter, *Adv. Water Resour.*, **28**, 135–147.
- Nakayama, K., et al. (1993), Estimation of soil moisture in the shallow root zone region, *J. Agric. Meteorol.*, **48**, 851–854.
- Ni-Meister, W., et al. (2006), Soil moisture initialization for climate prediction: Assimilation of scanning multifrequency microwave radiometer soil moisture data into a land surface model, *J. Geophys. Res.*, **111**, D20102, doi:10.1029/2006JD007190.
- Njoku, E. G., et al. (2003), Soil moisture retrieval from AMSR-E, *Geosci. Remote Sens., IEEE Trans.*, **41**, 215–229.
- Reichle, R. H., et al. (2001), Downscaling of radio brightness measurements for soil moisture estimation: A four-dimensional variational data assimilation approach, *Water Resour. Res.*, **37**, 2353–2364.
- Reichle, R. H., et al. (2002), Hydrologic data assimilation with the ensemble Kalman filter, *Mon. Weather Rev.*, **130**, 103–114.
- Reichle, R. H., et al. (2007), Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR), *J. Geophys. Res.*, **112**, D09108, doi:10.1029/2006JD008033.
- Rodell, M., et al. (2004), The global land data assimilation system, *Bull. Am. Meteorol. Soc.*, **85**, 381–394.
- Sau, F., et al. (2004), Testing and improving evapotranspiration and soil water balance of the DSSAT crop models, *Agron. J.*, **96**, 1243–1257.
- Schmugge, T. J., et al. (2002), Remote sensing in hydrology, *Adv. Water Resour.*, **25**, 1367–1385.
- Sellers, P. J., et al. (1996), A revised land surface parameterization (SiB2) for atmospheric GCMs: Part I. Model formulation, *J. Clim.*, **9**, 676–705.
- Sheffield, J., and E. F. Wood (2007), Characteristics of global and regional drought, 1950–2000: Analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle, *J. Geophys. Res.*, **112**, D17115, doi:10.1029/2006JD008288.

- Texier, D., et al. (1997), Quantifying the role of biosphere-atmosphere feedbacks in climate change: Coupled model simulations for 6000 years BP and comparison with palaeodata for northern Eurasia and northern Africa, *Clim. Dyn.*, *13*, 865–881.
- Thomas, B. S. (2006), Estimation of nonlinear dynamic systems - theory and applications, Ph.D. thesis, Linköpings Univ., Linköpings, Sweden.
- Wagner, W., K. Scipal, C. Pathe, D. Gerten, W. Lucht, and B. Rudolf (2003), Evaluation of the agreement between the first global remotely sensed soil moisture data with model and precipitation data, *J. Geophys. Res.*, *108*(D19), 4611, doi:10.1029/2003JD003663.
- Walker, J. P., and P. R. Houser (2001), A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations, *J. Geophys. Res.*, *106*, 11,761–11,774.
- Yang, K., et al. (2007), Auto-calibration system developed to assimilate AMSR-E data into a land surface model for estimating soil moisture and the surface energy budget, *J. Meteorol. Soc. Jpn.*, *85*, 229–242.
- Zhou, Y., et al. (2006), Assessing the performance of the ensemble Kalman filter for land surface data assimilation, *Mon. Weather Rev.*, *134*, 2128–2142.
- I. Kaihotsu, Department of Natural and Environmental Sciences, Faculty of Integrated Arts and Sciences, Hiroshima University, Kagamiyama 1-7-1, Higashi-Hiroshima 739-8521, Japan.
- T. Koike, Department of Civil Engineering, School of Engineering, University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan.
- S. Liang, Department of Geography, University of Maryland, 2181 LeFrak Hall, College Park, MD 20742, USA.
- R. Liu, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, No. 11A, Datun Road, Chaoyang District, Beijing 100101, China.
- J. Qin and K. Yang, Key Laboratory of Tibetan Environment Changes and Land Surface Processes, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, P.O. Box 2871, Beijing 100085, China. (shuairnqin@gmail.com)

Other Key Reference Papers for Prof. Qin's Lecture:

- [1] Li, X., et al., (2007), Auto-calibration system developed to assimilate AMSR-E data into a land surface model for estimating soil moisture and the surface energy budget, *Journal of the Meteorological Society of Japan*, *85A*: 229-242
- [2] liang, S., J. Qin (2008), Chapter 12: Data assimilation methods for land surface variable estimation, *Advances in Land Remote Sensing: System, Modeling, Inversion and Application*, Springer

Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework

Yuqiong Liu¹ and Hoshin V. Gupta¹

Received 20 November 2006; revised 20 April 2007; accepted 30 April 2007; published 3 July 2007.

[1] Despite significant recent developments in computational power and distributed hydrologic modeling, the issue of how to adequately address the uncertainty associated with hydrological predictions remains a critical and challenging one. This issue needs to be properly addressed for hydrological modeling to realize its maximum practical potential in environmental decision-making processes. Arguably, the key to properly addressing hydrologic uncertainty is to understand, quantify, and reduce uncertainty involved in hydrologic modeling in a cohesive, systematic manner. Although general principles and techniques on addressing hydrologic uncertainty are emerging in the literature, there exist no well-accepted guidelines about how to actually implement these principles and techniques in various hydrologic settings in an integrated manner. This paper reviews, in relevant detail, the common data assimilation methods that have been used in hydrologic modeling to address problems of state estimation, parameter estimation, and system identification. In particular, the paper discusses concepts, methods, and issues involved in hydrologic data assimilation from a systems perspective. An integrated hierarchical framework is proposed for pursuing hydrologic data assimilation in several progressive steps to maximally reduce uncertainty in hydrologic predictions.

Citation: Liu, Y., and H. V. Gupta (2007), Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, *Water Resour. Res.*, 43, W07401, doi:10.1029/2006WR005756.

1. Introduction

[2] Hydrologic modeling has benefited from significant developments over the past two decades, including dramatic growths in computational power, ever increasing availability of distributed hydrologic observations, and improved understanding of the physics and dynamics of the hydrologic system. This has led to the building of higher levels of complexity into hydrologic models, and an advance from lumped, conceptual models toward semidistributed and distributed physics-based models. Paradoxically, while these advances reflect our growing understanding, they have also increased the need for concrete methods to deal with the increasing uncertainty associated with the models themselves, and with the observations required for driving and evaluating the models. It is now being broadly recognized that proper consideration of uncertainty in hydrologic predictions is essential for purposes of both research and operational modeling [Wagner and Gupta, 2005]. The value of a hydrologic prediction to water resources and other relevant decision-making processes is limited if reasonable estimates of the corresponding predictive uncertainty are not provided [e.g., Georgakakos et al., 2004].

[3] To adequately address uncertainty in hydrologic modeling, there are three distinct yet related aspects to be considered: understanding, quantification, and reduction of uncertainty. Arguably, understanding uncertainty is an inte-

gral part of any application of uncertainty quantification and/or reduction. Many uncertainty analysis frameworks have been introduced in the hydrologic literature, including the generalized likelihood uncertainty estimation (GLUE) methodology [Beven and Binley, 1992], the Bayesian recursive estimation technique (BaRE) [Thiemann et al., 2001], the Shuffled Complex Evolution Metropolis algorithm (SCEM) [Vrugt et al., 2003a], the multiobjective extension of SCEM [Vrugt et al., 2003b], the dynamic identifiability analysis framework (DYNIA) [Wagner et al., 2003], the maximum likelihood Bayesian averaging method (MLBMA) [Neuman, 2003], the dual state-parameter estimation methods [Moradkhani et al., 2005a, 2005b], and the simultaneous optimization and data assimilation algorithm (SODA) [Vrugt et al., 2005]. However, few of these methods completely address all the above three critical aspects of uncertainty analysis in an explicit and cohesive way.

[4] Methods of probabilistic prediction and data assimilation (DA) for quantification and reduction of state uncertainty have been extensively explored in the atmospheric and oceanic sciences [e.g., Daley, 1991; Courtier et al., 1993; Anderson and Anderson, 1999]. Their application in the hydrological sciences is relatively new, although deterministic hydrological prediction and parameter estimation have become reasonably mature. Nevertheless, the hydrologic literature has seen various applications of data assimilation and/or uncertainty analysis in hydrology ranging from characterization of soil moisture and/or surface energy balance [e.g., Entekhabi et al., 1994; Houser et al., 1998; Entekhabi et al., 1999; Galantowicz et al., 1999; Boni et al., 2001; Walker et al., 2001; Reichle et al., 2001a, 2001b, 2002a, 2002b; Margulis et al., 2002; Dunne and Entekhabi,

¹Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona, USA.

2005], to rainfall-runoff modeling [e.g., *Restrepo*, 1985; *Moradkhani et al.*, 2005a, 2005b; *Vrugt et al.*, 2005], to flood forecasting [e.g., *Kitanidis and Bras*, 1980; *Young*, 2002], to estimation of hydraulic conductivity [e.g., *Katul et al.*, 1993; *Lee et al.*, 1993], to groundwater flow and transport problems [e.g., *Eigbe et al.*, 1998; *Graham and McLaughlin*, 1991; *McLaughlin et al.*, 1993], to estimation of water table elevations [e.g., *Van Geer et al.*, 1991; *Yangxiao et al.*, 1991], and to water quality modeling [e.g., *Beck*, 1987].

[5] One critical issue for hydrologic modeling is how the DA methods used in atmospheric and related sciences can best be adapted and combined with hydrologic methods to cope with the uncertainties arising from hydrologic modeling in a cohesive, systematic way to maximally reduce and adequately quantify the predictive hydrologic uncertainty [Krzysztofowicz, 1999; Mantovan and Todini, 2006]. Although general principles and techniques on addressing hydrologic uncertainty are emerging in the literature, there exist no well-accepted guidelines about how to actually implement these principles and techniques in various hydrologic settings. In this paper we discuss the sources of uncertainty in hydrological modeling from a systems perspective, illustrate in detail some of the common DA methods that have been used to quantify and reduce hydrological uncertainty, and propose a (preliminary) hierarchical data assimilation framework for systematically addressing the various types of uncertainties as a way to move forward. It is worth noting that this paper does not attempt to provide a comprehensive review of the literature regarding all the methods, applications, and issues related to data assimilation in hydrology; instead, we aim to present to the readers an illustrative and integrated (rather than fragmented) picture of the state of the art of hydrological data assimilation from a systems perspective.

[6] The paper is organized as follows: Section 2 discusses the three important aspects in addressing hydrologic uncertainty, i.e., understanding, quantifying, and reducing uncertainty; in section 3 we present an integrated view of uncertainty in hydrologic modeling from a systems perspective; Bayes' theorem and its application to data assimilation are discussed in section 4; sections 5, 6, 7, and 8 are devoted to reviews of the common methods that have been used to approach problems of system identification, parameter estimation, state estimation, and simultaneous state and parameter estimation, respectively; an integrated Bayesian hierarchical framework for handling all hydrologic uncertainty in a cohesive, systematic manner is proposed in section 9; and the paper closes with some general discussions and recommendations for future research in section 10.

2. Understanding, Quantifying, and Reducing Hydrologic Uncertainty

[7] As mentioned in the introduction, understanding, quantification, and reduction of uncertainty are the three critical aspects to be considered in order to adequately address uncertainty in hydrologic modeling and prediction. For a full uncertainty analyses one may argue that there exists an additional aspect where uncertainty in the predictions are analyzed and interpreted to infer the deficiencies in the model and data, a process that *Wagener and Gupta* [2005] referred to as "uncertainty communication." This,

however, is beyond the scope of the current paper, which focuses on hydrologic data assimilation.

[8] Obviously, without first adequately understanding all the different uncertainty sources and the relationships between them, it is difficult to conduct uncertainty quantification and reduction in a meaningful way. This is because different uncertainty sources may introduce significantly different error characteristics that require different techniques to deal with; and missing important uncertainty sources may lead to misleading uncertainty predictions in the hydrologic outputs. As of today, our understanding of hydrologic uncertainty is still far from complete and there is much room for further efforts in search of cohesive, systematic means to approach this. It is also very important to distinguish modeling uncertainty from predictive uncertainty: While modeling uncertainty comes mainly from the imperfect fit to the truth of the past, predictive uncertainty can also arise from extrapolation errors or temporal prediction errors due to the fact that the future typically does not look exactly like the past [e.g., *Morgan et al.*, 1990; *Krupnick et al.*, 2006]. In other words, predictive uncertainty is related to, but not necessarily equivalent to, modeling uncertainty; and reduction in modeling uncertainty does not necessarily lead to enhanced predictability of the model under changing conditions. In decision-making processes, there may exist other types of uncertainty, such as decision uncertainty, which arises "whenever there is ambiguity and controversy about how to quantify or compare social objective" [Finkel, 1990, p. 16], and scenario uncertainty, which is related to the inability of the scenarios to account for all the factors affecting the key output/decision variables [Cullen and Frey, 1999]. In the context of hydrological data assimilation, addressing modeling uncertainty is of primary interest, which, in turn, will have an impact on predictive uncertainty.

[9] As far as quantifying uncertainty is concerned, a classical and straightforward way presented in the literature is to represent the predictions in terms of a probability distribution, computed by performing probabilistic instead of deterministic prediction/modeling [e.g., *Kuczera and Parent*, 1998; *Krzysztofowicz*, 1999; *Montanari and Brath*, 2004; *Tamea et al.*, 2005]. For example, by producing an ensemble of hydrologic predictions (instead of a single deterministic prediction as does traditional hydrologic modeling), probabilistic prediction seeks to take into account uncertainties in the equations and/or parameters that are used to describe the physical system and in the hydrologic observations that are made on the system and used in the prediction/modeling process. Of course, for effective quantification of the uncertainty, some prior knowledge (estimate) about the error characteristics that describe the probability distribution of the uncertainties is required, indicating that quantification of uncertainty is, indeed, to a large degree dependent on the understanding of uncertainty. In practical applications of probabilistic prediction, the high nonlinearity of the hydrologic system and the complex interactions between different components of the system result in it being highly difficult to estimate and apply probability distributions that accurately represent the true joint distributions of the uncertainties without creating computational and/or mathematical difficulties. Hence, in practice, locally linear assumptions are usually made about

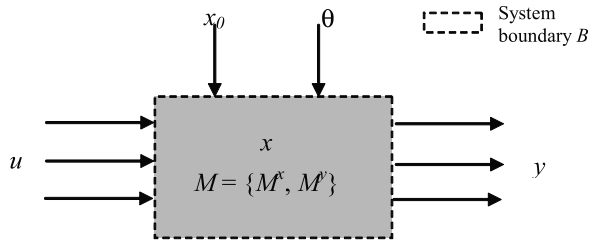


Figure 1. Schematic diagram of model components from a systems perspective.

the system, and uniform or (truncated) Gaussian/normal distributions are typically used to quantitatively represent various sources of uncertainties [e.g., *Moradkhani et al.*, 2005a, 2005b]. To quantify the uncertainty in hydrologic outputs, sampling (sometimes called ensemble) methods are now widely used by taking samples from the assumed error probability density functions (PDFs) and running the model forward for a certain amount of time. With a sufficiently large sample of predictions, statistics describing the uncertainties in model outputs can be easily derived from the sample. In most cases, quantification of uncertainty is embedded in the data assimilation processes aiming to reduce predictive uncertainty as discussed below.

[10] There are three main areas where actions can be taken toward reducing uncertainty in hydrologic predictions: (1) acquisition of more informative and higher quality hydrological data (including data of new types) by developing improved measurement techniques and observation networks; (2) development of improved hydrologic models by incorporating better representations of physical processes and using better mathematical techniques; and (3) development of efficient and effective techniques that can better extract and assimilate information from the available data via the model identification and prediction processes.

[11] While hydrologic science has witnessed astonishing advances in the availability of hydrologic data (area 1) and the complexity/reliability of hydrological models (area 2), there is an urgent need for techniques that effectively and efficiently assimilate important information from the data into the models to produce improved hydrological predictions (area 3). We will generally refer to such techniques as data assimilation (DA) methods, defined here as

procedures that aim to produce physically consistent representations or estimates of the dynamical behavior of a system by merging the information present in imperfect models and uncertain data in an optimal way to achieve uncertainty quantification and reduction.

[12] It is worth mentioning that this description of the DA problem is broadly encompassing, not being limited only to problems of “state estimation” as the term is often applied to in the literature. Instead, it describes the more comprehensive problem of “merging models with data” and therefore includes the three related problems of system (structure) identification, parameter estimation, and state estimation, which are all critical to the reduction of uncertainty in model predictions. More details on these concepts are provided in section 3.

[13] Arguably, understanding uncertainty should always be an integral part of any application of uncertainty

quantification/reduction; and given the continual arrival of different kinds of observations, one should not stop at the quantification step but continue to reduce the uncertainty by assimilating new observations. In most cases of DA applications, the process of uncertainty reduction inherently involves the quantification of uncertainty in the model inputs, parameters, structure, and observations, and preferably provides quantitative information about uncertainty in model predictions or forecasts. In recognition of this, the focus of this paper is given to understanding and reduction of hydrologic uncertainty from a systems perspective (sections 3–9).

3. Hydrologic Uncertainty From a Systems Perspective

[14] Uncertainty in hydrologic modeling may arise from several sources: model structure, parameters, initial conditions, and observational data used to drive and evaluate the model. In this section, to formally specify the different sources, we will describe a model as being composed of multiple components from the perspective of systems theory. Errors in each of these model components can give rise to uncertainty in hydrologic modeling. In this sense we include within the realm of data assimilation any procedure that assimilates information from observations to reduce the uncertainty associated with one or more of the model components, be it the state, the parameters, or the system structure.

3.1. Model Components in Systems Theory

[15] For the purpose of communication, here we consider a model to be composed of seven different components (Figure 1): system boundary (B), inputs (u), initial states (x_0), parameters (θ), structure (M), states (x), and outputs (y). Note not all the hydrologic applications existing in the literature comply with this definition/terminology of system components (see below).

[16] In this exposition we define the inputs u and outputs y as fluxes of mass and/or energy into and out of the system across the system boundary B ; states x as time-varying quantities of mass and/or energy stored within the system boundary B ; and parameters θ as characteristic properties of the system that are assumed to be “time-invariant” (remain constant over the time duration of interest). Note that in some fields, the system “state” x^* is taken to be some other quantity somehow related to the mass or energy state x ; in such cases the same general equations hold but with some modifications to account for the relationship of x^* to x . Also, we shall return to the issue of time-invariance of model parameters in a moment. For example, in catchment modeling, u may refer to the time-varying two-dimensional spatial distribution of precipitation flux over the catchment; y may refer to the time-varying two-dimensional distribution of streamflow flux at all points along the river network and of evaporation and transpiration from the surface of the catchment; x may refer to the three-dimensional time-varying spatial distribution of surface and subsurface moisture stored within the catchment boundary; and θ may refer to the time-invariant three-dimensional spatial distribution of catchment characteristics such as the soil hydraulic properties.

[17] The model structure M consists of two components: M^x and M^y (i.e., $M = \{M^x, M^y\}$). Here M^x and M^y are (in general) nonlinear vector functional relationships, where M^x represents the input-to-state mapping and M^y represents the state-to-output mapping. For example, M^x may refer to the coupled equations describing the three-dimensional evolution of surface and subsurface moisture in response to catchment inputs and outputs (precipitation, evaporation, transpiration, and outflow), and M^y may refer to the coupled equations describing the dependence of catchment outputs (evaporation, transpiration, and outflow) on the system states. These mappings can be described or constructed in a variety of different ways, including the continuous-time differential equation formulation (using t to represent continuously varying time):

$$dx/dt = M^x(x, u|\theta) \quad (1)$$

$$y = M^y(x|\theta), \quad (2)$$

and the discrete-time difference equation formulation (using k to represent discrete moments in continuously varying time t):

$$x_{k+1} = M_{k+1}^x(x_k, u_{k+1}|\theta) \quad (3)$$

$$y_{k+1} = M_{k+1}^y(x_{k+1}|\theta). \quad (4)$$

[18] Since computer-based implementations are usually constructed to make predictions at discrete moments of time, we shall (without loss of generality) use the discrete time formulation described by equations (3) and (4) in all subsequent discussion. Note that the formulation must implicitly employ the continuity equation $dx/dt = u - y$ to ensure physical consistency in the time-dependent accounting for mass and energy fluxes.

[19] As mentioned above, this formulation assumes that the model parameters θ do not vary with time over the duration of interest. As a conceptual extension, one might wish more generally to permit the system characteristics represented as “parameters” to vary slowly with time, in response to changes in the model state and/or system inputs. In general, we would expect (for reasons of physical consistency) that the rate of this “parameter” variation is slower than that of the variation of the state. To complete the mathematical description, we would then introduce an additional set of mapping relationships that describes, in a manner analogous to the input-state relationship, the time-evolution of the parameters θ (see equations (5)–(7)).

$$\theta_{k+1} = M_{k+1}^\theta(x_k, u_{k+1}, \theta_k|\phi) \quad (5)$$

$$x_{k+1} = M_{k+1}^x(x_k, u_{k+1}|\theta_k) \quad (6)$$

$$y_{k+1} = M_{k+1}^y(x_{k+1}|\theta_k). \quad (7)$$

[20] Note that this revised formulation introduces a new set of (uncertain) time-invariant coefficients ϕ which must be specified a priori or estimated from data; for example, if

θ is believed to take the same Gaussian distribution at all time steps, ϕ might represent the (time-invariant) mean and covariance of that distribution ($\phi \sim N(\mu_\phi, \sigma_\phi^2)$). However, if we define an extended “state” vector $x' = [x, \theta]$ by adjoining the time-varying quantities x and θ into a single variable, and define a new “parameter” vector ϕ , the formulation in (5)–(7) is not fundamentally different from that given in equations (3)–(4). For simplicity of notation, we will therefore proceed by adopting the representation of equations (3)–(4) and let the reader make the appropriate substitutions for the more general case as necessary.

3.2. Errors in Different Model Components

[21] Of the seven model components illustrated in Figure 1, five of them (i.e., B , u , x_0 , θ , and M) must be specified, estimated, or defined before the model can be actually run, while the remaining two (x and y) are computed by running the model. Each of the five predefined components may be uncertain in various characteristic ways, and the consequence of these uncertainties will be mapped into the model states and outputs. Hence input data, parameters, the model structure, initial conditions, and the system boundary represent five major sources of uncertainties in hydrologic modeling. In most cases, model inputs and initial conditions are specified or estimated from in situ observations. Accordingly, errors in these two sources can be collectively considered as data errors or observation errors. Errors in output observations that are used to evaluate the model results should be considered as data errors as well. Note in cases where x_0 are treated as model parameters, errors associated with x_0 can be considered as parameter errors [e.g., *Liu et al.*, 2003]. Definition of the system boundary is part of the model conceptualization process; hence the uncertainty associated with B can be considered as one source of structural uncertainty. In summary, there are three primary types of uncertainties in hydrologic modeling: structural errors, parameter errors, and data errors (see also discussions by *Wagener and Gupta* [2005]).

[22] 1. Models are assemblies of assumptions and simplifications and thus inevitably imperfect approximations to the complex reality, i.e., the true system that a model seeks to characterize. Conceptualization with inappropriate approximations and omissions can result in large (albeit poorly understood) errors in the conceptual structure of a numerical model. Structure errors can also arise from the mathematical implementation (e.g., spatial and temporal discretizations) that transforms a conceptual model into a numerical model [*Neuman*, 2003].

[23] 2. Model parameters are conceptual aggregate representations of spatially and temporally heterogeneous properties of the real system. Parameters are an integral part of the equation-based modeling approach, and the use of “effective” parameter values in hydrologic modeling is essential. Errors in the estimates of parameter values can result in huge errors in the model outputs as shown in many modeling studies [e.g., *Gupta et al.*, 1998; *Liu et al.*, 2005]. However, the “conceptual” and spatiotemporal aggregate nature of parameters sometimes makes it difficult to specify them directly and unambiguously from observations made in the field (of course exceptions exist, such as pumping tests to estimate system conductivities). In other words, parameters are not often easily measurable, and must

generally be estimated by indirect means (e.g., prior knowledge or model calibration) with consequent introduction of errors and uncertainties.

[24] 3. Data errors can generate uncertainties in hydrologic predictions through the model inputs and initial conditions, both of which can be estimated from observations [e.g., *Clark and Slater*, 2006]. A data error is also referred to as a measurement error if, as typically is the case, the data of concern is measured. A measurement error usually consists of two components: (1) instrument error due to imperfect measurement devices that do not accurately record the variables they are designed to measure and (2) representativeness error due to scale incompatibility or differences (in time or space) between the variable measured by a device and the corresponding model variable. Representativeness error can be discussed in terms of spacing (distance or interval between samples), extent (overall coverage of measurements in space or time), and support (averaging volume or area of samples) [*Blöschl and Grayson*, 2000]. These two error components tend to have very different characteristics which may vary from variable to variable. To effectively quantify or reduce uncertainty in the predictions, statistics of both errors should be considered and adequately specified.

[25] Structural, parameter, and data errors collectively lead to uncertainties in hydrologic predictions of model outputs and states. Among these three types of errors, structural errors are generally the most poorly understood and the most difficult to cope with; nevertheless, their impacts on hydrologic predictions can be far more detrimental than those of parameter errors and data errors [*Carrera and Neuman*, 1986; *Abramowitz et al.*, 2006].

3.3. Addressing Uncertainty in Different Model Components

[26] Viewing model components and the errors in them within a dynamic systems framework (as described in sections 3.1 and 3.2) helps to better understand and organize the different uncertainty sources in hydrologic modeling. The next critical issue is how to adequately represent (or quantify) the uncertainties in these sources and feed them into a DA framework to effectively and efficiently reduce the predictive uncertainty. A DA application usually requires proper specifications or assumptions of the characteristics of errors associated with the four major error sources (initial conditions, inputs, parameters, and model structure). This, however, is not a trivial issue, because prior knowledge on the error characteristics is usually not available, especially for errors associated with poor specification of the model structure. In the meantime, one should realize that different DA problems may require different techniques/algorithms that best fit into the specific problem setting.

[27] Loosely speaking, there are three types of data assimilation problems based on the model component being considered: state estimation, parameter estimation, and system identification, described as follows.

[28] 1. State estimation seeks to characterize the true “state” of the system by optimally combining state information represented by the model with that inferable from all kinds of available data sources, quantitative or qualitative. In the literature the term data assimilation is commonly used to refer specifically to state estimation only [e.g., *McLaughlin*, 1995, 2002]. Note that the definition (including dimension)

and computation of the model state are conditional on the specification of a model structure and values for the parameters. While current data assimilation methods are praised for their ability to deal with all the three types of errors mentioned above (i.e., structural errors, parameter errors, and measurement errors), most applications of state estimation have been focused on the measurement errors only, without rigorous treatment of structural and parameter errors [e.g., *Reichle et al.*, 2002a, 2002b].

[29] 2. Parameter estimation aims to estimate proper values of the model “parameters” based on available data, so that the model makes sufficiently accurate simulations or predictions of the true input-state-output response. Note that the definition, dimension, and specification of the model parameter set are conditional on the specification of a model structure (the form of the input-state-output relationship). Traditionally, parameter estimation has been conducted by using deterministic (manual or automatic) calibration techniques that tend to ignore model structural errors and measurement errors [e.g., *Duan et al.*, 1992; *Yapo et al.*, 1998]. Recently, stochastic data assimilation methods have been developed and applied to parameter estimation problems [e.g., *Thiemann et al.*, 2001; *Moradkhani et al.*, 2005a, 2005b].

[30] 3. System identification involves the selection of appropriate structures (i.e., conceptual models) for a mathematical or numerical model that aims to represent the real system. More specifically, a system identification process aims to define a set of proper mappings (typically equations, e.g., equations (3) and (4)) that accurately represent the relationships between the model inputs, parameters, states, and outputs [e.g., *Neuman*, 2003].

[31] Among the three types of DA problems, system identification is the most important and, typically, also the most difficult, as it may involve the development of qualitative diagnostic measures and include the use of expert knowledge and subjectivity. Regardless of that, to maximally reduce the final total uncertainty in hydrologic predictions, all these three types of problems should be addressed, with order of importance being system identification, parameter estimation, and state estimation, and, when necessary, in an iterative manner. In the meantime, all types of errors (i.e., structural, parameter, and data errors) should be properly considered in each of three types of DA processes to reduce bias and uncertainty in the final predictions.

[32] As mentioned above, we define data assimilation as a process that assimilates information from observational data (quantitative or qualitative) in such a way as to improve estimation/representation of any of the three major model components of concern (i.e., model states, parameters, and structure). The following sections review the methods typically used in hydrological modeling, including methods for system identification, parameter estimation, and state estimation. In general, Bayes’ theorem has been employed as the foundation of various data assimilation methods and is therefore discussed first (section 4).

4. Bayes’ Theorem and Its Application to Data Assimilation

[33] We consider two events A and B , which we expect (from empirical observation or for reasons of physical

consistency) to be related in some manner. We further assume that the probability in the occurrence (or observation) of events A and B can be described by $P(A)$ and $P(B)$. Then, the cooccurrence (or observation) of A and B is represented by the joint probability function $P(A \cap B)$, and this can be further expressed as

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A), \quad (8)$$

where $P(A|B)$ is the conditional probability of occurrence of event A given knowledge that event B has occurred (and similarly for $P(B|A)$). This leads directly to Bayes' theorem:

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}. \quad (9)$$

[34] In the Bayesian use of probabilities, the marginal probabilities ($P(A)$ and $P(B)$) and the conditional probabilities ($P(A|B)$ and $P(B|A)$) are referred to as the prior and posterior PDFs, respectively. Bayes' law provides a powerful basis for a full stochastic representation of all the uncertainties in the model and the data in hydrologic modeling. Using Bayes' theory, equation (9) can be reformulated to describe all three aspects of data assimilation, including system identification, parameter estimation, and state estimation.

[35] As discussed in section 2, the five uncertain quantities (B , u , x_0 , θ , and M) must be specified in order to use equations (3) and (4) to compute estimates of the two remaining quantities (x and y). We represent the prior knowledge of the quantities (B , u , x_0 , θ , and M) by the probabilities $p_{\text{prior}}(B)$, $p_{\text{prior}}(u)$, $p_{\text{prior}}(x_0)$, $p_{\text{prior}}(\theta)$, and $p_{\text{prior}}(M)$, respectively. We further assume that there may become available a set of uncertain observations z which may contain information about any of the system aspects of interest. For example, z may consist of direct or indirect measurements on any of the system fluxes (u and y), state variables (x), parameters (θ), or initial conditions (x_0 and B), or more generally can consist of qualitative assessments of any of these quantities, including the model structure (M). Of course, such observations will generally be incomplete, in the sense that they refer to values at a limited and discrete set of points in the four dimensions of space and time. Further, such observations may generally be indirect, in the sense that they actually describe some quantity that is related to the uncertain model quantity of interest. For example, this indirect relationship might arise from inexact correspondence such as scaling differences (e.g., point scale observations are made of spatially distributed soil hydraulic properties, whereas the model representation describes the mean spatial value over some larger scale). Alternatively, it may arise from observing some closely related quantity (e.g., remotely sensed observations are made of radiances which are then related to the system properties of interest via radiative-transfer models). By considering all such factors, we define the following general observation equation:

$$z_k = M^z(y_k, x_k, u_k, x_0, \theta, B, M). \quad (10)$$

[36] To solve the DA problem, we are now interested in the posterior probability distribution (PPD) of the various

quantities of interest: the model structure, parameters, state variables, and outputs. By application of Bayes' theory it can be shown that for given observations z ,

$$p_{\text{posterior}}(M|z) = \frac{p(z|M) \cdot p_{\text{prior}}(M)}{p(z)}. \quad (11)$$

[37] Equation (11) provides a means for identifying appropriate model structures M by describing the posterior probability associated with a selected model structure in terms of the "likelihood" $p(z|M)$ that the observations z might have been generated by the model assuming that the structure M is the correct one, multiplied by the probability that the model structural assumption is correct ($p_{\text{prior}}(M)$). Acknowledging the logical progressive chain of conditional dependence described as $\{M \rightarrow \theta \rightarrow x \rightarrow y\}$, we can further derive

$$p_{\text{posterior}}(\theta|z, M) = \frac{p(z|M, \theta) \cdot p_{\text{prior}}(\theta|M) \cdot p_{\text{prior}}(M)}{p(z, M)} \quad (12)$$

$$\begin{aligned} p_{\text{posterior}}(x|z, M, \theta) \\ = \frac{p(z|x, M, \theta) \cdot p(x|M, \theta) \cdot p_{\text{prior}}(\theta|M) \cdot p_{\text{prior}}(M)}{p(z, M, \theta)} \end{aligned} \quad (13)$$

$$\begin{aligned} p_{\text{posterior}}(y|z, x, M, \theta) \\ = \frac{p(z|y, x, M, \theta) \cdot p(y|x, M, \theta) \cdot p(x|M, \theta) \cdot p_{\text{prior}}(\theta|M) \cdot p_{\text{prior}}(M)}{p(x, z, M, \theta)}. \end{aligned} \quad (14)$$

[38] For simplicity of presentation we have ignored the additional dependence on the system boundary B and initial conditions x_0 . In equation (11), $p(z)$ is a constant that normalizes the posterior probability mass to unity. Equation (12) describes how to compute posterior estimates of the model parameters θ given the model structure M and observations z ; equation (13) describes how to compute posterior estimates of the model states x given the model structure M , parameters θ , and observations z ; and equation (14) describes how to compute posterior estimates of the model outputs y given the model structure M , parameters θ , states x , and observations z . In data assimilation these equations serve as the fundamental basis for system identification (equation (11)), parameter estimation (equation (12)), state estimation (equation (13)), and quantification of uncertainty in hydrologic predictions (equation (14)).

5. Methods for System Identification

[39] In hydrologic modeling or analysis, a system identification problem typically involves selecting or constructing a valid model structure or a set of equally valid model structures (i.e., conceptual models and mathematical implementations) for the hydrologic system of concern. Historically, hydrologic modeling has relied on a single conceptual model of a particular hydrologic environment. *Beven and Freer* [2001, p. 1] point out that for a complex environmental system, there may actually exist "many different

model structures and many different parameter sets within a chosen model structure that may be behavioral or acceptable in reproducing the observed behavior of that system,” a phenomenon that *Beven* [1993] has termed as “equifinality.” This may be partially, if not primarily, due to the limited ability of current conceptual models in representing the complex, heterogeneous hydrologic systems that have unknown, and possibly unique system characteristics [Beven, 2000]. In this sense, hydrologic predictions based on a single conceptual model or model structure are invariably subject to statistical bias (if an invalid model is chosen) and underestimation of uncertainty (if equivalent valid models are not included) [Neuman, 2003]. Hence the “system identification” problem in hydrologic modeling can be approached through using a suite of “independent” plausible model structures with probability of each structure properly defined so that collectively, these model structures adequately and unambiguously approximate the true underlying system.

[40] In most data assimilation techniques (such as those described later in sections 6 and 7), errors in model structures are usually accounted for by adding an (unbiased) error term to the model transition equation (see section 7.1). However, because of equifinality of models as described above, a full consideration of the model structure error requires involving at least several “independent” alternative model structures that encompass a range of different assumptions [Beven and Young, 2003]. In this sense, a multimodel approach based on a suite of conceptual models is better suited for handling uncertainty associated with model structure errors than single-model approaches [National Research Council, 2001; Neuman, 2003; Georgakakos et al., 2004]. In the hydrologic literature, *Beven and Binley* [1992] introduced the generalized likelihood uncertainty estimation methodology (GLUE) where multiple competing model structures and parameter sets are allowed to account for the possibility of equifinality of models, producing a likelihood-weighted probability distribution of output predictions. GLUE is described in more detail as a model calibration and uncertainty estimation methodology in section 6.2.1.

[41] Along the same line of reasoning, a coherent mechanism for handling structural uncertainty is the concept of Bayesian model averaging (BMA) [Hoeting et al., 1999]. In BMA, the posterior distribution of the prediction on a quantity y given the observation z is approximated by the weighted sum of the posterior distributions of a set of K independent (or mutually exclusive) models $\mathbf{M} = \{M_1, \dots, M_K\}$, i.e.,

$$p(y|z) = \sum_{k=1}^K p(y|M_k, z)p(M_k|z), \quad (15)$$

where the weights are determined by the posterior distributions of the models $p(M_k|z)$ given by Bayes’ theorem as expressed in (11), where the normalization factor $p(z)$ is obtained by

$$p(z) = \sum_{k=1}^K p(z|M_k)p(M_k) \quad (16)$$

and the likelihood of each model M_k (given by $p(z|M_k)$) is calculated as

$$p(z|M_k) = \int p(z|\theta_k, M_k)p(\theta_k|M_k)d\theta_k. \quad (17)$$

[42] In equations (16) and (17), $p(z|M_k)$ is the likelihood of observing the data z given the model M_k ; $p(z|\theta_k, M_k)$ is the joint likelihood of model M_k and its parameter set θ_k ; $p(\theta_k|M_k)$ is the prior density of θ_k given the model structure M_k ; and $p(M_k)$ is the prior probability that the model structure M_k is valid. The BMA framework provides a cohesive way to jointly assess model structure and parameter uncertainties; however, it tends to be computationally demanding/cumbersome and also requires reliable prior information about model parameters. *Neuman* [2002, 2003] proposed a maximum likelihood version of BMA (MLBMA) that proves to be more computationally feasible and capable of dealing with situations where reliable prior information is lacking [Ye et al., 2004, 2005].

[43] In atmospheric science it has recently become very popular to use a multimodel ensemble method (MME) for weather and climate forecasting [e.g., *Doblas-Reyes et al.*, 2000; *Palmer et al.*, 2000; *Ziehmann, 2000*; *Palmer, 2004*; *Hagedorn et al.*, 2005a, 2005b]. In an MME approach a larger ensemble of predictions is composed from a suite of smaller ensembles, each generated based on an independent, plausible model (i.e., several ensembles are generated using each model structure). Instead of computing the probability of each model as in a BMA approach, the goal of MME is to account for uncertainty in the model structure, the assimilated data, and, in particular, the uncertainty associated with knowledge of initial conditions, by means of sampling from the output distributions of several different models. Most MME-based studies have reported that the performance of (properly selected) multimodel ensembles is superior to that of single-model ensembles, due not only to error compensation among different models, but also to the greater consistency and reliability of multimodel ensembles that cover a broad range of possible solutions [e.g., *Georgakakos et al.*, 2004; *Hagedorn et al.*, 2005a, 2005b].

[44] For all the three methodologies mentioned above (i.e., GLUE, BMA, and MME), all probabilities (including the final posterior) are implicitly conditioned on the set of selected models \mathbf{M} . Hence it is critical to select a set of relatively independent, plausible models that are most strongly supported by available data. Otherwise, there is no confidence about whether uncertainty is overestimated or underestimated, and there is no guarantee that the truth will even lie within the range of a model ensemble. This, however, is not straightforward, for there exist no well-accepted guidelines in the literature about how to define “independent” model structures or how many “independent” models are needed to adequately span the model space.

6. Methods for Parameter Estimation

[45] Despite the physical basis of many hydrological models, their parameters are often conceptual, effective quantities that cannot be measured in the field, and must therefore be estimated indirectly. The parameter estimation

problem is referred to by different names in the literature, including model calibration, parameter optimization, data assimilation, inverse problem, parameter tuning, among others. Arguably, an adequate parameter sensitivity analysis should always precede a parameter estimation study to identify sensitive parameters, for including insensitive parameters may render a parameter estimation process ineffective and cumbersome, especially for a complex model that has a large number of parameters [e.g., *Liu et al.*, 2004, 2005]. In this section we review traditional, deterministic model calibration methods as well as the newly emerging, stochastic data assimilation methods for parameter estimation.

6.1. Model Calibration Methods

[46] As an illustration to the general concept of model calibration, we consider a physically based model with p parameters ($\theta = \{\theta_1, \dots, \theta_p\}$), which is to be calibrated by assimilating the information from N different time series of observations $\{Z_n, n = 1, \dots, N\}$ corresponding to N model outputs $\{Y_n, n = 1, \dots, N\}$. The parameter estimation problem can be most generally stated as a vector optimization problem as follows [*Gupta et al.*, 1998]:

$$\text{Minimize } \mathbf{F}(\theta) = \{f_n(\theta), n = 1, \dots, N\} \text{ subject to } \theta \in \Theta, \quad (18)$$

where $f_n(\theta)$ is an objective function (also called a criterion) for measuring the distance between the n th model output and the n th observation; Θ is the physically feasible p -dimensional parameter space; and $\mathbf{F}(\theta)$ is a vector in the case of a multiobjective parameter estimation problem ($N \geq 2$) and a scalar in single-objective cases ($N = 1$).

[47] The two major strategies used for parameter estimation have been the “manual-expert” approach and the “automatic” approach. While manual-expert strategies rely on the informed but subjective judgment and skill of an experienced hydrologist, automatic strategies utilize the power of computer-based optimization techniques based in nonlinear regression theory. With the emergence of increasingly complex hydrological models with larger numbers of model parameters, effective and efficient automatic approaches have become more popular than the time-consuming, expertise-demanding manual approaches. *Duan et al.* [1992] introduced the Shuffled Complex Evolution algorithm (SCE), a global optimization strategy applicable to a broad class of single-criterion calibration problems. This algorithm was extended to the multiobjective complex optimization method (MOCOM) by *Yapo et al.* [1998], thereby enabling the use of multiple complementary measures for better extraction of information from the data, resulting in improved parameter estimates.

[48] The single- and multiple-criteria methods for parameter estimation mentioned above rely on deterministic nonlinear optimization techniques that seek to identify a single (few) “best” parameter set (sets), thus implicitly ignoring the uncertainties associated with observed data, model structure and parameters. In the case of significant system and data noise or bias, such methods can lead to parameter estimates that provide biased model predictions. Recently, *Vrugt et al.* [2003a] presented an efficient Markov Chain Monte Carlo (MCMC) sampler called the Shuffled Complex Evolution Metropolis algorithm (SCEM, with a multiobjective extension MOSCEM presented by *Vrugt et*

al. [2003b]), which converges to an ensemble of parameter sets that approximates the posterior distribution of model parameters. This posterior description of parameter uncertainty obtained through SCEM or MOSCEM can be used to assess the uncertainty in hydrological outputs arising from parameter uncertainty, representing an improvement over traditional deterministic optimization methods (e.g., SCE and MOCOM) in accounting for uncertainties associated with model parameters.

[49] Nevertheless, application of any of the methods mentioned above is implicitly based on an assumption that there exists a feasible parameter set for which the specific model structure under consideration is able to provide unbiased estimates of the model states and outputs at each time step. When this is not true (as is generally the case), we must acknowledge the existence of model structural and data errors and combine the (stochastic) parameter estimation methods with methods for system identification as described below in section 6.2. In addition, it should be mentioned that most (traditional) parameter estimation methods do not exploit the full power of the Bayesian framework, because they rely on “batch” processing of long-term historical data, and therefore lack the ability to recursively reduce parameter (and hence prediction) uncertainty as new data become available. An exception is the data-based mechanistic (DBM) approach to stochastic modeling, which is based on advanced recursive methods of time series analysis and has been successfully applied to hydrological systems modeling and data assimilation [e.g., *Young*, 2003, and references therein]. When considered in Bayesian terms, the DBM approach has the advantage of quantifying the uncertainty in the model and the data without resort to Monte Carlo methods, resulting in comparatively simple online implementation for flood forecasting and warning [e.g., *Young*, 2002; *Romanowicz et al.*, 2006].

6.2. Parameter Estimation Based on Stochastic Methods

[50] In recognition of the two major limitations of the model calibration methods mentioned above, there has been recent growing interest in the use of stochastic, sequential data assimilation techniques for parameter estimation. Such techniques operate within the Bayesian updating framework for estimation of predictive uncertainty. Examples include the generalized likelihood uncertainty estimation method (GLUE, [*Beven and Binley*, 1992]), the Bayesian recursive estimation method (BaRE [*Thiemann et al.* [2001]]), and other more recent techniques for simultaneous state and parameter estimation (see relevant details in section 8).

6.2.1. Generalized Likelihood Uncertainty Estimation (GLUE)

[51] *Beven and Binley* [1992] introduced the generalized likelihood uncertainty estimation (GLUE) methodology for model calibration that takes into account the effects of uncertainty associated the model structure and parameters. A fundamental assumption underlying GLUE is the “equifinality” or “nonuniqueness” concept [*Beven*, 1993], where multiple model structures and many parameter sets within a chosen structure are considered equally likely as simulators of the system. In other words, it is assumed that there exists no optimal model or parameter set due to structural and parameter uncertainties. This has introduced a different

philosophy to the venue of model calibration where the primary goal had historically been identifying an optimal parameter set based on a single model.

[52] To implement the GLUE methodology, several alternative model structures are selected and appropriate prior parameter uncertainty distributions are assumed for each model. Samples are then taken from these parameter distributions (coupled with their corresponding model structures) to generate Monte Carlo simulations. To evaluate the degree of correspondence between each simulation and the observed system behavior, a likelihood value is calculated based on a predefined likelihood measure (i.e., a measure of goodness of fit). The likelihood values are then used to determine whether a model structure-parameter set is “behavioral” or “nonbehavioral” according to a subjectively defined threshold of likelihood values; and only behavioral model structure-parameter sets are retained to provide predictions of the system behavior. To assess the uncertainty associated with the predictions, weights of the behavioral sets of model structure and parameters are calculated by normalizing the corresponding likelihood values so that all the weights sum up to one; the distribution of these weights is then taken as the probabilistic distribution of the predicted variables to reflect the uncertainty impacts of structural and parameter errors on model predictions.

[53] The primary improvement of the GLUE methodology over the deterministic calibration methods lies in its ability to explicitly account for the combined effects of model structure and parameter uncertainty, by using multiple models and assuming proper prior distributions for each parameter. Moreover, when a new observation period arrives or there exist different observation types (quantitative or qualitative), the likelihood values can be updated to estimate the posterior distribution of parameter sets (and thus that of model predictions), based on Bayes’ theorem. One concern that has been raised is that the Bayesian equation may not properly apply in GLUE in certain cases, because in GLUE, a certain likelihood measure, or essentially an objective function, is used in place of a formal likelihood function that is consistent within the framework of Bayes’ theorem [e.g., *Thiemann et al.*, 2001; *Mantovan and Todini*, 2006]. In addition, in the GLUE procedure, uncertainties associated with input data and output data (i.e., data errors) are not explicitly and/or formally considered.

6.2.2. Bayesian Recursive Estimation (BaRE)

[54] *Thiemann et al.* [2001] introduced the Bayesian recursive parameter estimation (BaRE) methodology that poses the parameter estimation problem within the context of a formal Bayesian framework. Unlike in GLUE where error sources are only implicitly considered with a likelihood measure, BaRE makes strong, explicit assumptions about the characteristics of errors in the observations by using an exponential power density error model. Like in GLUE, proper parameter ranges and prior probability distributions are specified; and the Monte Carlo approach is used to sample from the predefined distributions to represent parameter uncertainty.

[55] Once the error model is defined and model structure-parameter selections are initialized from their prior distributions, the BaRE methodology consists of two recursive steps that are common to the other data assimilation methods for state estimation (see section 7): prediction

and update. At time t_k , BaRE predicts the outputs and the uncertainty in the outputs by running the model forward to the next observation time t_{k+1} (i.e., when the observation z_{k+1} is available) for each set in the model structure-parameter ensemble. To update the probability of the model structure-parameter sets, a recursive version of the Bayesian equation for parameter estimation (equation (12)) is used to obtain the posterior probability of each model structure-parameter set i as follows:

$$P_{\text{posterior}}(M^i, \theta_{k+1}^i | z_{k+1}) \propto p(z_{k+1} | M^i, \theta_k^i) p_{\text{prior}}(\theta_k^i | M^i) p_{\text{prior}}(M^i). \quad (19)$$

[56] After updating, the model system continues to run forward to the next observation time, using the posterior model structure-parameter distribution at time t_{k+1} as the prior distribution. With a well-posed modeling system, this recursive process of conditioning parameters on available observations would gradually reduce uncertainty associated with the model structure-parameter set and lead to a progressively smaller region of high probability density (HPD) in the model-parameter space. In some cases, the sampling limitation of the Monte Carlo approach may lead to the HPD parameter region converging to one single point [*Beven and Young*, 2003; *Gupta et al.*, 2003]. *Misirli* [2003] proposed an improvement on the BaRE methodology by including a resampling technique to reduce the effect of the sampling limitation.

[57] Like GLUE, the BaRE methodology introduced a broader paradigm for parameter estimation without resorting to traditional optimization techniques. By adopting a recursive rather than “batch” approach, BaRE allows model parameters to behave as though time-variant and also reduces the dependence on availability of substantial input and output data before estimation can begin. More important, BaRE explicitly considers the uncertainties associated with model-parameter selection and output measurements, which has not been possible for most previous model calibration studies through parameter optimization, and explicitly represents these in the state and output predictions.

[58] Nevertheless, the BaRE methodology is not the final word on what can be achieved for model-parameter estimation. First of all, input data uncertainty and model structural uncertainty are not specifically separated out and are only implicitly considered, by expanding the predictive uncertainty bounds in a somewhat subjective manner. In addition, in the current BaRE methodology for which parameter estimation is the primary focus, the outputs and associated uncertainty remain un-updated after the posterior parameter distributions are obtained; in other words, the effects of reduction on parameter uncertainty (through incorporating new knowledge from available observation) do not properly propagate to the estimation of outputs and associated uncertainty in a timely manner. Accordingly, it would be beneficial to conduct simultaneous state and parameter estimation to generate unbiased parameter estimates, as well as more accurate state estimates. Several such approaches are reviewed in section 8. Finally, given that system structures and model parameters naturally vary slowly in time, it would be more appropriate to employ a time interval

sufficiently larger than the typical observation time step when performing model structure-parameter estimation. In other words, better results may be achieved by adopting an estimation algorithm that combines the advantages of batch and recursive methods through using an assimilation time interval of proper length.

7. Methods for State Estimation

[59] State estimation for dynamic systems is a process where information is extracted from observations and accumulated in time into the model, propagating to all state variables. For a well-behaved model with consistent constraints of physical properties of the system, improved state estimates can be obtained through data assimilation. This section focuses on state estimation methods assimilating observations that are distributed in time. Given observations available up to the current time, there are three types of state estimation problems: (1) smoothing problems that seek to characterize system states at a past time; (2) filtering problems that seek to characterize system states at the current time; and (3) forecasting problems that seek to characterize system states at a future time point [Gelb, 1974; McLaughlin, 2002]. Smoothing problems are usually found in reanalysis or retrospective studies, while filtering and forecasting problems are most commonly seen in real-time or operational forecasting applications. In dealing with these problems, batch-processing methods (or smoothers) are employed to estimate model states in a batch mode through least squares approximations, while sequential methods (or filters) are typically used for recursive estimation/correction of the states of a system each time an observation becomes available. In hydrologic data assimilation the most commonly used methods are Kalman filtering, particle filtering, and variational data assimilation. These methods are explained in detail below, with an introduction to the state-space formulation commonly used for state estimation applications.

7.1. State-Space Formulation

[60] For the convenience of illustrating the different state estimation methods, let us consider the following generic dynamic state-space formulation of a stochastic model:

$$x_{k+1} = M_{k+1}(x_k, \theta, u_{k+1}) + \eta_{k+1} \quad (20)$$

$$z_{k+1} = H_{k+1}(x_{k+1}, \theta) + \varepsilon_{k+1}, \quad (21)$$

where x_k and x_{k+1} represent the true system state vectors at time t_k and t_{k+1} , respectively; the nonlinear operator M_{k+1} (equivalent to the model structure mentioned earlier in section 3) expresses the system propagation from time t_k to t_{k+1} in response to the model input vector u_{k+1} ; θ is a vector of time-invariant model parameters; the observation vector z_{k+1} is related to the model parameters and states through an observation operator H_{k+1} (equivalent to M^f mentioned in equation (10)); η_{k+1} denotes the model error with mean $\bar{\eta}_{k+1}$ and covariance Q_{k+1} ; and ε_{k+1} denotes the observation error with mean $\bar{\varepsilon}_{k+1}$ and covariance R_{k+1} . In the context of Bayesian updating (equation (13)), the state equation (20) represents the model prior at time t_{k+1} , while the observation equation (21) can be used to calculate the likelihood of the

observation z_{k+1} . Note in the literature, the state equation is also referred to as “transition equation,” “forward model,” “forecast model,” or “dynamic system”; and the observation equation is often referred to as “measurement equation/model/system.”

[61] To set up the assimilation system using the above state-space formulation, some assumptions have to be made on the statistics of the two error terms η and ε , based on the prior knowledge of the deficiencies in the assimilating system. For example, the mean values of η and ε (i.e., biases) reflect the systematic errors in the modeling and observation systems, while the error covariances Q_k and R_k in particular reflect the uncertainty in the model predictions and observations. In practice, since these error characteristics cannot be observed directly and are difficult to estimate via indirect methods such as calibration, approximations to the error PDFs are typically unavoidable [e.g., Reichle *et al.*, 2001a, 2001b]. One popular approach is to assume that the errors are zero-mean white noise sequences with a normal (i.e., Gaussian) probability distribution. In addition, it is typically assumed that the model error and observation error are uncorrelated in order to obtain optimal estimates.

7.2. Kalman Filtering

[62] In the case of Gaussian model and measurement errors and linear model and observation operators, the data assimilation problem presented in (20) and (21) can be easily solved by an optimal recursive data processing algorithm known as the Kalman filter (or KF [Kalman, 1960]). The KF algorithm originates from the optimal least squares analysis and consists of recursive implementation of a prediction step (equations (22) and (23)) and an update step (equations (24) and (25)) as follows:

$$x_{k+1}^- = M_{k+1}(x_k^+, \theta, u_{k+1}) \quad (22)$$

$$P_{k+1}^- = \mathbf{M}_{k+1} P_k^+ \mathbf{M}_{k+1}^T + Q_{k+1} \quad (23)$$

$$x_{k+1}^+ = x_{k+1}^- + K_{k+1} d_{k+1} \quad (24)$$

$$P_{k+1}^+ = P_{k+1}^- - K_{k+1} \mathbf{H}_{k+1} P_{k+1}^-, \quad (25)$$

where P is the error covariance matrix of the state variables; \mathbf{M} and \mathbf{H} stand for the linear (or “linearized” in nonlinear cases) model operator and observation operator presented in matrix forms, respectively; the minus and plus superscripts are used to discriminate the states and the error covariance matrix before and after updating, respectively; T stands for transpose; and d is the innovation vector and is defined as the difference between the actual observation z and the model forecast of z (denoted as z^-), i.e.,

$$d_{k+1} = z_{k+1} - z_{k+1}^- \quad (26)$$

$$z_{k+1}^- = H_{k+1}(x_{k+1}^-, \theta). \quad (27)$$

K is called the Kalman gain and can be calculated as follows:

$$K_{k+1} = \frac{P_{k+1}^- \mathbf{H}_{k+1}^T}{\mathbf{H}_{k+1} P_{k+1}^- \mathbf{H}_{k+1}^T + R_{k+1}}. \quad (28)$$

[63] The calculations of (22)–(28) can be repeated at the next time step $k + 2$ to assimilate a new observation available at that time; and this process can progress sequentially into the future to assimilate all available observations if desired. Note by updating the states with equation (24), the assimilation algorithm does not explicitly comply with fundamental physical principles such as conservation of mass, momentum, and energy within the model system.

[64] Equation (28) shows that the Kalman gain K is determined by the relative magnitudes of the state error covariance P and the observation error covariance R and acts as a weighting factor on the innovation term. In other words, the larger the observation error covariance, the smaller the Kalman gain, and the smaller the update correction applied to the forecast state vector. This indicates that the assimilation results can be highly sensitive to the choice of the priors, i.e., the statistics of model structural, parameter, and measurement errors. It is worth noting that low correlation between model states and observations will also result in a small Kalman gain, suggesting the importance of using appropriate observations in an assimilation study.

[65] The KF algorithm described above is easy to implement and has proved effective and efficient in the case of linear system dynamics [e.g., *Eigbe et al.*, 1998; *Galantowicz et al.*, 1999]. However, in practice, hydrologic systems are often inevitably highly nonlinear, limiting the use of Kalman filtering. Hence variations of the KF algorithm have been developed to make it applicable to nonlinear problems, including the commonly used extended Kalman filter (EnKF [Jazwinski, 1970]) and ensemble Kalman filter (EnKF [Evensen, 1994]).

[66] In the EKF algorithm, local (tangent linear) approximation of the nonlinear state and measurement equations (i.e., the model operator M and the observation operator H) is performed each time data assimilation is conducted. When implementing the EKF, the same equations (22)–(27) for the KF algorithm will be used; however, the linearized forms of the model and observation operators (\mathbf{M} and \mathbf{H}) will be used in those equations. Some successful applications of the EKF have been seen in the hydrological literature [Katul et al., 1993; Entekhabi et al., 1994; Walker and Houser, 2001]; the EKF, however, may produce instabilities or even divergence due to closure approximation by neglecting the second- and higher-order derivatives of the model [Evensen, 1994].

[67] Evensen [1994] introduced the ensemble Kalman filtering (or EnKF) algorithm as an alternative to the EKF to address difficulties arising from high-dimensional nonlinear filtering problems. By making a Monte Carlo generation from random input perturbations, EnKF nonlinearly propagates an ensemble of model states using (20), maps them to an ensemble of prior estimates of the observations using (21), and then updates the prior ensemble based on the Kalman gain. The EnKF still consists of a prediction step

(equation (29)) and an update step (equation (30)) as follows:

$$\mathbf{x}_{k+1}^{-i} = M_{k+1}(\mathbf{x}_k^{+i}, \theta, \mathbf{u}_{k+1}^i) \quad i = 1, \dots, n \quad (29)$$

$$\mathbf{x}_{k+1}^{+i} = \mathbf{x}_{k+1}^{-i} + K_{k+1} d_{k+1}^i \quad i = 1, \dots, n, \quad (30)$$

where n is the size of the ensemble; the input ensemble \mathbf{u}_{k+1}^i is obtained by adding a noise term ζ_{k+1}^i to the nominal input \mathbf{u}_{k+1} , i.e., $\mathbf{u}_{k+1}^i = \mathbf{u}_{k+1} + \zeta_{k+1}^i$ ($\zeta_{k+1}^i \sim N(0, U_{k+1})$); and U_{k+1} is the error covariance of \mathbf{u}_{k+1} . A noise term ε_{k+1}^i can also be added to the nominal observation z_{k+1} to calculate the innovation ensemble using the following two equations (as compared with (26) and (27)):

$$d_{k+1}^i = z_{k+1} + \varepsilon_{k+1}^i - z_{k+1}^{-i} \quad \varepsilon_{k+1}^i \sim N(0, R_{k+1}), i = 1, \dots, n \quad (31)$$

$$z_{k+1}^{-i} = H_{k+1}(\mathbf{x}_{k+1}^{-i}, \theta) \quad i = 1, \dots, n. \quad (32)$$

[68] Unlike in the EKF, no linearization of M or H is needed. More important, the prior (or prediction) error covariance P_{k+1}^- of the state variables can be directly calculated from the ensemble $\{\mathbf{x}_{k+1}^{-i}\}$ as expressed in (33), saving substantial computation resources in propagating and updating P using (23) and (25),

$$P_{k+1}^- = \Sigma_{k+1}^{\text{xx}} = E[\mathbf{X}_{k+1}^- (\mathbf{X}_{k+1}^-)^T], \quad (33)$$

where Σ denotes covariance and $\mathbf{X}_{k+1}^- = \{\mathbf{x}_{k+1}^{-i}\}_{i=1}^n$. In fact, the state error covariance P is never explicitly needed in EnKF, for the $P_{k+1}^- H_{k+1}^T$ term in (28) is essentially the cross error covariance of the state prediction $\{\mathbf{x}_{k+1}^{-i}\}$ and the observation prediction $\{z_{k+1}^{-i}\}$, i.e.,

$$P_{k+1}^- H_{k+1}^T = \Sigma_{k+1}^{\text{xz}} = E[\mathbf{X}_{k+1}^- \mathbf{Z}_{k+1}^-], \quad (34)$$

where $\mathbf{Z}_{k+1}^- = \{z_{k+1}^{-i}\}_{i=1}^n$. Similarly, the prediction error covariance in the observation space (i.e., the $H_{k+1} P_{k+1}^- H_{k+1}^T$ term in (28)) can be calculated from $\{z_{k+1}^{-i}\}$ as follows:

$$H_{k+1} P_{k+1}^- H_{k+1}^T = \Sigma_{k+1}^{\text{zz}} = E[\mathbf{Z}_{k+1}^- (\mathbf{Z}_{k+1}^-)^T]. \quad (35)$$

Consequently, the Kalman gain in the EnKF algorithm can be easily derived by substituting (34) and (35) into the following equation:

$$K_{k+1} = \Sigma_{k+1}^{\text{xz}} (\Sigma_{k+1}^{\text{zz}} + R_{k+1})^{-1}. \quad (36)$$

Similar to the standard KF and EKF algorithms, EnKF can also be implemented recursively in time to sequentially assimilate observations as they become available.

[69] The applicability to nonlinear problems and easy implementation of the EnKF method has led to extensive applications of this DA technique in hydrology, meteorology, and other fields [e.g., *Burgers et al.*, 1998; *Margulis et al.*, 2002; *Reichle et al.*, 2002a, 2002b; *Moradkhani et al.*, 2005a; *Vrugt et al.*, 2005].

7.3. Particle Filtering

[70] Particle filtering (PF) is another commonly used data assimilation algorithm for recursive estimation of model states. In the literature the algorithm is also known as bootstrap filtering, the condensation algorithm, sequential Monte Carlo (SMC) sampling, interacting particle approximations, and survival of the fittest [Arulampalam *et al.*, 2002]. In particle filtering, the posterior probability distribution (PPD) of model states at time t_{k+1} is characterized by a set of discrete random particles ($\{x_{k+1}^i\}_{i=1}^n$) with associated importance weights ($\{w_{k+1}^i\}_{i=1}^n$) as follows:

$$p(x_{k+1}|z_{1:k+1}) \approx \sum_{i=1}^n w_{k+1}^i \delta(x_{k+1} - x_{k+1}^i), \quad (37)$$

where n is the number of particles and δ denotes the Dirac delta function. If n is sufficiently large, the discrete expression on the left-hand side of (37) becomes an effective approximation to the PPD of the true state space at time t_{k+1} .

[71] The PPD is best represented if the particles are directly sampled from the posterior distribution of the states, which, however, is generally not possible. To circumvent this obstacle, a sequential importance sampling (SIS) strategy has typically been adopted, where a proposal distribution $q()$ (referred to as importance density in the literature) is used and the importance weights are calculated as follows:

$$w_{k+1}^{i(*)} = p(x_{k+1}^i|z_{1:k+1})/q(x_{k+1}^i|z_{1:k+1}) \quad i = 1, \dots, n, \quad (38)$$

where $\{w_{k+1}^{i(*)}\}_{i=1}^n$ are the weights before normalization (i.e., $w_{k+1}^i = w_{k+1}^{i(*)}/\sum_{i=1}^n w_{k+1}^{i(*)}$). In practice, equation (38) can be rearranged as below to allow recursive evaluation of the importance weights as successive observations become available (see Arulampalam *et al.* [2002] for detailed derivation):

$$w_{k+1}^{i(*)} = w_k^{i(*)} \frac{p(z_{k+1}|x_{k+1}^i)p(x_{k+1}^i|x_k^i)}{q(x_{k+1}^i|x_k^i, z_{k+1})} \quad i = 1, \dots, n. \quad (39)$$

[72] Choice of an appropriate proposal importance density is crucial in the SIS algorithm as reported by several studies [e.g., Doucet *et al.*, 2000; Arulampalam *et al.*, 2002]. In a generic approach the importance density is often conveniently chosen to be the prior; and the weight calculation in (39) simplifies to

$$w_{k+1}^{i(*)} = w_k^{i(*)} p(z_{k+1}|x_{k+1}^i) \quad i = 1, \dots, n. \quad (40)$$

This renders the importance weights proportional to the likelihood $p(z_{k+1}|x_{k+1}^i)$ calculated using the observation equation (21).

[73] Particle filtering based on the above SIS algorithm consist of recursively propagating the particles using (20) and updating the importance weights associated with each particle using (21) and (40) as successive observations become available in time. Compared with the Kalman filtering algorithms discussed earlier (i.e., the standard KF,

EKF, and EnKF), PF performs updating on the particle weights instead of the state variables. In addition, PF has the desirable characteristics of being applicable to any state-space model of any form, linear or nonlinear, Gaussian or non-Gaussian.

[74] Implementation of the SIS particle filter in practice, however, may often be complicated by the well-known degeneracy problem where many particles are found to have negligible weights after a few iterations, thus making little or no contribution to the final representation of the posterior distribution [Doucet *et al.*, 2000] (Note that this same problem arose in the implementation of the BaRE algorithm, which has conceptual and implementational similarities.) As a result, only a small number of particles effectively participate in the filtering process according to the following measure [Doucet *et al.*, 2000; Arulampalam *et al.*, 2002]:

$$N_{eff} \approx 1 / \sum_{i=1}^n (w_{k+1}^i)^2, \quad (41)$$

where N_{eff} is the effective sample size that can be used to measure the degree of degeneracy in the filter. In general, the required number of particles n is likely to increase with the dimension of the state vector, the overlap between the prior and the likelihood, and the required number of time steps for filter operation; there exists, however, no universal provable criterion for defining the minimum effective sample size required to achieve a satisfactory approximation to the true PPD of the state vectors [Gordon *et al.*, 1993].

[75] In practice, to reduce the effect of the degeneracy problem, a resampling procedure is usually added to the SIS algorithm when there exists significant degeneracy (i.e., when N_{eff} is below a certain predefined threshold). The resampling step involves eliminating particles with small weights by replacing them with high-weight particles and then applying uniform weights to all the particles [e.g., Arulampalam *et al.*, 2002; Moradkhani *et al.*, 2005b]. When resampling is applied at each step (without evaluating N_{eff}), the standard SIS algorithm becomes the sampling importance resampling (SIR) filter, a special case of the SIS filter.

[76] Although resampling can reduce the effect of degeneracy, it also introduces another practical problem known as sample impoverishment due to loss of diversity among particles, especially for systems with small noises. In the case of severe sample impoverishment, all particles may converge to one single point in the state space, rendering poor final representation of the posterior distribution. Musso *et al.* [2001] introduced a modified PF known as the regularized particle filter (RPF) to solve the above problem by resampling from a continuous approximation to the importance density, instead of a discrete approximation as the SIR does.

[77] For more details on the implementation and applications of particle filtering and various SMC methods, the readers are referred to Gordon *et al.* [1993], Carpenter *et al.* [1999], Crisan *et al.* [1999], Doucet *et al.* [2001], Arulampalam *et al.* [2002], and Djurić *et al.* [2003].

7.4. Variational Data Assimilation (VDA)

[78] Unlike Kalman filtering and particle filtering, which approach the assimilation in a sequential manner, variational

methods operate in a batch-processing manner over a given time window which contains a sequence of observation time points. Hence variational methods are smoothers and mostly suitable for solving smoothing problems. Theoretically, VDA methods can also be used for filtering problems if a new smoothing problem is defined sequentially at each observation time point; this, however, can be computationally inefficient for real-time applications where the measurement vector z needs to be expanded indefinitely as new observations arrive continually. Depending on the spatial and temporal dimensions of the state variable, VDA methods can be one-dimensional (1D-Var), three-dimensional (3D-Var), or four-dimensional (4D-Var) (see unpublished lecture available at www.ecmwf.int/newsevents/training/rcourse_notes/pdf_files/Assim_concepts.pdf).

[79] For illustration purposes we assume that the prior estimate of state variables at time t_0 is x_0^- (with error covariance Q_0); and the assimilation is to operate over the time interval $[t_1, t_n]$, with observations $[z_1, z_2, \dots, z_n]$ available at the n discrete time points $[t_1, t_2, \dots, t_n]$. A general variational data assimilation problem can then be defined as the minimization of the following cost function J , which represents the aggregated error over the entire assimilation window (assuming that errors at different times are independent and additive):

$$\begin{aligned} J(x, u, \theta) = & \sum_{i=1}^n \eta_i^T Q_i^{-1} \eta_i + \sum_{i=1}^n (z_i - H_i[x_i])^T R_i^{-1} (z_i - H_i[x_i]) \\ & + (x_0 - x_0^-)^T Q_0^{-1} (x_0 - x_0^-) \\ & + \sum_{i=1}^n (u_i - u_i^-)^T C_{uu}^{-1} (u_i - u_i^-) \\ & + (\theta - \theta^-)^T C_\theta^{-1} (\theta - \theta^-) \\ = & J_M + J_O + J_0 + J_u + J_\theta, \end{aligned} \quad (42)$$

where η_i represents the model error at t_i ; u_i^- and θ^- denote the prior model inputs at t_i and the prior time-invariant parameters, respectively; and C_{uu} and C_θ are the time-invariant error covariances of inputs and parameters, respectively. The purpose of variational data assimilation is, by means of minimizing J , to obtain the least squares estimates of state variables x_i and input variables u_i for each time point within the assimilation window and the time-invariant parameters θ . The minimization problem is subject to the strong constraint that the state, input, and parameter estimates obtained by VDA must be consistent with the state equation (20). Alternatively, one can turn the constrained minimization problem into an unconstrained one by adjoining the state equation to the cost function (42) with a Lagrange multiplier λ as follows:

$$J(x, u, \theta) = J_M + J_O + J_0 + J_u + J_\theta + \sum_{i=1}^n \lambda_i^T [x_i - M_i(x_{i-1}, u_i, \theta)]. \quad (43)$$

[80] In the above general formulations of the cost function, the first term J_M penalizes the difference between the estimated model error vector η_i and its prior mean (assumed to be zero in this case); the second term J_O is used to penalize the differences between model predictions and

observations at all time points within the assimilation window; and J_0 , J_u , and J_θ are included to measure the errors associated with the initial conditions, model inputs and parameters, respectively. When summed together to form the aggregated cost function, each of the errors is weighted by the corresponding error covariance (i.e., Q , R , Q_0 , C_{uu} , or C_θ). In this general VDA framework, errors from various sources (e.g., the model, observations, initial conditions, inputs, and parameters) can be collectively taken into account.

[81] In practice, however, nonlinear, high-dimensional hydrologic applications render the comprehensive optimization problem as represented by (42) very difficult, and often impossible, to solve. Consequently, simplifications and approximations are often introduced by, for example, neglecting model/parameter errors and/or linearizing the state and observation equations. Even with simplifications, solving a VDA problem analytically is not easy, and often a numerical algorithm such as the adjoint model technique is used to obtain solutions in an iterative manner.

[82] To illustrate the implementation process of variational data assimilation, we consider a simple VDA system where the objective is to minimize the following cost function with only the measurement term J_O considered:

$$J(x) = J_O = \sum_{i=1}^n (z_i - H_i[x_i])^T R_i^{-1} (z_i - H_i[x_i]). \quad (44)$$

[83] According to the state equation (20), given θ , u_i , and η_i (assumed to be zero in this case), the state prediction at t_i (i.e., x_i) is solely dependent on prediction at the previous time step t_{i-1} (i.e., x_{i-1}), which is in turn solely dependent on x_{i-2} . This indicates that x_i is ultimately determined by the initial condition x_0 , the only fundamental unknown in this VDA problem. The objective of VDA is then to find the best estimate of x_0 that minimizes $J_O(x_0)$. The optimization process requires the evaluation of both the cost function and its gradient $\nabla J_O(x_0)$, which can be computed as follows using the adjoint technique (see detail derivations by *Huang and Yang* [1996]; see also unpublished note available at <http://citeseer.ist.psu.edu/huang96variational.html>):

$$\nabla J_O(x_0) = \sum_{i=1}^n \left[\prod_{k=0}^{i-1} \mathbf{M}_k^T \right] \mathbf{H}_i^T d_i \quad (45)$$

$$d_i = R_i^{-1} (z_i - H_i(x_i, \theta))$$

where \mathbf{M}_k^T is the transpose of the tangent linear model of M_k at point x_k ; similarly, \mathbf{H}_i^T denotes the transpose of the tangent linear model of H_i at point x_i and d_i is the normalized difference between model prediction and the observation at time t_i . For computational efficiency, we define the adjoint model at time t_i as

$$\tilde{x}_{i-1} = \mathbf{M}_i^T (\tilde{x}_i + \mathbf{H}_i^T d_i), \quad (46)$$

where \tilde{x}_i is called the adjoint variable. It can be proved that if we start from the end point of the assimilation interval t_n with \tilde{x}_n initialized to zero and then integrate the adjoint model

(46) backward in time to the initial time t_0 , we obtain \tilde{x}_0 , which is exactly equal to $\nabla J_O(x_0)$ defined by (45). With this method of computing the cost function gradient, a VDA problem can be solved through an iterative minimization process to identify a best estimate of x_0 , which can then be used to compute the value of x at any time point within the assimilation window by integrating the state equation (20) forward in time.

[84] A number of examples of designing and solving a VDA problem can be found in the literature. *Huang and Yang* [1996] discussed in detail the general procedure to construct a VDA system using the adjoint technique based on a nonlinear mathematical model, with only observation errors considered in the cost function. Applications of similar VDA techniques are given by *McLaughlin* [1995, 2002], *Bouttier and Courtier* [1999], *Reichle et al.* [2001a, 2001b], and *Seo et al.* [2003]. The readers are referred to these references for more details of using the VDA algorithms.

[85] Compared with the sequential equivalents KF and EKF, the VAR methods are preferable for data assimilation in a realistic, complex system (e.g., a numerical weather prediction framework) because they are much less expensive computationally than KF and EKF methods. In addition, by using observations inside the assimilation interval all at once, VDA methods are also more optimal than KF and EKF methods inside (within) the interval (at the end of the interval, VDA and KF methods are expected to give the same results for linear systems; in the presence of high nonlinearity, the results from the two methods may diverge because VDA gives the mode of an uncertain variable while KF estimates the expected value). However, the sequential KF methods are more suitable for real-time data assimilation to process observations that arrive continuously in time, while VDA methods can only be run for a finite time interval; also, KF methods provide error covariance estimates for the prediction, while a VDA method itself does not provide any estimate of the predictive uncertainty. When the assimilation system is nonlinear, both EKF and VDA methods rely on using the tangent-linear models M and H to approximate the state and observation equations; if the nonlinearity is important, it makes more sense to use ensemble (or Monte Carlo) approaches such as EnKF and PF for data assimilation.

8. Simultaneous State and Parameter Estimation

[86] In general, parameter estimation tends to focus on uncertainty in the parameter estimates only, while neglecting partial or all of the other uncertainty sources. On the other hand, state estimation via data assimilation methods, although having the potential for explicitly handling various uncertainties arising from model inputs and observations, typically does not take into account the uncertainties associated with model parameters. In either case, there is a tendency to generate biased model predictions due to biased parameter and/or state estimates. Hence it would be desirable to combine parameter estimation with state estimation to account for all kinds of uncertainties.

[87] *Vrugt et al.* [2005] applied a simultaneous optimization and data assimilation (SODA) approach to estimate both states and parameters of two hydrological models. The SODA approach estimates model parameters using the

batch calibration strategy SCEM, with EnKF updating of state estimates performed at each time step in each model run during the calibration process. This way of combining optimization with data assimilation is conceptually simple and easy to implement. Preliminary results show that the SODA approach is able to produce both less biased parameters and less biased model states, compared with the results from using only SCEM or EnKF. The SODA approach is certainly a step forward from traditional parameter estimation and data assimilation methodologies, in that it reasonably uses calibration to correct long-term systematic biases due to parameter uncertainties and uses ensemble data assimilation to correct short-term or instantaneous system biases associated with model states, data, and other sources of errors. The SODA approach, however, is still not wholly satisfactory in that it optimizes model parameters in one single batch, without allowing parameters to vary over time while also requiring considerable computational time.

[88] *Moradkhani et al.* [2005a, 2005b] presented two dual state-parameter estimation methods based on EnKF and PF, respectively. These two methods were designed to recursively estimate both states and parameters using two parallel filters. In these methods, Monte Carlo sampling and sequential updating (via EnKF or PF techniques) are applied to not only a vector of state variables, but also to a different vector of model parameters at each assimilation time step. Accordingly, the probability distributions of both model states and parameters are (independently) recursively updated each time a new observation is available. In these approaches, better state and parameter estimates enable the modeling system to evolve consistently over time and make improved predictions with proper uncertainty bounds. Along the same lines, *Labarre et al.* [2006] presented an approach to jointly estimate the model states and the hyperparameters of the data assimilation algorithm using the mutually interactive state and parameter estimation (MISP) technique [*Todini*, 1978a, 1978b] with two conditionally linked Kalman filters running in parallel.

[89] Another way of conducting joint state-parameter estimation is to extend the current state vector with the model parameters, a technique known as “state augmentation” [e.g., *Gelb*, 1974; *Drécourt et al.*, 2005]. Here the model parameters are recast as state variables to form an extended state vector; and the simultaneous state and parameter estimation problem is reduced to a state estimation problem. If we assume the parameters are time-variant with normally distributed errors ξ_k at time t_k ($\xi_k \sim N(0, V_k)$), then with state augmentation, the new model equation and observation equation can be expressed as follows (as opposed to the original equations (20) and (21)):

$$\begin{bmatrix} x_{k+1} \\ \theta_{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{M}_k & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} x_k \\ \theta_k \end{bmatrix} + \begin{bmatrix} \eta_k \\ \xi_k \end{bmatrix} \text{ or } x'_{k+1} = \mathbf{M}'_k x'_k + \eta'_k \quad (47)$$

$$\begin{bmatrix} y_{k+1} \\ \theta_{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_k & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} x_k \\ \theta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_k \\ \xi_k \end{bmatrix} \text{ or } y'_{k+1} = \mathbf{H}'_k x'_k + \varepsilon'_k, \quad (48)$$

where x' , y' , \mathbf{M}' , \mathbf{H}' , η' , and ε' are the new state vector, observation vector, model operator, observation operator, model error, and observation error, respectively. These new

quantities should be used in place of the old ones when a data assimilation algorithm (e.g., EKF or EnKF) is applied to recursively update the states and parameters simultaneously. Being conceptually simple, this method, however, may render the estimation process unstable and intractable because of complex interactions between states and parameters in nonlinear dynamic systems [Todini, 1978a, 1978b]. In addition, since parameters generally vary much more slowly than the system states, unstable problems may also result from the fact that both model states and parameters are updated at each observation time step in this method. This same argument may apply to the dual state-parameter estimation methods presented by Moradkhani *et al.* [2005a, 2005b].

9. An Integrated Uncertainty Framework for Hydrological Modeling

[90] The data assimilation methods introduced in sections 5–8 are designed for system identification, parameter estimation, state estimation, and combined state and parameter estimation, respectively. As described in section 4, Bayes' theorem is the fundamental basis of these DA methods. However, one criticism of Bayesian methods is that the computation of posteriors depends on prescribed priors that could be wrong, rendering the possibility of unrealistic uncertainty estimation. For example, state estimation is often conducted under the assumption that the model structure and parameters are correct (i.e., assuming the model/parameter priors to be unity), which is hardly the case in hydrologic modeling. To adequately quantify the total uncertainty in hydrologic predictions and to maximally reduce it, we shall consider an integrated uncertainty framework that can facilitate the implementation of all the three types of DA applications in a cohesive, systematic manner.

[91] Berliner [1996] introduced a Bayesian hierarchical modeling (BHM) approach to complex environmental DA problems, where one can obtain the joint distribution of the model process x and parameters θ , given observational data z , by computing a hierarchy of conditional models based on Bayesian rules as follows (see also Wike [2003]):

$$p(x, \theta | z) \propto p(z | x, \theta) p(x | \theta) p(\theta). \quad (49)$$

With the logical progressive chain of conditional dependence $\{M \rightarrow \theta \rightarrow x \rightarrow y\}$ as described in section 4, we can derive the following integrated hierarchical framework in a manner analogous to the BHM approach described above:

$$\begin{aligned} p(y, x, \theta, M | z^y, z^x, z^\theta, z^M) &\propto p(M | z^M) \quad (1) \text{system identification} \\ &\cdot p(\theta | z^\theta, M) \quad (2) \text{parameter estimation} \\ &\cdot p(x | z^x, \theta, M) \quad (3) \text{state estimation} \\ &\cdot p(y | z^y, x, \theta, M) \quad (4) \text{output prediction} \end{aligned} \quad (50)$$

where z^M , z^θ , z^x , z^y denote independent observations on which the model structure, parameters, states, and outputs are progressively conditioned, respectively. To assess the total uncertainty of a hydrological model, it is really critical

that the joint output-state-parameter-structure distribution $p(y, x, \theta, M)$ (instead of the individual distributions) be examined, considering the complex interactions between the model outputs, states, parameters, and structure. The purpose of the hierarchical approach is to obtain this complicated joint distribution, which is difficult to compute directly, by factoring it into a sequence of conditional probabilities that are easier to characterize utilizing available knowledge and data.

[92] In the framework defined in (50), one starts with (1) system identification by computing $p(M | z^M)$ given the observation z^M and some prior knowledge of the model parameters; (2) the model parameters are then estimated through a parameter estimation technique given a different set of observations z^θ and the models M (and their probability distribution obtained in the previous step; (3) with the observation z^x , and M and θ obtained in (2), the distribution of the model states x can be updated; and (4) finally, with M , θ , and x defined in the last three steps and the observation z^y , one can predict the final uncertainty reflected in the model outputs y . The joint distribution of the model outputs, states, parameters, and structures, which captures the complex interactions among these components, can be obtained by multiplying the four conditional probability terms together as shown in (50). When a period of new observations arrives, the above progressive steps can be repeated to update the conditional probability distributions and the overall joint distribution.

[93] In implementing this uncertainty framework, one should choose a most suitable data assimilation method for each of the four steps. In particular, special attention should be paid to the appropriate timescales when deciding which DA method to use. For example, as discussed in section 3, we expect model parameters to vary much more slowly in time than the states and outputs. In this sense, it would be appropriate to adopt a smoothing or variational approach for parameter estimation in step 2 and a sequential or filtering approach for state estimation in step 3. In principle, a method used for state estimation is expected to be applicable to output prediction as well, for model states and outputs generally share the same dynamic characteristics/frequencies. We can also reasonably assume that the model structure does not vary with time or varies at a very low frequency (even lower than that of the parameters). Precisely how to adjust the existing DA methods so that they can be cohesively nested within the above integrated hierarchical framework remains an important topic for further research.

10. Summary and Discussions

[94] Application of data assimilation techniques to hydrologic modeling is relatively new, and there is a lack of general guidance in the hydrologic literature on how to choose and implement a suitable DA technique so that the hydrologic uncertainty is properly considered. This may come to limit the extensive application of hydrologic data assimilation. On the other hand, it has not been realized by the general hydrologic community that the traditional data assimilation focus on state estimation alone is not sufficient for adequate consideration of the uncertainties associated with all sources in hydrologic modeling. In most cases,

uncertainty in model structures and parameters is ignored in data assimilation applications.

[95] In this analysis we have discussed the three critical aspects of addressing hydrologic uncertainty, namely, understanding, quantifying, and reducing uncertainty, to arrive at a general context for hydrologic data assimilation. The intention of this paper is to provide not an extensive review of all the data assimilation techniques and applications existing in the hydrologic literature, but a discussion of the main recent developments, potential future directions, and some open issues in hydrologic data assimilation. We explore uncertainties associated with different sources from a systems perspective, leading to the definition of the three major types of DA problems: system identification, parameter estimation, and state estimation. Bayesian techniques for addressing these uncertainty problems and typical methods used in the hydrologic literature were then described in relevant detail to provide sufficient guidance on how to properly implement these methods. To adequately quantify the hydrologic predictive uncertainty and reduce it to a maximum degree, we call for the adoption of an integrated framework such as the one proposed in section 9, where system identification, parameter estimation, state estimation, and ultimately output prediction are progressively conducted in a cohesive, systematic manner. Proper implementation of all these DA problems within such a single, integrated framework would greatly improve the effectiveness and efficiency in extracting information from available data and assimilating it into hydrologic predictions.

[96] Nevertheless, there remain critical issues that need to be properly addressed before the proposed integrated framework can be implemented to realize its maximum potential. For example, the exposition assumes that we have successfully described the information extraction process via the definition of suitable likelihood functions [e.g., *Beven and Young*, 2003; *Gupta et al.*, 2003], a topic that merits rethinking and further research. Also, the fundamental Bayesian rule requires that the error models used as inputs to data assimilation applications be properly prescribed from prior information, which is often difficult to satisfy for real-world hydrologic applications. Although we all recognize that real hydrologic systems are seldom linear or close to linear, there often has been no other choice but to prescribe the error distributions as Gaussians; better strategies (e.g., via using mixtures of Gaussians [*Wójcik et al.*, 2006]) are desired to avoid this subjectivity in prescribing various types of errors, and new mathematical developments might be necessary to circumvent this difficulty. Also, we must recognize that a hydrological system (i.e., a model structure) and its physical properties (i.e., model parameters) naturally tend to vary much more slowly in time than the states and fluxes of the system. In other words, proper timescales (or time intervals) should be identified and utilized in the different steps within the framework. In addition, implementing such an integrated framework would require a substantial amount of data, encompassing various observation types that are suitable for different kinds of DA problems including system identification, parameter estimation, and state estimation. Other issues to be addressed may include, for example, handling temporally and spatially correlated errors in hydrological variables [e.g., *Drécourt et al.*, 2005] and resolving the scale differences between model states and

observations. To facilitate addressing these issues, the next generation of hydrologic models should be developed in coordination with developments in DA techniques to facilitate the implementation of the proposed integrated hierarchical framework. Finally, it is useful to be aware that complete accounting for model structure errors might never be achieved, as there exists no means to define a model space (with a set of truly independent model structures) that can perfectly represent the reality.

[97] **Acknowledgments.** This study was supported by SAHRA (an NSF center for Sustainability of Semi-Arid Hydrology and Riparian Areas) under grant EAR-9876800. We thank Gab Abramowitz, Andras Bardossy, Martyn Clark, Shlomo Neuman, Amilcare Porporato, Peter Troch, Ezio Todini, Thorsten Wagener, and the anonymous reviewer for their constructive comments that greatly improved this paper.

References

- Abramowitz, G., H. Gupta, A. Pitman, Y. Wang, R. Leuning, H. Cleugh, and K.-L. Hsu (2006), Neural error regression diagnosis (NERD): A tool for model bias identification and prognostic data assimilation, *J. Hydrometeorol.*, 7, 160–177.
- Anderson, J. L., and S. L. Anderson (1999), A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts, *Mon. Weather Rev.*, 127, 2741–2758.
- Arulampalam, M. S., S. Maskell, N. Gordon, and T. Clapp (2002), A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, *IEEE Trans. Signal Processes*, 50(2), 174–188.
- Beck, M. B. (1987), Water quality modeling: A review of the analysis of uncertainty, *Water Resour. Res.*, 23(8), 1393–1442.
- Berliner, L. M. (1996), Hierarchical Bayesian time series models, in *Maximum Entropy and Bayesian Methods*, edited by K. Hanson and R. Silver, pp. 15–22, Springer, New York.
- Beven, K. J. (1993), Prophecy, reality, and uncertainty in distributed hydrological modeling, *Adv. Water Resour.*, 16, 41–51.
- Beven, K. J. (2000), Uniqueness of place and non-uniqueness of models in assessing predictive uncertainty, in *Computational Methods in Water Resources XIII*, edited by L. R. Bentley et al., pp. 1085–1091, A. A. Balkema, Rotterdam, Netherlands.
- Beven, K. J., and A. M. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Processes*, 6, 279–298.
- Beven, K. J., and J. Freer (2001), Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology, *J. Hydrol.*, 249, 11–29.
- Beven, K. J., and P. Young (2003), Comment on “Bayesian recursive parameter estimation for hydrologic models” by M. Thieman, M. Trosset, H. Gupta, and S. Sorooshian, *Water Resour. Res.*, 39(5), 1116, doi:10.1029/2001WR001183.
- Blöschl, G., and R. Grayson (2000), Spatial observations and interpolation, in *Spatial Patterns in Catchment Hydrology: Observations and Modeling*, edited by A. Grayson and G. Blöschl, pp. 17–50, Cambridge Univ. Press, New York.
- Boni, G., F. Castelli, and D. Entekhabi (2001), Sampling strategies and assimilation of ground temperature for the estimation of surface energy balance components, *IEEE Trans. Geosci. Remote Sens.*, 39(1), 165–172.
- Burgers, G. P., J. Van Leeuwen, and G. Evensen (1998), Analysis scheme in the ensemble Kalman filter, *Mon. Weather Rev.*, 126, 1719–1724.
- Carpenter, J., P. Clifford, and P. Fearnhead (1999), Improved particle filter for nonlinear problems, *IEE Proc. Radar Sonar Navig.*, 146(1), 2–7.
- Carrera, J., and S. P. Neuman (1986), Estimation of aquifer parameters under transient and steady state conditions: 3. Application to synthetic and field data, *Water Resour. Res.*, 22(2), 228–242.
- Clark, M. P., and A. G. Slater (2006), Probabilistic quantitative precipitation estimation in complex terrain, *J. Hydrometeorol.*, 7(1), 3–22.
- Courtier, P., J. Derber, R. Errico, J.-F. Louis, and T. Vukicevic (1993), Important literature on the use of adjoint, variational methods and the Kalman filter in meteorology, *Tellus, Ser. A*, 45, 342–357.
- Crisan, D., P. Del Moral, and T. J. Lyons (1999), Non-linear filtering using branching and interacting particle systems, *Markov Processes Related Fields*, 5(3), 293–319.
- Cullen, A. C., and H. C. Frey (1999), *Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing With Variability and Uncertainty in Models and Inputs*, Springer, New York.

- Daley, R. (1991), *Atmospheric Data Analysis*, Cambridge Univ. Press, New York.
- Djurić, P. M., J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Miguez (2003), Particle filtering, *IEEE Signal Processing Mag.*, 9, 19–38.
- Doblas-Reyes, F. J., M. Deque, and J.-P. Pielikevire (2000), Multi-model spread and probabilistic seasonal forecasts in PROVOST, *Q. J. R. Meteorol. Soc.*, 126(567), 2069–2088.
- Doucet, A., S. Godsill, and C. Andrieu (2000), On sequential Monte Carlo sampling methods for Bayesian filtering, *Stat. Comput.*, 10, 197–208.
- Doucet, A., N. D. Freitas, and N. Gordon (Eds.) (2001), *Sequential Monte Carlo Methods in Practice*, 581 pp., Springer, New York.
- Drécourt, J.-P., H. Madsen, and D. Rosbjerg (2005), Bias aware Kalman filters: Comparison and improvements, *Adv. Water Resour.*, 29(5), 707–718.
- Duan, Q., S. Sorooshian, and V. K. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, 28, 1015–1031.
- Dunne, S., and D. Entekhabi (2005), An ensemble-based reanalysis approach to land data assimilation, *Water Resour. Res.*, 41, W02013, doi:10.1029/2004WR003449.
- Eigbe, U., M. B. Beck, H. S. Wheatner, and F. Hirano (1998), Kalman filtering in groundwater flow modeling: Problems and prospects, *Stochastic Hydrol. Hydraul.*, 12, 15–32.
- Entekhabi, D., H. Nakamura, and E. Njoku (1994), Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely sensed observations, *IEEE Trans. Geosci. Remote Sens.*, 32, 438–448.
- Entekhabi, D., G. R. Asrar, A. K. Betts, K. J. Beven, R. L. Bras, C. J. Duffy, T. Dunne, R. D. Koster, D. P. Lettenmaier, and D. B. McLaughlin (1999), An agenda for land surface hydrology research and a call for the second international hydrological decade, *Bull. Am. Meteorol. Soc.*, 80(10), 2043–2058.
- Evensen, G. (1994), Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, 99, 10,143–10,162.
- Finkel, A. M. (1990), Confronting uncertainty in risk management: A guide for decision-makers, report, Cent. for Risk Manage., Resour. for the Future, Washington, D. C.
- Galantowicz, J., D. Entekhabi, and E. Njoku (1999), Tests of sequential data assimilation for retrieving profile soil moisture and temperature from observed L-band radiobrightness, *IEEE Trans. Geosci. Remote Sens.*, 37, 1860–1870.
- Gelb, A. (Ed.) (1974), *Applied Optimal Estimation*, 374 pp., MIT Press, Cambridge, Mass.
- Georgakakos, K. P., D.-J. Seo, H. V. Gupta, J. Schaake, and M. B. Butts (2004), Towards the characterization of streamflow simulation uncertainty through multimodel ensembles, *J. Hydrol.*, 298, 222–241.
- Gordon, N., D. Salmond, and A. F. M. Smith (1993), Novel approach to nonlinear and non-Gaussian Bayesian state estimation, *Proc. IEE, Part F*, 140, 107–113.
- Graham, W., and D. McLaughlin (1991), A stochastic model of solute transport in groundwater: Applications to the Borden Ontario tracer test, *Water Resour. Res.*, 27(6), 995–1004.
- Gupta, H. V., S. Sorooshian, and P. O. Yapo (1998), Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resour. Res.*, 34, 751–763.
- Gupta, H., M. Thieman, M. Trosset, and S. Sorooshian (2003), Reply to comment by K. Beven and P. Young on “Bayesian recursive parameter estimation for hydrologic models,” *Water Resour. Res.*, 39(5), 1117, doi:10.1029/2002WR001405.
- Hagedorn, R., F. J. Doblas-Reyes, and T. N. Palmer (2005a), The rationale behind the success of multi-model ensemble in seasonal forecasting: I. Basic concept, *Tellus, Ser. A*, 57, 219–233.
- Hagedorn, R., F. J. Doblas-Reyes, and T. N. Palmer (2005b), The rationale behind the success of multi-model ensemble in seasonal forecasting: II. Calibration and combination, *Tellus, Ser. A*, 57, 234–252.
- Hoeting, J. A., D. Madigan, A. E. Raftery, and C. T. Volinsky (1999), Bayesian model averaging: A tutorial, *Stat. Sci.*, 14(4), 382–417.
- Houser, P. R., W. J. Shuttleworth, J. S. Famiglietti, H. V. Gupta, K. H. Syed, and D. C. Goodrich (1998), Integration of soil moisture remote sensing and hydrological modeling using data assimilation, *Water Resour. Res.*, 34(12), 3405–3420.
- Huang, X.-Y., and X. Yang (1996), Variational data assimilation with the Lorenz model: High-resolution limited area model (HIRLAM), *Tech. Rep. 26*, 42 pp., Danish Meteorol. Inst., Copenhagen.
- Jazwinski, A. H. (1970), *Stochastic Processes and Filtering Theory*, Elsevier, New York.
- Kalman, R. E. (1960), A new approach to linear filtering and prediction problems, *J. Basic Eng.*, 82(1), 35–64.
- Katul, G. G., O. B. Wendroth, M. B. Parlange, C. E. Puente, M. Folegatti, and D. R. Nielsen (1993), Estimation of in situ hydraulic conductivity function from nonlinear filtering theory, *Water Resour. Res.*, 29(4), 1063–1070.
- Kitanidis, P. K., and R. L. Bras (1980), Real-time forecasting with a conceptual hydrologic model: 1. Analysis of uncertainty, *Water Resour. Res.*, 16(6), 1025–1033.
- Krupnick, A., R. Morgenstern, M. Batz, P. Nelson, D. Burtraw, J.-S. Shih, and M. McWilliams (2006), *Not a Sure Thing: Making Regulatory Choices Under Uncertainty*, 233 pp., Resour. for the Future, Washington, D. C.
- Krzysztofowicz, R. (1999), Bayesian theory of probabilistic forecasting via deterministic hydrologic model, *Water Resour. Res.*, 35(9), 2739–2750.
- Kuczera, G., and E. Parent (1998), Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The Metropolis algorithm, *J. Hydrol.*, 211(1–4), 69–85.
- Labarre, D., E. Grivel, Y. Berthoumieu, E. Todini, and M. Najim (2006), Consistent estimation of autoregressive parameters from noisy observations based on two interacting Kalman filters, *Signal Processing*, 86(10), 2863–2876.
- Lee, K. K., F. W. Deng, and J. H. Cushman (1993), Multiscale adaptive estimation of the conductivity field from head and tracer data, *Stochastic Environ. Res. Risk Assess.*, 7(1), 66–82.
- Liu, Y., L. A. Bastidas, H. V. Gupta, and S. Sorooshian (2003), Impacts of a parameterization deficiency on off-line and coupled land-surface model simulations, *J. Hydrometeorol.*, 4(5), 901–914.
- Liu, Y., H. V. Gupta, S. Sorooshian, L. A. Bastidas, and J. W. Shuttleworth (2004), Exploring parameter sensitivities of the land surface using a locally coupled land-atmosphere model, *J. Geophys. Res.*, 109, D21101, doi:10.1029/2004JD004730.
- Liu, Y., H. V. Gupta, S. Sorooshian, L. A. Bastidas, and W. J. Shuttleworth (2005), Constraining land surface and atmospheric parameters of a locally coupled model using observational data, *J. Hydrometeorol.*, 6, 156–172.
- Mantovan, P., and E. Todini (2006), Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology, *J. Hydrol.*, 330, 368–381.
- Margulis, S. A., D. McLaughlin, D. Entekhabi, and S. Dunne (2002), Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment, *Water Resour. Res.*, 38(12), 1299, doi:10.1029/2001WR001114.
- McLaughlin, D. (1995), Recent developments in hydrologic data assimilation, *U.S. Natl. Rep. Int. Union Geod. Geophys. 1991–1994, Rev. Geophys.*, 33, 977–984.
- McLaughlin, D. (2002), An integrate approach to hydrologic data assimilation: Interpolation, smoothing, and filtering, *Adv. Water Resour.*, 25, 1275–1286.
- McLaughlin, D., L. B. Reid, S. G. Li, and J. Hyman (1993), A stochastic method for characterizing groundwater contamination, *Ground Water*, 31(2), 237–249.
- Misirli, B. F. (2003), Improving efficiency and effectiveness of Bayesian recursive parameter estimation for hydrologic models, Ph.D. dissertation, Univ. of Ariz., Tucson.
- Montanari, A., and A. Brath (2004), A stochastic approach for assessing the uncertainty of rainfall-runoff simulations, *Water Resour. Res.*, 40, W01106, doi:10.1029/2003WR002540.
- Moradkhani, H., S. Sorooshian, H. V. Gupta, and P. Houser (2005a), Dual state–parameter estimation of hydrologic models using ensemble Kalman filter, *Adv. Water Resour.*, 28, 135–147.
- Moradkhani, H., K.-L. Hsu, H. V. Gupta, and S. Sorooshian (2005b), Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, *Water Resour. Res.*, 41, W05012, doi:10.1029/2004WR003604.
- Morgan, M. G., M. Henrion, and M. Small (1990), *Uncertainty: A Guide to Dealing With Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge Univ. Press, New York.
- Musso, C., N. Oudjane, and F. LeGland (2001), Improving regularized particle filters, in *Sequential Monte Carlo Methods in Practice*, edited by A. Doucet et al., pp. 247–271, Springer, New York.
- National Research Council (2001), *Conceptual Models of Flow and Transport in the Vadose Zone*, 392 pp., Natl. Acad. Press, Washington, D. C.
- Neuman, S. P. (2002), Accounting for conceptual model uncertainty via maximum likelihood model averaging, paper presented at the 4th International Conference on Calibration and Reliability in Groundwater Mod-

- elling (ModelCARE 2002), Charles Univ., Prague, Czech Republic, 16 – 20 June.
- Neuman, S. P. (2003), Maximum likelihood Bayesian averaging of uncertain model predictions, *Stochastic Environ. Res. Risk Assess.*, 17, 291–305.
- Palmer, T. N. (2004), Development of a European multi-model ensemble system for seasonal-to-interannual prediction (DEMETER), *Bull. Am. Meteorol. Soc.*, 85(6), 853–872.
- Palmer, T. N., C. Brankovic, and D. S. Richardson (2000), A probability and decision-model analysis of PROVOST seasonal multi-model ensemble integrations, *Q. J. R. Meteorol. Soc.*, 126(567), 2013–2033.
- Reichle, R., D. Entekhabi, and D. McLaughlin (2001a), Downscaling of radio brightness measurements for soil moisture estimation: A four-dimensional variational data assimilation approach, *Water Resour. Res.*, 37, 2353–2364.
- Reichle, R., D. B. McLaughlin, and D. Entekhabi (2001b), Variational data assimilation of microwave radiobrightness observations for land surface hydrology applications, *IEEE Trans. Geosci. Remote Sens.*, 39, 1708–1718.
- Reichle, R., D. McLaughlin, and D. Entekhabi (2002a), Hydrologic data assimilation with the ensemble Kalman filter, *Mon. Weather Rev.*, 130(1), 103–114.
- Reichle, R., J. P. Walker, R. D. Koster, and P. R. Houser (2002b), Extended versus ensemble Kalman filtering for land data assimilation, *J. Hydrometeorol.*, 3, 728–740.
- Restrepo, P. J. (1985), A view of maximum-likelihood estimation with large conceptual hydrologic models, *Appl. Math. Comput.*, 17(4), 375–403.
- Romanowicz, R. J., P. C. Young, and K. J. Beven (2006), Data assimilation and adaptive forecasting of water levels in the River Severn catchment, *Water Resour. Res.*, 42, W06407, doi:10.1029/2005WR004373.
- Seo, D. J., V. Koren, and N. Cajina (2003), Real-time variational assimilation of hydrologic and hydrometeorological data into operational hydrologic forecasting, *J. Hydrometeorol.*, 4(3), 627–641.
- Tamea, S., F. Laio, and L. Ridolfi (2005), Probabilistic nonlinear prediction of river flows, *Water Resour. Res.*, 41, W09421, doi:10.1029/2005WR004136.
- Thiemann, T., M. Trosset, H. Gupta, and S. Sorooshian (2001), Bayesian recursive parameter estimation for hydrologic models, *Water Resour. Res.*, 37(10), 2521–2535.
- Todini, E. (1978a), Mutually interactive state/parameter estimation (MISP) in hydrological applications, in *Identification and Control in Environmental Systems*, edited by G. C. Vansteenkiste, Elsevier, New York.
- Todini, E. (1978b), Mutually interactive state/parameter estimation (MISP), in *Application of Kalman Filter to Hydrology, Hydraulics and Water Resources*, edited by C.-L. Chiu, Univ. of Pittsburgh, Pittsburgh, Pa.
- Van Geer, F. C., C. B. M. Te Stroet, and Z. Yangxiao (1991), Using Kalman filtering to improve and quantify the uncertainty of numerical ground-water simulations: 1. The role of system noise and its calibration, *Water Resour. Res.*, 27(8), 1987–1994.
- Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003a), A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, 39(8), 1201, doi:10.1029/2002WR001642.
- Vrugt, J. A., H. V. Gupta, L. A. Bastidas, W. Bouten, and S. Sorooshian (2003b), Effective and efficient algorithm for multiobjective optimization of hydrologic models, *Water Resour. Res.*, 39(8), 1214, doi:10.1029/2002WR001746.
- Vrugt, J. A., C. G. H. Diks, H. V. Gupta, W. Bouten, and J. M. Verstraten (2005), Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, *Water Resour. Res.*, 41, W01017, doi:10.1029/2004WR003059.
- Wagener, T., and H. V. Gupta (2005), Model identification for hydrological forecasting under uncertainty, *Stochastic Environ. Res. Risk Assess.*, 19, doi:10.1007/s00477-005-0006-5.
- Wagener, T., N. McIntyre, M. J. Lees, H. S. Wheater, and H. V. Gupta (2003), Towards reduced uncertainty in conceptual rainfall-runoff modeling: Dynamic identifiability analysis, *Hydrol. Processes*, 17, 455–476.
- Walker, J. P., and P. R. Houser (2001), A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations, *J. Geophys. Res.*, 106, 11,761–774.
- Walker, J. P., G. R. Willgoose, and J. D. Kalma (2001), One-dimensional soil moisture profile retrieval by assimilation of near-surface observations: A comparison of retrieval algorithms, *Adv. Water Resour.*, 24(6), 631–650.
- Wikle, C. K. (2003), Hierarchical Bayesian models for predicting the spread of ecological processes, *Ecology*, 84(6), 1382–1394.
- Wójcik, R., P. A. Troch, H. Stricker, P. Torfs, E. Wood, H. Su, and Z. Su (2006), Mixtures of Gaussians for uncertainty description in bivariate latent heat flux proxies, *J. Hydrometeorol.*, 7, 330–345.
- Yangxiao, Z., C. B. M. Te Stroet, and F. C. Van Geer (1991), Using Kalman filtering to improve and quantify the uncertainty of numerical ground-water simulations: 2. Application to monitoring network design, *Water Resour. Res.*, 27(8), 1995–2006.
- Yapo, P. O., H. V. Gupta, and S. Sorooshian (1998), Multi-objective global optimization for hydrological models, *J. Hydrol.*, 204, 83–97.
- Ye, M., S. P. Neuman, and P. D. Meyer (2004), Maximum likelihood Bayesian averaging of spatial variability models in unsaturated fractured tuff, *Water Resour. Res.*, 40, W05113, doi:10.1029/2003WR002557.
- Ye, M., S. P. Neuman, P. D. Meyer, and K. Pohlmann (2005), Sensitivity analysis and assessment of prior model probabilities in MLBMA with application to unsaturated fractured tuff, *Water Resour. Res.*, 41, W12429, doi:10.1029/2005WR004260.
- Young, P. C. (2002), Advances in real-time flood forecasting, *Philos. Trans. R. Soc. London, Ser. A*, 360, 1433–1450.
- Young, P. C. (2003), Top-down and data-based mechanistic modeling of rainfall-flow dynamics at the catchment scale, *Hydrol. Processes*, 17, 2195–2217.
- Ziehmann, C. (2000), Comparison of a single-model EPS with a multi-model ensemble consisting of a few operational models, *Tellus, Ser. A*, 52(3), 280–299.

H. Gupta and Y. Liu, Department of Hydrology and Water Resources, University of Arizona, 845 North Park Avenue, Marshall Building, Room 535, Tucson, AZ 85721-0158, USA. (yqliu@hwr.arizona.edu)

Other Key Reference Papers for Prof. Sorooshian's Lecture:

- [1] Duan, Q., S. SOROOSHIAN (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resources Research*, 28(4): 1015-1031
- Vrugt JA, Gupta HV, et al. (2003), A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resources Research*, 39(8)
- [2] Gupta, HV, S, Sorooshian, P, Yapo (1998), Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resources Research*, 34(4): 751-763
- [3] Yapo, P., H. Gupta, S. Sorooshian (1998), Multi-objective global optimization for hydrologic models, *JOURNAL OF HYDROLOGY*, 204(1-4): 83-97
- [4] Duan, Q., N. Ajami, X. Gao, et al. (2007), Title: Multi-model ensemble hydrologic prediction using Bayesian model averaging, *Advances in Water Resources*, 30(5): 1371-1386
- [5] Moradkhani, H., H. Gupta, et al. (2005), Dual state-parameter estimation of hydrological models using ensemble Kalman filter, *Advances in Water Resources*, 28(2): 135-147

An Adaptive Estimation of Forecast Error Covariance Parameters for Kalman Filtering Data Assimilation

Xiaogu ZHENG*

National Institute of Water and Atmospheric Research, Wellington, New Zealand

(Received 5 December 2007; revised 22 May 2008)

ABSTRACT

An adaptive estimation of forecast error covariance matrices is proposed for Kalman filtering data assimilation. A forecast error covariance matrix is initially estimated using an ensemble of perturbation forecasts. This initially estimated matrix is then adjusted with scale parameters that are adaptively estimated by minimizing $-2\log$ -likelihood of observed-minus-forecast residuals. The proposed approach could be applied to Kalman filtering data assimilation with imperfect models when the model error statistics are not known. A simple nonlinear model (Burgers' equation model) is used to demonstrate the efficacy of the proposed approach.

Key words: data assimilation, Kalman filter, ensemble prediction, estimation

Citation: Zheng, X. G., 2009: An adaptive estimation of forecast error covariance parameters for Kalman filtering data assimilation. *Adv. Atmos. Sci.*, **26**(1), 154–160, doi: 10.1007/s00376-009-0154-5.

1. Introduction

Kalman filtering (Cohn, 1997) is a popular sequential analysis scheme for data assimilation. It is well known that estimation of forecast error covariance matrices plays a key role in the performance of the Kalman filtering schemes (e.g. Miller et al., 1994).

An approach for estimating forecast error covariance matrices in the Kalman filtering assimilation schemes is to parameterize the error covariance matrices, and then to estimate the parameters by minimizing the $-2\log$ -likelihood of observed-minus-forecast residuals (e.g., Dee and da Silva, 1999; Ozaki et al., 2000). A major obstacle for this approach is that it is difficult to parameterize forecast error covariance matrices, especially when they are not stationary in time.

A more popular approach is to generate an ensemble of perturbation forecasts by perturbing an initial state, and then the forecast error covariance matrix is estimated as the sampling covariance matrix of the ensemble (for example, Bengtsson et al., 2003). While this approach does not require parameterization of the forecast error covariance matrix, the estimation may depend on magnitude and number of the perturbations. Therefore, there is no guarantee that the sampling covariance matrix is a good estimation of the

forecast error covariance matrix.

In this paper, we propose an approach for estimating forecast error covariance matrices by combining both approaches. A real-time forecast error covariance matrix is initially estimated using an ensemble of perturbation forecasts. This initially estimated matrix is then adjusted with scale parameters that are estimated by minimizing the $-2\log$ -likelihood of observed-minus-forecast residuals. The proposed scheme permits imperfect models, but knowledge of the model error covariance matrices is not required.

The paper is arranged as follows. In section 2, the details of the proposed approach are described. Section 3 is devoted to the tests of the proposed approach using the data sets simulated by Burgers' equation model. Our conclusions are given in section 4.

2. Methodology

In this section, the proposed Kalman filtering data assimilation is outlined and a method for the adaptive estimation of the real time forecast error covariance matrices is proposed.

2.1 Proposed Kalman filtering

Using the notation similar to that proposed by Ide et al. (1997), a nonlinear discrete time prediction-

*Corresponding author: Xiaogu ZHENG, x.zheng@niwa.co.nz

observation system is of the form

$$\mathbf{x}_{i+1,t} = M_i[\mathbf{x}_{i,t}] + \boldsymbol{\eta}_i, \quad (1)$$

$$\mathbf{y}_{i,o} = \mathbf{H}_i \mathbf{x}_{i,t} + \boldsymbol{\varepsilon}_i, \quad (2)$$

where i is the time step index, the $\mathbf{x}_{i,t}$ is the true state vector with dimension n at the time step i with t abbreviating for “true”; M_i is an prediction operator such as a numerical weather forecasting model; $\mathbf{y}_{i,o}$ is the observation vector with dimension p_i ; \mathbf{H}_i is a $p_i \times n$ matrix indicating which linear combination of states is observed; $\boldsymbol{\eta}_i$ and $\boldsymbol{\varepsilon}_i$ are the model error and the observation error vectors, which are assumed to be statistically independent of both each other and time and have zero mean vectors and covariance matrices \mathbf{Q}_i and \mathbf{R}_i respectively. The goal of Kalman filtering data assimilation is to find a series of analysis states $\{\mathbf{x}_{i,a}\}$ that is sufficiently close to the true states $\{\mathbf{x}_{i,t}\}$ by using the information provided by the operators $\{M_i\}$ and the observations $\{\mathbf{y}_{i,o}\}$.

Suppose an initial analysis state $\mathbf{x}_{0,a}$ is known, our proposed Kalman filtering data assimilation comprises the following steps. Except for Step (ii) below for the estimation of the real time forecast error covariance matrices, the steps are those of the standard Kalman filter.

Step (i). Forecast the model state at time i :

$$\mathbf{x}_{i,f} = M_{i-1}[\mathbf{x}_{i-1,a}], \quad (3)$$

where $\mathbf{x}_{i,f}$ is assumed to be a Gaussian random vector with mean vector $\mathbf{x}_{i,t}$ and covariance matrix $\mathbf{P}_{i,f}$.

Step (ii). Estimate the forecast error statistics:

The forecast error covariance matrix could be estimated as the sampling covariance matrix which is defined by

$$\mathbf{P}_i \equiv \sum_{j=1}^m w_j \left(\mathbf{j} \mathbf{x}_{i,f} - \sum_{k=1}^m k \mathbf{x}_{i,f} w_k \right) \times \left(\mathbf{j} \mathbf{x}_{i,f} - \sum_{k=1}^m k \mathbf{x}_{i,f} w_k \right)^T \quad (4)$$

where $\{\mathbf{j} \mathbf{x}_{i,f}, j = 1, \dots, m\}$ are the perturbation forecast states from the perturbed analysis states $\{\mathbf{j} \mathbf{x}_{i-1,a}, j = 1, \dots, m\}$, and $\{w_j, j = 1, \dots, m\}$ are the weights. All of the existing approaches for deriving perturbed analysis states are able to be applied here, including the approach by perturbing observations used in the ensemble Kalman filtering assimilation (Anderson, 2001).

For the standard ensemble Kalman filtering assimilation, w_j is chosen as $1/m$, and \mathbf{P}_i is regarded as an estimation of $\mathbf{P}_{i,f}$ (Bengtsson et al., 2003). However,

if $\{\mathbf{j} \mathbf{x}_{i,f}, j = 1, \dots, m\}$ are not samples of population of the forecast state $\mathbf{x}_{i,f}$, and/or m is not sufficiently large, \mathbf{P}_i can be far from the forecast error covariance matrix. To mitigate this shortfall, the estimated forecast error covariance matrix is rescaled to

$$\hat{\mathbf{P}}_{i,f} = [\boldsymbol{\lambda}_i] \mathbf{P}_i [\boldsymbol{\lambda}_i], \quad (5)$$

where $\boldsymbol{\lambda}_i$ is a vector representing a scale change of \mathbf{P}_i and $[\boldsymbol{\lambda}_i]$ is a diagonal matrix with diagonal vector $\boldsymbol{\lambda}_i$. By choosing an appropriate $\boldsymbol{\lambda}_i$, $\hat{\mathbf{P}}_{i,f}$ could be a better approximation of $\mathbf{P}_{i,f}$ than \mathbf{P}_i could be. The detailed estimation procedure for $\boldsymbol{\lambda}_i$ is proposed in this paper, and is documented in section 2.2. Intuitively, $[\boldsymbol{\lambda}_i]$ can be viewed as a multivariate covariance inflation operator. While the common covariance inflation factor is a scalar value and is estimated by trials (Constantinescu et al., 2007), this study extends it to the diagonal version, and also optimizes it.

Step (iii). Calculate the observed-minus-forecast residuals:

$$\mathbf{d}_i = \mathbf{y}_{i,o} - \mathbf{H}_i \mathbf{x}_{i,f}, \quad (6)$$

where \mathbf{d}_i is assumed to be Gaussian with a zero mean vector and covariance matrix $\mathbf{H}_i [\boldsymbol{\lambda}_i] \mathbf{P}_i [\boldsymbol{\lambda}_i] \mathbf{H}_i^T + \mathbf{R}_i$.

Step (iv). Calculate the analysis state:

$$\mathbf{x}_{i,a} = \mathbf{x}_{i,f} + \hat{\mathbf{P}}_{i,f} \mathbf{H}_i^T (\mathbf{H}_i \hat{\mathbf{P}}_{i,f} \mathbf{H}_i^T + \mathbf{R}_i)^{-1} \mathbf{d}_i \quad (7)$$

Step (v). If $\mathbf{y}_{i,o}$ is not the last observation, put $i = i + 1$ and return to step (i). Otherwise, stop the filtering. $\mathbf{x}_{i,a}$ are the filtered states. Each $\mathbf{x}_{i,a}$ is Gaussian with mean $\mathbf{x}_{i,t}$ and covariance matrices

$$\mathbf{P}_{i,a} = \hat{\mathbf{P}}_{i,f} - \hat{\mathbf{P}}_{i,f} \mathbf{H}_i^T (\mathbf{H}_i \hat{\mathbf{P}}_{i,f} \mathbf{H}_i^T + \mathbf{R}_i)^{-1} \mathbf{H}_i \hat{\mathbf{P}}_{i,f}. \quad (8)$$

From Eq. (7), $\mathbf{H}_i \hat{\mathbf{P}}_{i,f}$ and $\mathbf{H}_i \hat{\mathbf{P}}_{i,f} \mathbf{H}_i^T$ are sufficient to determine $\mathbf{x}_{i,a}$. They can be estimated by

$$\mathbf{H}_i \hat{\mathbf{P}}_{i,f} = \sum_{j=1}^m w_j \left(\mathbf{H}_i [\boldsymbol{\lambda}_i] (\mathbf{j} \mathbf{x}_{i,f} - \sum_{k=1}^m k \mathbf{x}_{i,f} w_k) \right) \times \left([\boldsymbol{\lambda}_i] (\mathbf{j} \mathbf{x}_{i,f} - \sum_{k=1}^m k \mathbf{x}_{i,f} w_k) \right)^T \quad (9)$$

and

$$\mathbf{H}_i \hat{\mathbf{P}}_{i,f} \mathbf{H}_i^T = \sum_{j=1}^m w_j \left(\mathbf{H}_i [\boldsymbol{\lambda}_i] (\mathbf{j} \mathbf{x}_{i,f} - \sum_{k=1}^m k \mathbf{x}_{i,f} w_k) \right) \times \left(\mathbf{H}_i [\boldsymbol{\lambda}_i] (\mathbf{j} \mathbf{x}_{i,f} - \sum_{k=1}^m k \mathbf{x}_{i,f} w_k) \right)^T. \quad (10)$$

Usually, the observational dimension p_i is significantly less than the model dimension n . Then the computational cost for $\mathbf{H}_i \hat{\mathbf{P}}_{i,f}$ and $\mathbf{H}_i \hat{\mathbf{P}}_{i,f} \mathbf{H}_i^T$ is more economical than that for $\hat{\mathbf{P}}_{i,f}$.

2.2 Estimation of scale parameter λ_i

We estimate λ_i by minimizing the $-2\log$ -likelihood of the observed-minus-forecast residual $\mathbf{d}_i = \mathbf{y}_{i,o} - \mathbf{H}_i \mathbf{x}_{i,f}$. Since \mathbf{d}_i is assumed to be Gaussian with a zero mean vector and covariance matrix $\mathbf{H}_i[\lambda_i] \mathbf{P}_i[\lambda_i] \mathbf{H}_i^T + \mathbf{R}_i$ [Eq. (2)], its $-2\log$ -likelihood function is

$$-2L_i(\lambda_i) = \ln[\det(\mathbf{H}_i[\lambda_i] \mathbf{P}_i[\lambda_i] \mathbf{H}_i^T + \mathbf{R}_i)] + \mathbf{d}_i^T (\mathbf{H}_i[\lambda_i] \mathbf{P}_i[\lambda_i] \mathbf{H}_i^T + \mathbf{R}_i)^{-1} \mathbf{d}_i, \quad (11)$$

where \det represents the determinant of a matrix (Dee and da Silva, 1999; Ozaki et al., 2000).

Vector λ_i is comprised of the two components: the observable component $\lambda_{i,o}$ (i.e., $\mathbf{H}_i[\lambda_i]$ depends on $\lambda_{i,o}$), and an unobservable component $\lambda_{i,u}$ (i.e., $\mathbf{H}_i[\lambda_i]$ is independent of $\lambda_{i,u}$). From Eq. (11), the $-2\log$ -likelihood function is the function of $\mathbf{H}_i[\lambda_i]$. Then it is only the function of the observable component $\lambda_{i,o}$. Therefore, the unobservable component $\lambda_{i,u}$ cannot be estimated by minimizing function (11). These two components should be estimated separately.

2.2.1 Estimate observable component $\lambda_{i,o}$

The observable component $\lambda_{i,o}$ can be estimated by minimizing $-2L_i(\lambda_i)$ ([see Eq. (11)]. To do this effectively, we need to calculate its first derivative $\nabla_{\lambda} L_i(\lambda)$ and its second derivative $\nabla_{\lambda}^2 L_i(\lambda)$. Then, the fastest descendent direction of the likelihood function is

$$\delta \lambda = -\nabla_{\lambda} L_i(\lambda) [\nabla_{\lambda}^2 L_i(\lambda)]^{-1}. \quad (12)$$

For this purpose, we introduce the following matrix notation. For a matrix \mathbf{A} , the i -th column is denoted as \mathbf{a}_i and its (i, j) -th entry is denoted as a_{ij} . Suppose matrices \mathbf{A} and \mathbf{B} have a same dimension. $\mathbf{A} \times \mathbf{B}$ represents the matrix with (i, j) -th entry $a_{ij} b_{ij}$ (i.e., element-by-element or Shur product).

Under these notations, the first derivative of $L_i(\lambda)$ at time step i is

$$\begin{aligned} \nabla_{\lambda} [-L_i(\lambda)] &= \begin{pmatrix} (\partial/\partial \lambda_1)(-L_i(\lambda)) \\ (\partial/\partial \lambda_2)(-L_i(\lambda)) \\ \vdots \\ (\partial/\partial \lambda_n)(-L_i(\lambda)) \end{pmatrix} \\ &= - \begin{pmatrix} \mathbf{p}_1^T[\lambda] \mathbf{q}_1 \\ \mathbf{p}_2^T[\lambda] \mathbf{q}_2 \\ \vdots \\ \mathbf{p}_n^T[\lambda] \mathbf{q}_n \end{pmatrix} + \mathbf{c} \times \begin{pmatrix} \mathbf{p}_1^T[\lambda] \mathbf{c} \\ \mathbf{p}_2^T[\lambda] \mathbf{c} \\ \vdots \\ \mathbf{p}_n^T[\lambda] \mathbf{c} \end{pmatrix}, \end{aligned} \quad (13)$$

where $\mathbf{P} \equiv \mathbf{P}_i$, $\mathbf{Q} \equiv \mathbf{H}^T (\mathbf{H}[\lambda] \mathbf{P}[\lambda] \mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H}$, and $\mathbf{c} \equiv \mathbf{H}^T (\mathbf{H}[\lambda] \mathbf{P}[\lambda] \mathbf{H}^T + \mathbf{R})^{-1} \mathbf{d}$ with $\mathbf{H} = \mathbf{H}_i$, $\mathbf{d} = \mathbf{d}_i$ and $\mathbf{R} = \mathbf{R}_i$. The detailed proof is documented in the Appendix.

The second derivative of $L_i(\lambda)$ [the Hessian matrix $\nabla_{\lambda}^2 L_i(\lambda)$] is

$$\begin{aligned} &\frac{\partial}{\partial \lambda} \begin{pmatrix} \mathbf{p}_1^T[\lambda] \mathbf{q}_1 & \mathbf{p}_2^T[\lambda] \mathbf{q}_2 & \cdots & \mathbf{p}_n^T[\lambda] \mathbf{q}_n \end{pmatrix} \\ &- \frac{\partial}{\partial \lambda} \begin{pmatrix} \mathbf{c}^T \times (\mathbf{p}_1^T[\lambda] \mathbf{c} & \mathbf{p}_2^T[\lambda] \mathbf{c} & \cdots & \mathbf{p}_n^T[\lambda] \mathbf{c}) \end{pmatrix} \\ &= \mathbf{Q} \times \mathbf{P} - (\mathbf{P}[\lambda] \mathbf{Q}) \times (\mathbf{Q}[\lambda] \mathbf{P}) - (\mathbf{P}[\lambda] \mathbf{Q})[\lambda] \mathbf{P} \times \mathbf{Q} \\ &+ \mathbf{Q} \times (\mathbf{P}[\lambda] \mathbf{c} \mathbf{c}^T [\lambda] \mathbf{P}) + (\mathbf{Q}[\lambda] \mathbf{P}) \times (\mathbf{P}[\lambda] \mathbf{c} \mathbf{c}^T) - \\ &\mathbf{P} \times (\mathbf{c} \mathbf{c}^T) + (\mathbf{P}[\lambda] \mathbf{Q}) \times (\mathbf{c} \mathbf{c}^T [\lambda] \mathbf{P}) + \\ &(\mathbf{P}[\lambda] \mathbf{Q}[\lambda] \mathbf{P}) \times (\mathbf{c} \mathbf{c}^T) \end{aligned} \quad (14)$$

The detailed proof is also documented in the Appendix. Note that both derivatives also only depend on $\lambda_{i,o}$.

After the fastest descendent direction is obtained, λ_i is substituted by $\lambda + \mu \delta \lambda_i$, where $0 < \mu \leq 1$ is a scalar to guarantee that $-2L_i(\lambda_i + \mu \delta \lambda_i)$ is less than $-2L_i(\lambda)$.

2.2.2 Estimate unobservable component $\lambda_{i,u}$

Although the likelihood at the time step i is independent of the unobservable component $\lambda_{i,u}$, the $-2\log$ -likelihood at time step $i+1$ ($-2L_{i+1}$) does depend on $\lambda_{i,u}$. This is because L_{i+1} depends on $\mathbf{x}_{i,a}$, the analysis state at time step i . From Eq. (7),

$$\mathbf{x}_{i,a} = \mathbf{x}_{i,f} + [\lambda_i] \mathbf{P}_i[\lambda_i] \mathbf{H}_i^T (\mathbf{H}_i[\lambda_i] \mathbf{P}_i[\lambda_i] \mathbf{H}_i^T + \mathbf{R}_i)^{-1} \mathbf{d}_i, \quad (15)$$

which depends on both $\lambda_{i,o}$ and $\lambda_{i,u}$.

We shall estimate $\lambda_{i,u}$ and $\lambda_{i+1,o}$ jointly. For a $\lambda_{i+1,o}$, find a new $\lambda_{i,u}$ to reduce $-2L_{i+1}$ by direct search. Then, apply the procedure documented in section 2.2.1 to find a new $\lambda_{i+1,o}$ that further reduces $-2L_{i+1}$. Continue this procedure iteratively until L_{i+1} converges.

2.2.3 Constraint on parameter λ

The dimension of λ is equivalent to the dimension of analysis states (n), that is often too large to estimate λ . In practice, constraints are often imposed on λ . As an example, we may assume components of λ at one vertical level are all the same in meteorological data assimilation.

The first derivative and the second derivative of the constrained $L(\lambda)$ can be easily derived from the first derivative and second derivative of

the unconstrained $L(\lambda)$. Suppose the analysis space is partitioned into $N(N \leq n)$ blocks, i.e., $\lambda_N = (\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_2, \dots, \lambda_N, \dots, \lambda_N)$. Then the first derivative of constrained $L(\lambda_N)$ is the N dimensional vector which element for a block is the sum of elements of $\nabla_\lambda L(\lambda_N)$ within that block. Similarly, the second derivative of constrained $L(\lambda_N)$ is the N by N matrix which element for a block pair is the sums of element of $\nabla_\lambda^2 L(\lambda_N)$ within the block pairs.

Because the constraint has always to be applied for large models, the dimension of the second derivatives $\nabla_\lambda^2 L(\lambda)$ will not be large. Therefore, the inverse of $\nabla_\lambda^2 L(\lambda)$ in Eq. (12) can be calculated directly.

2.3 Computing $(H[\lambda]P[\lambda]H^T + R)^{-1}$

$(H[\lambda]P[\lambda]H^T + R)^{-1}$ can be calculated as follows. Decompose $H[\lambda]P[\lambda]H^T$ into

$$H[\lambda]P[\lambda]H^T = \sum_{i=1}^m z_i z_i^T, \quad (16)$$

where

$$z_j = H[\lambda] \sqrt{w_j} \left(j x_{i,f} - \sum_{k=1}^m k x_{i,f} w_k \right). \quad (17)$$

Define $A_i = \sum_{k=1}^i z_k z_k^T + R$. It is easy to check that

$$A_1^{-1} = R^{-1} - (R^{-1} z_1 z_1^T R^{-1}) / (1 + z_1^T R^{-1} z_1), \quad (18)$$

and for $1 \leq j \leq m-1$

$$A_{j+1}^{-1} = A_j^{-1} - (A_j^{-1} z_j z_j^T A_j^{-1}) / (1 + z_j^T A_j^{-1} z_j). \quad (19)$$

Thus, $(H[\lambda]P[\lambda]H^T + R)^{-1} = A_m^{-1}$ can be calculated iteratively.

In this way, the inverse $(H[\lambda]P[\lambda]H^T + R)^{-1}$ can be calculated without significantly computational cost, providing the inverse of the observational matrix R^{-1} is known. However, R^{-1} is also required for any variational approach. Most ensemble based Kalman filtering are likely to be as computationally expensive as 4D-VAR, and perhaps significantly more expensive when there are an overwhelmingly large number of observations, such as very high resolution satellite images (Hamill, 2006). Our proposed approach for calculating the inverse provides a solution to overcome this obstacle.

3. Application to Burgers' equation model

3.1 Burgers' equation model

Burgers (1974) proposed the following equation

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = v \frac{\partial^2 u}{\partial x^2} \quad (20)$$

to describe the one-dimensional advection-diffusion process over an infinite spatial domain. Recently, Zhu and Kamachi (2000) used it as a test bed for several data assimilation schemes. In the present study, we also use it to test our proposed methodology. For the readers' convenience, we briefly introduce how to simulate the solution of a Burgers equation following Zhu and Kamachi (2000). This solution is then used as the series of true states in the present study. Knowing the true states, the root mean square error (RMSE) of the assimilated analysis states can be estimated.

The numerical model is defined as a finite-difference leapfrog scheme with a forward step every 15 time steps. This simple scheme gives accurate results by comparing with one analytical solution (Uboldi and Kamachi, 2000). We will use the following abbreviations: m for meter, km for kilometer, s for second and h for hour. The computation spatial domain is $[-1000, 1000]$ km, but only the solution within $[-100, 150]$ km is considered. A large computation domain is used to reduce the boundary effects on the interior solution. The model parameters are: time step $\Delta t = 60$ s; spatial grid resolution $\Delta x = 5$ km. Therefore, n (the dimension of state vector) is $51 (= 150/5 + 1 + 100/5)$. The simulation time is $T = 16$ h. The "true" solution is generated by running the model with the initial condition

$$u(x, 0) = \begin{cases} 0 & x < -L; \quad x > L \\ u_0(1 + 2x/L)(1 - 2x/L)^2 & 0 \leq x < L \\ u_0(1 - 2x/L)(1 + 2x/L)^2 & -L < x < 0 \end{cases} \quad (21)$$

where $u_0 = 5 \text{ m s}^{-1}$, and $L = 50$ km and with diffusion coefficient $v = 1.0 \times 10^4 \text{ m}^2 \text{ s}^{-1}$. Denote the true solution by $u_t(x, t)$, where $-100 \text{ km} \leq x \leq 150 \text{ km}$ and $0 \text{ h} \leq t \leq 16 \text{ h}$.

In the data assimilation experiments with the imperfect model, the diffusion coefficient is set to be $v = 1.4 \times 10^4 \text{ m}^2 \text{ s}^{-1}$ to simulate the model error. The model error caused by the wrong diffusion coefficient is estimated by running the model twice with the same above initial condition, but with different diffusion coefficients. The spatially averaged root mean square of the model error is shown in Fig. 2 of Zhu and Kamachi (2000) with an overall averaged error of about 0.13 m s^{-1} . The overall averaged signal is about 0.9 m s^{-1} in the time-space domain. The averaged model error is about 14% of the average signal. The observations are assumed to be available at every other model grid point from $x = -100$ km to 150 km. In total there are 26 observation stations (total of 51 model grid points). The observations are available only at hour 1, 4, 7 and

10, and are statistically independent of each other with standard deviation 0.15 m s^{-1} .

3.2 Results

To construct \mathbf{P}_i in Eq. (4), the number of perturbations m and the weights $\{w_j, j = 1, \dots, m\}$ are selected similar to that in the unscented ensemble Kalman filter (Julier and Uhlmann, 2004), where $m = 2n + 1$ and

$$w_j = \begin{cases} k/(n+k) & j = 0, \\ 1/2(n+k) & j = 2, \dots, n+1, \end{cases} \quad (22)$$

where k is a parameter. Julier and Uhlmann (2004) suggested that if the forecast error is Gaussian, set $n+k=3$. Otherwise set $n+k < 3$. Since the forecast error is assumed Gaussian and $n=51$, we set $k=-48$. The perturbed analysis states is selected as

$${}_j\mathbf{x}_{i,a} = \begin{cases} \mathbf{x}_{i,a} & j = 1, \\ \mathbf{x}_{i,a} + \alpha \boldsymbol{\delta}_j & j = 2, \dots, n+1, \\ \mathbf{x}_{i,a} - \alpha \boldsymbol{\delta}_j & j = n+2, \dots, 2n+1, \end{cases} \quad (23)$$

where $\boldsymbol{\delta}_j$ is j -th column of the n by n identity matrix and $\alpha > 0$ is the perturbation parameter. Here α is set as 0.1 m s^{-1} to match the overall averaged initial error of 0.13 m s^{-1} (Zhu and Kamachi, 2000). Scale parameters for the 26 observational points $\boldsymbol{\lambda}_{i,o}$ and the scale parameters for the other 25 unobservable points $\boldsymbol{\lambda}_{i,u}$ are constraint as one parameter respectively.

The estimated scale parameters, the $-2\log$ -likelihood, and the RMSE of the analysis state at each time step with observation are listed in Table 1. The RMSE of the analysis state from hour 1 to hour 16 are plotted in Fig. 1. As a comparison, the $-2\log$ -likelihood, and the RMSE of analysis without the parameter adjustment (i.e., $\boldsymbol{\lambda} = \mathbf{I}$) are listed in Table 2, and the corresponding RMSE are also plotted in Fig. 1.

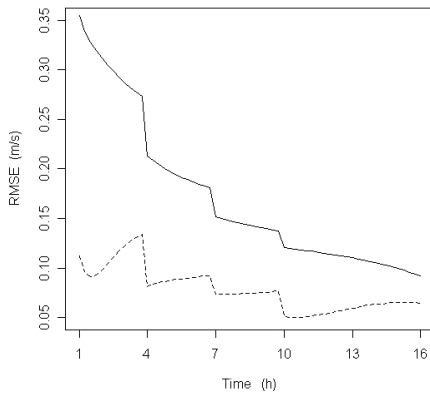


Fig. 1. Root mean square errors of the analysis states with the parameter adjustment (dashed) and without the parameter adjustment (solid).

Table 1. Results with the parameter adjustment.

	Hour 1	Hour 4	Hour 7	Hour 10
RMSR (m s^{-1})	0.12	0.097	0.068	0.04
$-2\log [L_i(\boldsymbol{\lambda}_i)]$	-389	-421	-429	-436
Observable $\lambda_{i,o}$	10	2	4	0.5
Unobservable $\lambda_{i,o}$	11	1	4	*

Note: Unobserved $\lambda_{i,o}$ is not required at the last time step.

Table 2. Results without the parameter adjustment.

	Hour 1	Hour 4	Hour 7	Hour 10
RMSR (m s^{-1})	0.35	0.19	0.14	0.09
$-2\log [L_i(\mathbf{I})]$	-251	-387	-411	-421

Tables 1 and 2 show that the $-2\log$ -likelihood of the observed-minus-forecast residual with the parameter adjustment is significantly less than those without parameter adjustment. Correspondingly, the RMSE for the analysis states with the parameter adjustment are less than the half of those without the parameter adjustment. This fact is also shown in Fig. 1.

Zhu and Kamachi (2000) proposed a number of adaptive variational assimilation schemes with imperfect models. They tested their schemes against the simplified 4D-VAR assimilation scheme and the space variable optimal nudging assimilation scheme, using Burgers' equation model as the test bed. The RSME derived by these schemes are shown in Fig. 3a of Zhu and Kamachi (2000). For these schemes, the RMSE for the reduced order adaptive variational method (ROAV) is the best one. Comparing the RMSE derived by ROAV with the RMSE shown in Fig. 1 and in Table 1, our RMSE is comparable to theirs at early hours (1 and 4), but ours is 0.03 m s^{-1} less (at hour 7) and 0.01 m s^{-1} less (at hour 10) than theirs. Furthermore there are only seven parameters (Table 1) to be estimated in our proposed schemes, while there are 51 parameters in ROAV. These facts indicate that the adaptive estimation of forecast error statistic proposed in this paper is at least a competitive scheme.

4. Conclusions

An adaptive estimation of forecast error statistics is proposed for Kalman filtering data assimilation for non-linear imperfect models. It has the advantage that the model error statistics do not need to be known and the forecast error statistics and observational error statistics can depend on time steps. The proposed scheme may have potential in data assimilation with large models. It is showed, by a case study, that the proposed inflation can improve the assimilation. In the future, we plan to further study the possibility of applying the proposed inflation approach to improve the

assimilation when the number of ensemble members is small.

Acknowledgements. This study was initiated during my visit to the Institute of Statistical Mathematics (ISM) in Tokyo under the support of a Visiting Fellowship offered by the Japanese Ministry of Education, Science, Sports and Culture. The study has been continued under the support of the Foundation for Research Science and Technology of New Zealand under contract C01X0401. I thank Prof. K. Tanabe for his collaboration on this paper and Prof. T. Ozaki for his support and many stimulating discussions. We would like to thank Prof. Zhidong Bai and Dr. Shouzheng Tang for their comments on mathematical algorithms, Prof. Masahide Kimoto and Dr. Phil Andrews for their comments on physical sciences and the two anonymous reviewers for improving the presentation.

APPENDIX

Proof of the Derivatives of the Log-Likelihood

1. Proof of the first derivative of the log-likelihood

Let subscripts r and s represent components of the model state vector.

$$\begin{aligned}
 & \frac{\partial}{\partial \lambda_r} \{ \ln[\det(\mathbf{H}[\lambda]\mathbf{P})[\lambda]\mathbf{H}^T + \mathbf{R})] \} \\
 &= \text{tr} \left[(\mathbf{H}[\lambda]\mathbf{P}[\lambda]\mathbf{H}^T + \mathbf{R})^{-1} \frac{\partial}{\partial \lambda_r} (\mathbf{H}[\lambda]\mathbf{P}[\lambda]\mathbf{H}^T + \mathbf{R}) \right] \\
 &= \text{tr} \left[(\mathbf{H}[\lambda]\mathbf{P}[\lambda]\mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H}([\delta_r]\mathbf{P}[\lambda] + [\lambda]\mathbf{P}[\delta_r])\mathbf{H}^T \right] \\
 &= \text{tr} \left[(\mathbf{H}[\lambda]\mathbf{P}[\lambda]\mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H}[\delta_r]\mathbf{P}[\lambda]\mathbf{H}^T \right] \\
 &+ \text{tr} \left[(\mathbf{H}[\lambda]\mathbf{P}[\lambda]\mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H}[\lambda]\mathbf{P}[\delta_r]\mathbf{H}^T \right] \\
 &= 2\text{tr} \left[(\mathbf{P}[\lambda][\delta_r](\mathbf{H}^T(\mathbf{H}[\lambda]\mathbf{P}[\lambda]\mathbf{H}^T + \mathbf{R})^{-1}\mathbf{H})) \right] \\
 &= 2\mathbf{p}_r^T[\lambda]\mathbf{q}_r. \tag{A1}
 \end{aligned}$$

$$\begin{aligned}
 & \frac{\partial}{\partial \lambda_r} [d^T(\mathbf{H}[\lambda]\mathbf{P}[\lambda]\mathbf{H}^T + \mathbf{R})^{-1}d] \\
 &= -d^T(\mathbf{H}[\lambda]\mathbf{P}[\lambda]\mathbf{H}^T + \mathbf{R})^{-1} \times \\
 & \frac{\partial}{\partial \lambda_r} (\mathbf{H}[\lambda]\mathbf{P}[\lambda]\mathbf{H}^T + \mathbf{R})(\mathbf{H}[\lambda]\mathbf{P}[\lambda]\mathbf{H}^T + \mathbf{R})^{-1}d \\
 &= -[d^T(\mathbf{H}[\lambda]\mathbf{P}[\lambda]\mathbf{H}^T + \mathbf{R})^{-1}\mathbf{H})([\delta_r]\mathbf{P}[\lambda] +
 \end{aligned}$$

$$\begin{aligned}
 & + [\lambda]\mathbf{P}[\delta_r])(\mathbf{H}^T(\mathbf{H}[\lambda]\mathbf{P}[\lambda]\mathbf{H}^T + \mathbf{R})^{-1}d) \\
 &= -\mathbf{c}^T[\delta_r]\mathbf{P}[\lambda]\mathbf{c} - \mathbf{c}^T[\lambda]\mathbf{P}[\delta_r]\mathbf{c} \\
 &= -c_r\mathbf{p}_r^T[\lambda]\mathbf{c} - \mathbf{c}^T[\lambda]\mathbf{p}_r c_r \\
 &= -2c_r\mathbf{p}_r^T[\lambda]\mathbf{c}. \tag{A2}
 \end{aligned}$$

From Eq. (11) and Eqs. (A1)–(A2)

$$\begin{aligned}
 \frac{\partial}{\partial \lambda_r} L_i(\lambda) &= -\frac{1}{2}(2\mathbf{p}_r^T[\lambda]\mathbf{q}_r - 2c_r\mathbf{p}_r^T[\lambda]\mathbf{c}) \\
 &= -\mathbf{p}_r^T[\lambda]\mathbf{q}_r + c_r\mathbf{p}_r^T[\lambda]\mathbf{c}.
 \end{aligned}$$

Thus, the first derivative [i.e., Eq. (13)] is derived.

2. Proof of the second derivative of the log-likelihood

$$\begin{aligned}
 \frac{\partial}{\partial \lambda_s} \mathbf{Q} &= \mathbf{H}^T \left\{ \frac{\partial}{\partial \lambda_s} (\mathbf{H}[\lambda]\mathbf{P}[\lambda]\mathbf{H}^T + \mathbf{R})^{-1} \right\} \mathbf{H} \\
 &= -\mathbf{H}^T \{ (\mathbf{H}[\lambda]\mathbf{P}[\lambda]\mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H}([\delta_s]\mathbf{P}[\lambda] + \\
 & [\lambda]\mathbf{P}[\delta_s])\mathbf{H}^T (\mathbf{H}([\lambda]\mathbf{P}[\lambda]\mathbf{H}^T + \mathbf{R})^{-1}\mathbf{H}) \} \mathbf{H} \\
 &= -\mathbf{Q}([\delta_s]\mathbf{P}[\lambda] + [\lambda]\mathbf{P}[\delta_s])\mathbf{Q}.
 \end{aligned}$$

Then

$$\begin{aligned}
 & \frac{\partial}{\partial \lambda_s} \mathbf{p}_r^T[\lambda]\mathbf{q}_r \\
 &= \mathbf{p}_r^T \left(\frac{\partial}{\partial \lambda_s} [\lambda] \right) \mathbf{q}_r + \mathbf{p}_r^T[\lambda] \frac{\partial}{\partial \lambda_s} \mathbf{q}_r \\
 &= \mathbf{p}_r^T[\delta_s]\mathbf{q}_r - \mathbf{p}_r^T[\lambda]\mathbf{Q}[\delta_s]\mathbf{P}[\lambda]\mathbf{q}_r - \\
 & \mathbf{p}_r^T[\lambda]\mathbf{Q}[\lambda]\mathbf{P}[\delta_s]\mathbf{q}_r \\
 &= p_{rs}q_{sr} - (\mathbf{p}_r^T[\lambda]\mathbf{q}_s)(\mathbf{p}_s^T[\lambda]\mathbf{q}_r) - \\
 & (\mathbf{p}_r^T[\lambda]\mathbf{Q}[\lambda]\mathbf{p}_s)q_{sr}.
 \end{aligned}$$

Thus,

$$\begin{aligned}
 & \frac{\partial}{\partial \lambda} \begin{pmatrix} \mathbf{p}_1^T[\lambda]\mathbf{q}_1 & \mathbf{p}_2^T[\lambda]\mathbf{q}_2 & \cdots & \mathbf{p}_n^T[\lambda]\mathbf{q}_n \end{pmatrix} \\
 &= \mathbf{Q} \times \mathbf{P} - (\mathbf{P}[\lambda]\mathbf{Q} \times (\mathbf{Q}[\lambda]\mathbf{P}) - \\
 & (\mathbf{P}[\lambda]\mathbf{Q}[\lambda]\mathbf{P}) \times \mathbf{Q} \tag{A3}
 \end{aligned}$$

Since

$$\begin{aligned}\frac{\partial}{\partial \lambda_s} \mathbf{c} &= \mathbf{H}^T \left\{ \frac{\partial}{\partial \lambda_s} (\mathbf{H}[\lambda] \mathbf{P}[\lambda] \mathbf{H}^T + \mathbf{R})^{-1} \right\} \mathbf{d} \\ &= -\mathbf{H}^T \{ (\mathbf{H}[\lambda] \mathbf{P}[\lambda] \mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H}([\delta_s] \mathbf{P}[\lambda] + \\ &\quad [\lambda] \mathbf{P}[\delta_s]) \mathbf{H}^T (\mathbf{H}[\lambda] \mathbf{P}[\lambda] \mathbf{H}^T + \mathbf{R})^{-1} \} \mathbf{d} \\ &= -\mathbf{Q}([\delta_s] \mathbf{P}[\lambda] + [\lambda] \mathbf{P}[\delta_s]) \mathbf{c},\end{aligned}$$

it follows that

$$\begin{aligned}\frac{\partial}{\partial \lambda_s} c_r \mathbf{p}_r^T [\lambda] \mathbf{c} &= \left(\frac{\partial}{\partial \lambda_s} c_r \right) \mathbf{p}_r^T [\lambda] \mathbf{c} + c_r \mathbf{p}_r^T \left(\frac{\partial}{\partial \lambda_s} [\lambda] \right) \mathbf{c} + \\ &\quad c_r \mathbf{p}_r^T [\lambda] \left(\frac{\partial}{\partial \lambda_s} \mathbf{c} \right) \\ &= -\mathbf{q}_r^T ([\delta_s] \mathbf{P}[\lambda] + [\lambda] \mathbf{P}[\delta_s]) \mathbf{c} \mathbf{p}_r^T [\lambda] \mathbf{c} + c_r \mathbf{p}_r^T [\delta_s] \mathbf{c} \\ &\quad - c_r \mathbf{p}_r^T [\lambda] \mathbf{Q}([\delta_s] \mathbf{P}[\lambda] + [\lambda] \mathbf{P}[\delta_s]) \mathbf{c} \\ &= -q_{rs} (\mathbf{P}[\lambda] \mathbf{c} \mathbf{c}^T [\lambda] \mathbf{P})_{sr} - (\mathbf{Q}[\lambda] \mathbf{P})_{rs} (\mathbf{P}[\lambda] \mathbf{c} \mathbf{c}^T)_{rs} + \\ &\quad \mathbf{p}_{rs} (\mathbf{c} \mathbf{c}^T)_{rs} - (\mathbf{P}[\lambda] \mathbf{Q})_{rs} (\mathbf{P}[\lambda] \mathbf{c} \mathbf{c}^T)_{sr} - \\ &\quad (\mathbf{P}[\lambda] \mathbf{Q}[\lambda] \mathbf{P})_{rs} (\mathbf{c} \mathbf{c}^T)_{sr}\end{aligned}$$

Therefore

$$\begin{aligned}\frac{\partial}{\partial \lambda} (\mathbf{c}^T \times (\mathbf{p}_1^T [\lambda] \mathbf{c} \quad \mathbf{p}_2^T [\lambda] \mathbf{c} \quad \cdots \quad \mathbf{p}_n^T [\lambda] \mathbf{c})) \\ &= -\mathbf{Q} \times (\mathbf{P}[\lambda] \mathbf{c} \mathbf{c}^T [\lambda] \mathbf{P}) - \\ &\quad (\mathbf{Q}[\lambda] \mathbf{P}) \times (\mathbf{P}[\lambda] \mathbf{c} \mathbf{c}^T) + \\ &\quad \mathbf{P} \times (\mathbf{c} \mathbf{c}^T) - (\mathbf{P}[\lambda] \mathbf{Q}) \times (\mathbf{c} \mathbf{c}^T [\lambda] \mathbf{P}) - \\ &\quad (\mathbf{P}[\lambda] \mathbf{Q}[\lambda] \mathbf{P}) \times (\mathbf{c} \mathbf{c}^T)\end{aligned}\tag{A4}$$

The Hessian matrix (Eq. 14) is derived from Eqs. (A3)–(A4).

REFERENCES

- Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2844–2903.
- Bengtsson, T., D. Nychka, and C. Snyder, 2003: A framework for data assimilation and forecasting in high dimensional non-linear dynamic systems. *J. Geophys. Res.*, **108**(D), 8875.
- Burgers, J. M., 1974: *The Nonlinear Diffusion*. D. Reidel Publ. Co., Dordrecht, Holland, 173pp.
- Cohn, S. E. 1997: An introduction to estimation theory. *J. Meteor. Soc. Japan*, **75**, 257–288.
- Constantinescu, M., A. Sandu, T. Chai, and G. R. Carmichael, 2007: Ensemble-based chemical data assimilation. I: General approach. *Quart. J. Roy. Meteor. Soc.*, **133**, 1229–1243.
- Dee, D. P., and A. M. da Silva, 1999: Maximum-likelihood estimation of forecast and observation error covariance parameters. Part 1: Methodology. *Mon. Wea. Rev.*, **127**, 1822–1849.
- Hamill, T. M., 2006: Ensemble-based atmospheric data assimilation. *Predictability of Weather and Climate*, Cambridge press, 123–156.
- Ide, K., P. Courtier, G. Michael, and A. C. Lorenc, 1997: Unified notation for data assimilation: operational, sequential and variational. *J. Meteor. Soc. Japan*, **75**, 71–79.
- Julier, S. J., and K. Uhlmann, 2004: Unscented filtering and nonlinear estimation. *Proc. IEEE Aerospace and Electronic Systems*, **92**, 410–422.
- Miller, R. N., M. Ghil, and F. Gauthiez, 1994: An advanced data assimilation in strongly nonlinear dynamical systems. *J. Atmos. Sci.*, **15**, 1037–1056.
- Ozaki, T., J. C. Jimenez, and V. H. Ozaki, 2000: The role of the likelihood function in the estimation of chaos models. *Journal of Time Series Analysis*, **21**, 363–387.
- Uboldi, F., and M. Kamachi, 2000: Time-space weak-constraint data assimilation for nonlinear models. *Tellus*, **52A**, 412–421.
- Zhu, J. and M. Kamachi, 2000: An adaptive variational method for data assimilation with imperfect models. *Tellus*, **52A**, 265–279.

Other Key Reference Papers for Prof. Zheng's Lecture:

- [1] Dee, D. P., A. Silva (1999), Maximum-likelihood estimation of forecast and observation error covariance parameters. Part I: Methodology, Monthly weather review, 127: 1822-1834
- [2] Dee, D. P., et al. (1999), Maximum-likelihood estimation of forecast and observation error covariance parameters. Part I: Application, Monthly weather review, 127: 1835-1849
- [3] Evensen, G. (1994), Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, Journal of Geophysical Research, 10143-10162

PART IV: APPLICATIONS OF DATA ASSIMILATION

This part introduces some novel and live application cases about data assimilation, covering the water and energy cycle, the crop monitoring and yield estimation. Some assimilation skills will be present. The multi-institution North American Land Data Assimilation System (NLDAS) is presented.

An EKF assimilation of AMSR-E soil moisture into the ISBA land surface scheme

C. S. Draper,¹ J.-F. Mahfouf,² and J. P. Walker¹

Received 21 December 2008; revised 3 March 2009; accepted 24 June 2009; published 20 October 2009.

[1] An Extended Kalman Filter (EKF) for the assimilation of remotely sensed near-surface soil moisture into the Interactions between Surface, Biosphere, and Atmosphere (ISBA) model is described. ISBA is the land surface scheme in Météo-France's Aire Limitée Adaptation Dynamique développement InterNational (ALADIN) Numerical Weather Prediction (NWP) model, and this work is directed toward providing initial conditions for NWP. The EKF is used to assimilate near-surface soil moisture observations retrieved from C-band Advanced Microwave Scanning Radiometer (AMSR-E) brightness temperatures into ISBA. The EKF can translate near-surface soil moisture observations into useful increments to the root-zone soil moisture. If the observation and model soil moisture errors are equal, the Kalman gain for the root-zone soil moisture is typically 20–30%, resulting in a mean net monthly increment for July 2006 of $0.025 \text{ m}^3 \text{ m}^{-3}$ over ALADIN's European domain. To test the benefit of evolving the background error, the EKF is compared to a Simplified EKF (SEKF), in which the background errors at the time of the analysis are constant. While the Kalman gains for the EKF and SEKF are derived from different model processes, they produce similar soil moisture analyses. Despite this similarity, the EKF is recommended for future work where the extra computational expense can be afforded. The method used to rescale the near-surface soil moisture data to the model climatology has a greater influence on the analysis than the error covariance evolution approach, highlighting the importance of developing appropriate methods for rescaling remotely sensed near-surface soil moisture data.

Citation: Draper, C. S., J.-F. Mahfouf, and J. P. Walker (2009), An EKF assimilation of AMSR-E soil moisture into the ISBA land surface scheme, *J. Geophys. Res.*, 114, D20104, doi:10.1029/2008JD011650.

1. Introduction

[2] Soil moisture can have a strong influence on Numerical Weather Prediction (NWP) forecasts, both at short [Baker *et al.*, 2001; Drusch and Viterbo, 2007] and medium range [Zhang and Frederiksen, 2003; Fischer *et al.*, 2007]. Currently, soil moisture is initialized in most operational NWP models based on errors in short-range forecasts of low-level humidity and temperature [e.g., Giard and Bazile, 2000; Hess, 2001; Bélair *et al.*, 2003]. While these schemes can in general produce reasonable boundary layer forecasts [Drusch and Viterbo, 2007], they assume a causative relationship between low-level atmospheric forecast errors and local soil moisture errors. As a result, soil moisture is often adjusted to compensate for errors elsewhere in the model, resulting in soil moisture fields that are frequently unrealistic [Seuffert *et al.*, 2004; Draper and Mills, 2008]. The accumulation of model errors in surface variables also makes it difficult to diagnose the source of these errors. Additionally,

these schemes cannot be sensibly applied to situations where the local soil moisture–atmospheric boundary layer feedback is weak; for example, during periods of strong advection, or weak radiative forcing. The effectiveness of a soil analysis based on screen-level variables is also limited by the availability of screen-level observations, which are particularly sparse across much of the Southern Hemisphere. A particularly promising approach to addressing some of the above mentioned shortcomings is the possibility of assimilating remotely sensed near-surface soil moisture into NWP models [e.g., Seuffert *et al.*, 2004; Balsamo *et al.*, 2007; Scipal *et al.*, 2008]. This approach is explored here, using an Extended Kalman Filter (EKF) to assimilate remotely sensed near-surface soil moisture into Météo-France's Aire Limitée Adaptation Dynamique développement InterNational (ALADIN) NWP model.

[3] Recent interest in the assimilation of remotely sensed near-surface soil moisture is anticipating the planned launch of the European Space Agency's Soil Moisture and Ocean Salinity (SMOS [Kerr *et al.*, 2001]) mission. SMOS is the first purpose designed soil moisture remote sensing mission, and will be followed by NASA's Soil Moisture Active Passive (SMAP [Entekhabi *et al.*, 2004]) mission. However, while SMOS and SMAP are expected to enhance the accuracy and utility of remotely sensed soil moisture data, currently orbiting microwave sensors can already provide

¹Department of Civil and Environmental Engineering, University of Melbourne, Parkville, Victoria, Australia.

²Météo France, CNRS, Toulouse, France.

useful soil moisture observations. For this study, near-surface soil moisture has been retrieved from passive microwave brightness temperatures observed by the Advanced Microwave Scanning Radiometer–Earth Observing System (AMSR-E). While it is difficult to quantitatively verify remotely sensed soil moisture due to the scarcity of soil moisture data at the appropriate scales [Reichle *et al.*, 2004], some encouraging comparisons have been made between soil moisture derived from AMSR-E and that from other sources. At the local scale, AMSR-E derived soil moisture has a good temporal association to in situ soil moisture data [Wagner *et al.*, 2007; Rüdiger *et al.*, 2009; Draper *et al.*, 2009], and to model data [Rüdiger *et al.*, 2009]. At the continental scale, it shows a clear response to precipitation [McCabe *et al.*, 2005; Draper *et al.*, 2009], and using a novel evaluation technique, Crow and Zhan [2007] showed that the assimilation of AMSR-E derived soil moisture into a simple water balance model added value to that model.

[4] In addition to recent advances in the remote sensing of soil moisture, there has also been focused development of suitable assimilation strategies for near-surface soil moisture observations. Early studies based on synthetic data showed that observation increments of near-surface soil moisture, or similarly microwave brightness temperature, can be propagated into the deeper soil layers [Reichle *et al.*, 2001; Walker and Houser, 2001]. Studies of single column models run over heavily instrumented field sites confirmed that such an assimilation can improve the model deep soil moisture [Seuffert *et al.*, 2004; Muñoz Sabater *et al.*, 2007]. Using remotely sensed data at the continental scale, Drusch [2007] used a simple nudging scheme to assimilate Tropical Rainfall Measuring Mission Microwave Imager derived near-surface soil moisture into the ECMWF Integrated Forecast System model over the southern United States, and Scipal *et al.* [2008] used the same method to assimilate European Remote Sensing scatterometer derived soil moisture globally. Both demonstrated that the nudging scheme improved the root-zone soil moisture (compared to ground data), and both recommended the development of a more sophisticated assimilation scheme. Using NASA's global Catchment Land Model (CLM), Ni-Meister *et al.* [2006] showed that an Ensemble Kalman Filter (EnKF) assimilation of near-surface soil moisture derived from Scanning Multichannel Microwave Radiometer (SMMR) generated improvements in the CLM soil moisture over Eurasia. Also using an EnKF with the CLM, Reichle *et al.* [2007] demonstrated modest improvements in the model root-zone soil moisture compared to ground data by assimilating near-surface soil moisture from SMMR and AMSR-E.

[5] A significant hurdle to the assimilation of near-surface soil moisture in NWP models has been the expense of the additional model integrations required by advanced assimilation methods. However, this expense can be reduced by assimilating the data into an off-line version of the land surface model. Using a simplified 2D-Variational assimilation approach, Balsamo *et al.* [2007] showed that the information content of different observation types (including screen-level variables and microwave brightness temperature) is similar for assimilation into either an off-line or atmospherically coupled land surface model. In the same experimental setup as used here, Mahfouf *et al.* [2009] developed a surface analysis for ALADIN, based on assimilating screen-

level temperature and humidity into an off-line version of its land surface scheme, the Interactions between Surface, Biosphere, and Atmosphere (ISBA) model. Mahfouf *et al.* [2009] used a Simplified EKF (SEKF), in which a static background error was assumed at the time of each analysis, and the observation operator was a 6-h ISBA integration. In an experiment over July 2006, they showed that the dynamic Kalman gain terms for the SEKF were similar to the analytically derived coefficients used in the operational Optimal Interpolation (OI) scheme [Giard and Bazile, 2000]. As well as confirming the viability of the SEKF, this demonstrates that the off-line system captures the necessary surface–screen-level interactions for the assimilation of screen-level observations, since the OI coefficients were derived using the full atmospheric model [Bouttier *et al.*, 1993].

[6] This work extends that of Mahfouf *et al.* [2009] to assimilate remotely sensed near-surface soil moisture derived from AMSR-E observations into ISBA, and is a preliminary step before the combined assimilation of remotely sensed near-surface soil moisture and screen-level observations. The screen-level data assimilated by Mahfouf *et al.* [2009] were reliably available for each analysis cycle, and the use of a static background error was based on the assumption that the increase in the background error during each forecast step was balanced by the reduction from the previous analysis. This assumption is less valid in this study, since the remotely sensed data used here are not available with the same regularity, motivating the development of a full EKF. Additionally, the dynamic error covariances will be of greater importance for the future combined assimilation of screen-level observations and near-surface soil moisture, which are available at different frequencies [Rüdiger *et al.*, 2007]. The aim of this study is to determine whether an off-line assimilation of remotely sensed near-surface soil moisture is a viable method for analyzing root-zone soil moisture in ISBA, and to understand how the near-surface soil moisture increments are translated into deeper-layer increments by the Kalman filter. Additionally, the benefit of using dynamic background error covariances is tested by comparing the SEKF and EKF assimilation of near-surface soil moisture.

2. Methodology

2.1. ISBA Land Surface Scheme

[7] ISBA [Noilhan and Mahfouf, 1996] is the land surface scheme used in the ALADIN NWP model. The moisture and energy dynamics in ISBA are modeled using a force-restore method [Deardorff, 1977], with eight prognostic variables: surface temperature, mean (deep layer) surface temperature, surface water content (liquid/frozen), total (deep layer) water content (liquid/frozen), vegetation intercepted water content, and snow water content. There is free drainage from the lower boundary, and each grid is divided into a vegetated and bare soil fraction, with evaporation calculated separately from each (although there is a single heat budget). For moisture, the near-surface water reservoir (w_1) is defined as the depth from which moisture can be extracted by bare soil evaporation (~ 10 mm), and the total water reservoir (w_2) is defined as the depth from which moisture can be extracted through bare-soil evaporation or transpiration (0.1 to 10 m, depending on the local soil type and climate). Both soil layers are forced by precipitation and

evaporation, and transpiration is applied to w_2 , with the atmospheric forcing acting more slowly on w_2 . The model restore term adjusts w_1 toward an equilibrium between capillary and gravity forces, while w_2 is restored toward field capacity by gravitational drainage. For ALADIN, the soil moisture and temperature states are currently analyzed from screen-level observations of humidity and temperature, using the OI technique of *Giard and Bazile* [2000].

[8] The EKF assimilation uses an off-line version of ISBA within the Surface Externalized (SURFEX) environment. In SURFEX the atmospheric forcing is applied at the first atmospheric model layer (17 m); this is higher than most off-line land surface models, to enable off-line assimilation of screen-level observations. For this experiment ISBA has been run in an environment that resembles the operational ALADIN model as closely as possible. One month of hourly forcing fields (precipitation, temperature, specific humidity, pressure, wind components, and short- and long-wave radiation) has been generated from ALADIN, and interpolated onto the ISBA time step (300 s). Eventually the surface analysis from SURFEX will be semicoupled to the NWP model, so that ALADIN is updated with the soil moisture analyses, and the SURFEX forcing supplied from the updated atmospheric forecast. However, for this initial investigation static forcing has been used, neglecting feedback between the soil moisture updates and the atmospheric forecasts. For further details of SURFEX, and how it would be coupled to the NWP for a land surface assimilation, refer to *Mahfouf et al.* [2009].

2.2. Soil Moisture From AMSR-E

[9] Near-surface soil moisture retrieved from AMSR-E brightness temperatures has been provided by the Vrije Universiteit Amsterdam (VUA) in collaboration with NASA-GSFC [*Owe et al.*, 2007]. C-band AMSR-E data is used here, since *Njoku et al.* [2005] showed that C-band Radio Frequency Interference (RFI) is not widespread across Europe (with the exception of isolated pockets over some urban areas). The descending AMSR-E data (approximate overpass time: 0130 LST) is used, since the nighttime soil moisture retrievals are more accurate [*Owe et al.*, 2001; *Draper et al.*, 2009]. The resolution of C-band AMSR-E data is 45×75 km [*Njoku et al.*, 2003], however the swath is oversampled at approximately every 5 km, and (level 2 and 3) C-band data is typically reported on a 0.25° grid, which is thought to approximate the scale of the information in the signal. The ALADIN France model has an irregular (stretched) grid, which covers most of Europe with resolution ~ 9.5 km. Rather than disaggregating the 0.25° AMSR-E data, the level 1 swath data have been regridded onto the ALADIN grid using a nearest neighbor approach.

[10] AMSR-E provides global cover in less than two days [*Njoku et al.*, 2003], with coverage decreasing toward the equator. For July 2006 the daily coverage over Europe is reduced from nearly 100% at 58°N , to 70% at 33°N . AMSR-E soil moisture data must be screened to remove data contaminated by RFI or open water, or where dense vegetation or frozen ground cover conceals the near-surface microwave signal. RFI contamination has been identified based on the RFI index of *Li et al.* [2004], which is provided by VUA with the soil moisture data. The only region shown as having significant RFI is Italy, and roughly

50% of the data over the Italian peninsula has been removed (in contrast to *Njoku et al.* [2005], who found limited C-band RFI over Europe, and X-band RFI over Italy in 2003, also using the index of *Li et al.* [2004]). Frozen ground cover is identified and removed during the moisture retrieval, although this is not expected to be significant in July. For vegetation, the VUA-NASA retrieval algorithm partitions the passive microwave signal into soil moisture and vegetation optical depth [*Owe et al.*, 2001]. The vegetation optical depth is linearly proportional to the vegetation water content, and the sensitivity of the microwave brightness temperature to soil moisture decreases with increasing vegetation optical depth [e.g., *de Jeu et al.*, 2008]. *Owe et al.* [2001] show that the soil moisture sensitivity is quite low for optical depths above about 0.75, and a mean monthly optical depth threshold of 0.8 has been adopted to screen out densely vegetated regions, following *de Jeu et al.* [2008].

2.3. Extended Kalman Filter

[11] The state forecast and update equations for the EKF are:

$$\mathbf{x}_t^b = \mathcal{M}_{t^-} [\mathbf{x}_{t^-}^a] \quad (1)$$

$$\mathbf{x}_t^a = \mathbf{x}_t^b + \mathbf{K}_t [\mathbf{y}_{t+6}^o - \mathcal{H}(\mathbf{x}_t^b)] \quad (2)$$

where

$$\mathbf{K}_t = \mathbf{B}_t^f \mathbf{H}_t^T (\mathbf{H}_t \mathbf{B}_t^f \mathbf{H}_t^T + \mathbf{R})^{-1} \quad (3)$$

\mathbf{x}_t indicates the model state at the time of the analysis, t , and the superscripts a and b indicate the analysis and background, respectively. \mathbf{y}_{t+6}^o is the observation vector (6 h after the analysis time). \mathcal{M}_{t^-} is the nonlinear state forecast model (ISBA) from the time of the previous analysis, t^- , \mathbf{K} is the Kalman gain, and \mathbf{B} and \mathbf{R} are the covariance matrices of the background and observation errors. \mathcal{H} is the nonlinear observation operator, and \mathbf{H} is its linearization. Here, the state variable consists of the superficial soil moisture (w_1) and the total soil moisture (w_2), and the observation operator is a 6-h integration of ISBA from time t . The AMSR-E near-surface soil moisture observations are assumed to occur at the end of the assimilation window, at 0000 UTC, and the quantity observed by AMSR-E is taken to be equivalent to the model w_1 (both represent the soil moisture in approximately the uppermost 10 mm of soil). The model update is made 6 h before the observation time (in contrast to *Mahfouf et al.* [2009] who add the increment at the observation time).

[12] The linearization of \mathcal{H} is obtained by finite differences, using a first-order Taylor expansion about \mathbf{x} . For each analysis cycle, this requires an additional (perturbed) 6-h model integration for each element of the state vector. For the i th observation, and the j th element of the control vector:

$$H_{ij,t} = \frac{\mathcal{H}(\mathbf{x}_t + \delta x_{j,t})_i - \mathcal{H}(\mathbf{x}_t)_i}{\delta x_j} \quad (4)$$

[13] The background error covariance matrix undergoes an analogous forecast and analysis cycle:

$$\mathbf{B}_t^f = \mathbf{M}_t \mathbf{B}_t^a \mathbf{M}_t^T + \mathbf{Q} \quad (5)$$

$$\mathbf{B}_t^a = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{B}_t^f \quad (6)$$

[14] In the forecast step (equation (5)), the previous analysis, \mathbf{B}_t^a , is forecast forward in time by the tangent linear of the state forecast model, \mathbf{M} , and the forecast error covariance matrix, \mathbf{Q} , is added to account for errors in the model forecast, giving the background error matrix forecast, \mathbf{B}_t^f . The model state analysis decreases the model error, and \mathbf{B} is reduced by an analysis step (equation (6)). The linearization of \mathcal{M} is obtained by the same finite difference method used for \mathbf{H} . The linearization of \mathcal{M} is made affordable by the assumption that there is no horizontal correlation in the model errors. The 24-h model Jacobian is estimated as the product of four 6-h Jacobians ($\mathbf{M}_{t \rightarrow t+24} = \mathbf{M}_{t+18 \rightarrow t+24} \mathbf{M}_{t+12 \rightarrow t+18} \mathbf{M}_{t+6 \rightarrow t+12} \mathbf{M}_{t \rightarrow t+6}$), to reduce the potential for nonlinearities.

[15] An alternative, and more common EKF formulation for the assimilation of near-surface soil moisture retrievals is to make the update at the time of the observations, and use $\mathcal{H} = (1 \ 0)$ rather than including the model in \mathcal{H} [e.g., Walker and Houser, 2001; Reichle et al., 2002; Muñoz Sabater et al., 2007]. This form of the EKF is analytically the same as that used here, except for the timing of the addition of \mathbf{Q} (see Appendix A). While the $\mathcal{H} = (1 \ 0)$ method avoids the additional integrations required here to linearize \mathcal{H} , it is dependent on the observed variable being included in the state vector. If the assimilation of remotely sensed soil moisture proves useful, it is intended that it be combined with the assimilation of screen-level observations. The screen-level observations cannot be sensibly included in the update vector, since they are not prognostic within SURFEX: they are diagnosed by interpolating the humidity and temperature between the ISBA surface and the (prescribed) value at the first atmospheric model layer. For consistency with Mahfouf et al. [2009] and future studies, the EKF form as initially described (ISBA included in \mathcal{H}) is used here.

[16] The error correlations for the AMSR-E and ALADIN soil moisture have been set based on the assumption that the standard deviation of the observed and modeled soil moisture errors are equal. The variance of the difference between the (rescaled) AMSR-E and ALADIN near-surface soil moisture (w_1) is 0.0061, which gives an error standard deviation of $0.055 \text{ m}^3 \text{ m}^{-3}$ for each, assuming that the observation and model error are independent and unbiased. This agrees closely with the error estimates that have been made for AMSR-E (see, for example, Table 1 of de Jeu et al [2008]), and the observation error standard deviation has been set at this value. Only the diagonal entries of the model error matrices (\mathbf{B} and \mathbf{Q}) have been prescribed (i.e., error cross correlations have not been applied), and following Mahfouf et al. [2009] the model soil moisture errors are assumed to be proportional to the soil moisture range (the difference between the volumetric field capacity (w_{fc}) and the wilting point (w_{wilt}), calculated as a function of soil type, as given by Noilhan and Mahfouf [1996]). For \mathbf{B} , the initial

model error standard deviations for both w_1 and w_2 have been set to $0.6 \times (w_{fc} - w_{wilt})$, which converts to a mean volumetric error standard deviation of $0.052 \text{ m}^3 \text{ m}^{-3}$, slightly lower than that used for the observation error. The magnitude of the diagonal elements of \mathbf{Q} were selected to minimize long-term tendencies in \mathbf{B} , on the assumption that \mathbf{B} should not dramatically increase or decrease over time. Using this method values of $0.3 \times (w_{fc} - w_{wilt})$ and $0.2 \times (w_{fc} - w_{wilt})$ were chosen for the w_1 and w_2 error standard deviations, respectively. The EKF assimilation is compared to a SEKF assimilation of soil moisture, which neglects the evolution of the background error (equations (5) and (6)), and assumes that \mathbf{B} is constant at the start of each analysis cycle (some dependence on the conditions of the day is introduced through the use of the model in \mathcal{H}). For the SEKF analysis, the same \mathbf{R} was used, and \mathbf{B} was set at the same initial value as was used for the EKF.

3. Results

3.1. Scaling the Observations to the Model Climatology

[17] Since the soil moisture quantity observed by remote sensors differs from that defined in models, soil moisture data must be rescaled before assimilation, so as to be consistent with the model climatology [Reichle et al., 2004]. Here, the AMSR-E data are rescaled by matching its Cumulative Distribution Function (CDF [Reichle and Koster, 2004; Drusch et al., 2005]) to that of the superficial soil moisture forecast by ALADIN for 0000 UTC each day (this is the 6-h forecast from 1800 UTC, which provides the first guess for the operational soil moisture analysis). Ideally, a long data set is used to sample the model and observation climatology and the CDF matching is performed on as localized a scale as possible, however for this study only 1 year of ALADIN soil moisture fields are available. Reichle and Koster [2004] demonstrated that the CDF matching operator can be estimated from 1 year of data by using spatial averaging to compensate for the reduced temporal sample size, and the CDF-matching operator has been estimated here using a one-degree window around each grid cell.

[18] CDF matching is based on the assumption that the differences between the model and observations are stationary, however for ALADIN and AMSR-E this is not the case. For example, Figure 1 shows a time series of the AMSR-E data before and after the CDF matching at a location in northern France. While both the AMSR-E and ALADIN time series have a similar range of short-term (up to several days) variability with amplitude between 0.1 and $0.2 \text{ m}^3 \text{ m}^{-3}$, the seasonal cycle in the AMSR-E data has a greater magnitude ($>0.2 \text{ m}^3 \text{ m}^{-3}$) than that in the ALADIN data ($\sim 0.1 \text{ m}^3 \text{ m}^{-3}$). To compensate for the variance generated by the enhanced seasonal cycle in the AMSR-E data, the CDF matching has overly dampened the short-term variability, resulting in a lessened response to rain events in the CDF-matched time series. Additionally, at the seasonal to monthly scale there are biases in the CDF-matched time series (e.g., around day 100 in Figure 1).

[19] To avoid these problems, the CDF matching has been repeated using seasonally bias-corrected AMSR-E data, generated by subtracting the observation-model difference in the 31-day moving average. The resulting time series in Figure 1 has retained an appropriate response to

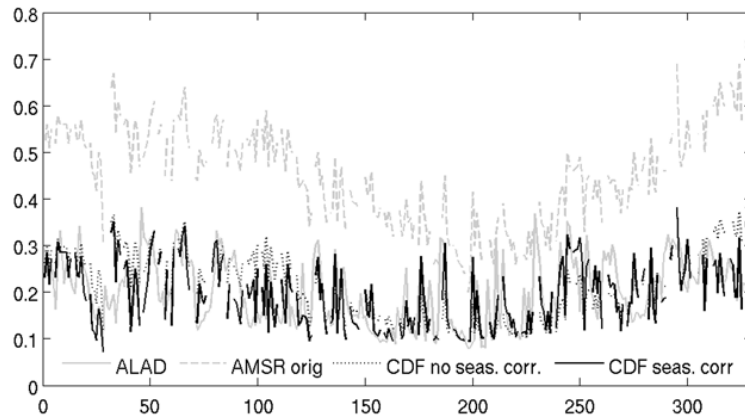


Figure 1. Time series of near-surface soil moisture ($\text{m}^3 \text{m}^{-3}$) for a grid cell in France (47.30 E/0.06 N) over 2006, from ALADIN (gray, solid), the original AMSR-E data (gray, dashed), and the seasonally corrected (black, solid) and nonseasonally corrected (black, dotted) CDF-matched AMSR-E data.

precipitation, and the monthly biases are reduced. For this experiment (July 2006) the mean monthly bias is reduced from $-0.014 \text{ m}^3 \text{m}^{-3}$ in the initial CDF-matched data to $0.001 \text{ m}^3 \text{m}^{-3}$ in the seasonally corrected and CDF-matched data, compared to $0.14 \text{ m}^3 \text{m}^{-3}$ in the original data. If a longer data set were available the difference in the seasonal cycles could be removed based on the climatological seasonal cycles, which would retain any seasonal bias anomalies in the observations. With just 1 year of data seasonal bias anomalies cannot be detected (regardless of the method used to rescale), and the approach used here is necessarily conservative, assuming that the ISBA 2006 seasonal cycle was correct.

[20] The mean monthly RMSD between the CDF-matched (and seasonally corrected) AMSR-E data and the ALADIN w_1 is $0.007 \text{ m}^3 \text{m}^{-3}$. Figure 2 shows that the RMSD is relatively large ($>0.09 \text{ m}^3 \text{m}^{-3}$) over most of the Italian peninsula, where much of the data was removed due to RFI, suggesting that the remaining data is of poor quality (all data were rejected at locations with less than 100 observations over 2006, and the poor match is unlikely to be due to reduced data coverage). The RMSD is also relatively high in many locations adjacent to regions screened for dense vegetation, most likely due to increased error in the AMSR-E data due to vegetation interference. Additionally, the higher RMSD over the Alps and the Pyrenees could be due to inaccuracies in the model and/or the data, since both have known problems in regions of steep terrain [Rüdiger *et al.*, 2009].

3.2. Tangent-Linear Approximation

[21] The magnitude of the perturbations used to estimate \mathbf{M} was chosen by examining the difference between the Jacobians estimated using positive and negative perturbations for a range of magnitudes, following Walker and Houser [2001] and Balsamo *et al.* [2004]. On the basis of this method, a perturbation of $10^{-4} \times (w_{fc} - w_{wilt})$ was selected for estimating \mathbf{M} , and also \mathbf{H} (the linearization of \mathcal{H} is not discussed here, since linearity over 24 h strongly suggests linearity over 6 h). The difference between the Jacobians estimated with the positive and negative perturbations gives a measure of the nonlinearity of \mathcal{M} for perturbations of that size. Scatterplots of the Jacobian terms estimated with positive and negative perturbations of magnitude $10^{-4} \times (w_{fc} - w_{wilt})$ for the analysis cycle on 1 July 2006 show virtually all of the points aligned along the one-to-one line (not shown), consistent with \mathcal{M} being well approximated by \mathbf{M} within the range of the applied perturbation. This is confirmed by the statistics in Table 1, which show little difference between the mean, standard deviation, and extreme values for the Jacobians estimated with the positive and negative perturbations. The extreme sensitivity causing the very large maximum values in Table 1 for perturbed w_1 is quite rare, and less than 0.2% of the grid cells have a $\partial w_2(t+24)/\partial w_1(t)$ or $\partial w_1(t+24)/\partial w_1(t)$ greater than 10. Mahfouf *et al.* [2009] used the same perturbation size to estimate the Jacobians for ISBA over 6 h, yet the occurrence of nonlinearities as observed by Mahfouf *et al.* [2009] does not occur here, since the (dissipative) land-surface component of the model is less prone to the nonlinearities that can occur in the atmosphere.

[22] The above analysis indicates that \mathcal{M} is well approximated by \mathbf{M} within the range of the very small perturbations

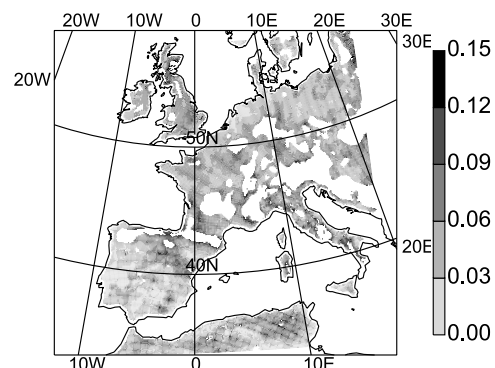


Figure 2. Root-mean-square difference ($\text{m}^3 \text{m}^{-3}$) between the CDF-matched AMSR-E near-surface soil moisture and ALADIN w_1 over 2006. White indicates that no AMSR-E soil moisture data are available.

Table 1. Statistics of 24-h Jacobian Terms From 1800 UTC on 1 July 2006^a

		Mean	SD	Minimum	Maximum
$\partial w_1(t+24)/\partial w_1(t)$	+ve	0.25	1.9	-0.11	189
	-ve	0.26	1.9	-0.11	189
	lrg	0.30	1.9	-0.10	67
$\partial w_2(t+24)/\partial w_1(t)$	+ve	-0.0024	0.0045	-0.15	0.0050
	-ve	-0.0024	0.0045	-0.15	0.0045
	lrg	-0.0025	0.0055	-0.19	0.0
$\partial w_1(t+24)/\partial w_2(t)$	+ve	0.60	0.96	-2.4	120
	-ve	0.60	0.97	-0.37	120
	lrg	0.60	0.53	-0.15	23
$\partial w_2(t+24)/\partial w_2(t)$	+ve	0.95	0.054	0.21	1.0
	-ve	0.95	0.053	0.42	1.0
	lrg	0.96	0.044	0.00	1.0

^aEstimated using a perturbation size of $+10^{-4} \times (w_{fc} - w_{wilt})$ (positive), $-10^{-4} \times (w_{fc} - w_{wilt})$ (negative), and $+10^{-1} \times (w_{fc} - w_{wilt})$ (large). Units are in (%/‰).

that were applied, however this does not guarantee that **M** approximates **M** well when applied to the errors in **B**, since these errors are typically much larger than the applied perturbations. To test the potential error generated when **M** is used to propagate **B**, the model Jacobians estimated using perturbations with magnitude similar to the expected model error ($10^{-1} \times (w_{fc} - w_{wilt}) \sim O(10^{-2})$) have been compared to the above estimates. Table 1 shows that the mean Jacobian estimates for this larger perturbation are very similar to those based on the smaller perturbations, although the distribution of values about the mean is different, with differences in their extreme values and variances. The difference between the Jacobians estimated with the smaller and larger perturbation is greater than 0.1 for 4% ($\partial w_1(t+24)/\partial w_1(t)$), 0% ($\partial w_2(t+24)/\partial w_1(t)$), 23% ($\partial w_1(t+24)/\partial w_2(t)$), and 10% ($\partial w_2(t+24)/\partial w_2(t)$) of the grid cells, indicating that the larger perturbation is outside the model's linear regime in more instances (for the positive and negative perturbations considered above the difference was greater than 0.1 for less than 1% of the grids for all of the Jacobian terms). However, the Jacobian estimates compare favorably over the majority of grid cells, indicating that **M** estimated with perturbations of $10^{-4} \times (w_{fc} - w_{wilt})$ leads to an acceptable approximation of nonlinear **M** for propagating **B** forward 24 h.

3.3. ISBA Jacobians

[23] The ISBA Jacobians reflect the force-restore dynamics of the model. The superficial soil layer responds rapidly to atmospheric forcing, so that a perturbation applied to w_1 is gradually reduced over 24 h. As a result the mean $\partial w_1/\partial w_1$ is reduced from 0.80 over 6 h to 0.25 over 24 h (with the Jacobians estimated from 1800 UTC on 1 July 2006). In addition to its short timescale, w_1 represents a very small physical reservoir, and cannot influence w_2 strongly, so that $\partial w_2/\partial w_1$ is insignificant (mean < 0.01 over 6 or 24 h). In contrast to w_1 , the atmospheric forcing is applied more slowly to the total soil moisture, and w_2 has a timescale of 10 days. Over a comparatively short 24-h period a w_2 perturbation is largely retained (mean $\partial w_2(t+24)/\partial w_2(t)$: 0.95). The influence of w_2 on w_1 increases over time, and the mean $\partial w_1(t+6)/\partial w_2(t)$ is 0.20, increasing to 0.60 for $\partial w_1(t+24)/\partial w_2(t)$. Since w_1 does not have a strong or persistent influence on the other surface variables, its accurate

analysis is less important than that of w_2 . While w_1 could then be excluded from the control variable, (to reduce the number of linearization required), this would result in an underestimation of the model w_1 error, since a large component of this is due to short-lived errors (i.e., the element \mathbf{q}_{11} of the **Q** matrix).

[24] The background error matrix used in each analysis is largely derived from the previous w_2 error correlations and the applied (static) **Q**, since the w_1 errors are short-lived and do not influence w_2 . The important terms in the 24-h linear tangent model are then $\partial w_1(t+24)/\partial w_2(t)$ and $\partial w_2(t+24)/\partial w_2(t)$, both of which are shown in Figure 3 for a 24-h period. The Soil Wetness Index (SWI; $\text{SWI} = (w_2 - w_{wilt})/(w_{fc} - w_{wilt})$), a measure of soil water availability in the root zone, is provided in Figure 4 for comparison. Over the full diurnal cycle the ISBA moisture dynamics, and hence Jacobians, are dominated by the force component (precipitation and evapotranspiration) of its force-restore scheme. The addition of moisture from precipitation reduces the sensitivity of w_1 to w_2 , and the reduced $\partial w_1(t+24)/\partial w_2(t)$ across much of southeast Europe and in smaller regions in northern Spain and along the Pyrenees was caused by rain (these locations have been excluded from the statistics given below). In the absence of precipitation, the 24-h Jacobians are most strongly influenced by evapotranspiration, specifically its sensitivity to w_2 . Under dry conditions the parameterization of transpiration depends

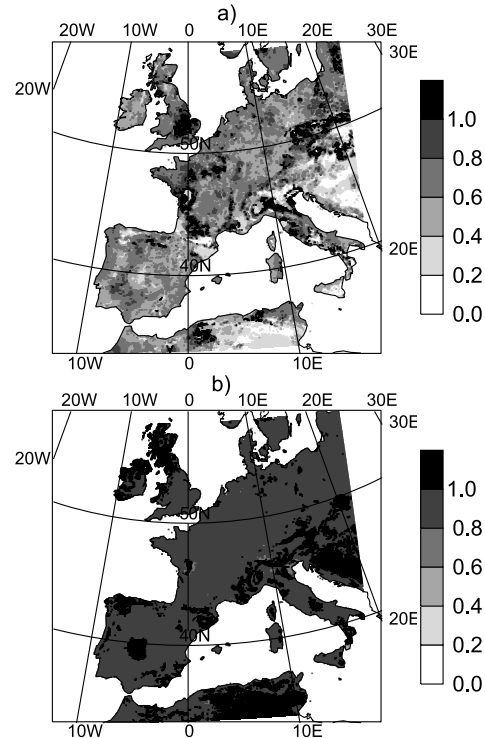


Figure 3. Jacobian terms for the forward model (**M**): (a) $\partial w_1(t+24)/\partial w_2(t)$ and (b) $\partial w_2(t+24)/\partial w_2(t)$ from 1800 UTC on 1 July to 1800 UTC on 2 July 2006. Note that $\partial w_1(t+24)/\partial w_1(t)$ and $\partial w_2(t+24)/\partial w_1(t)$ are not shown, as w_1 has little memory over 24 h.

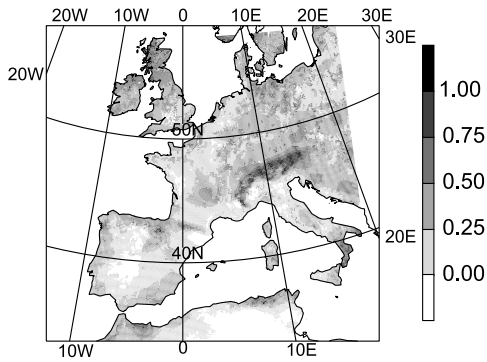


Figure 4. Surface wetness index at 1800 UTC on 2 July 2006 from the EKF analysis.

strongly on w_2 , with the dependence increasing as w_2 approaches w_{wilt} . In these moisture limited conditions a small increase in w_2 generates a relatively large increase in transpiration, reducing w_2 , and hence $\partial w_2(t+24)/\partial w_2(t)$. In turn, this leads to an increase in $\partial w_1(t+24)/\partial w_2(t)$ when the enhanced transpiration reduces the surface temperature, which reduces the depletion of w_1 by bare-ground evaporation, giving a relative increase in w_1 . This mechanism has been confirmed by testing the impact on the model forecasts of switching off aspects of the model physics. It is also evidenced in Figure 3. Most obviously, where w_2 is below the wilting point, transpiration ceases and w_2 perturbations are not communicated to w_1 . As a result, the regions of negative SWI in Figure 4 in North Africa, and also in Spain and France, correspond to $\partial w_2(t+24)/\partial w_2(t)$ close to 1 (for negative SWI, the mean $\partial w_2(t+24)/\partial w_2(t)$ is 1, compared to 0.96 across the whole domain), and reduced $\partial w_1(t+24)/\partial w_2(t)$ (to a mean of 0.31, compared to 0.64 for the whole domain). For the rest of the domain where the SWI is positive, the Jacobians are most sensitive to moisture availability where there is sufficient vegetation present to generate substantial transpiration. Where the fractional vegetation cover in ISBA is greater than 0.5, $\partial w_2(t+24)/\partial w_2(t)$ is reduced (mean: 0.92) and $\partial w_1(t+24)/\partial w_2(t)$ is increased (mean: 0.79) where the SWI is below 0.25 (compared to means of 0.96 and 0.68, respectively for all grids with positive SWI and fractional vegetation greater than 0.5). In contrast, where the fractional vegetation cover is less than 0.5, there is no obvious difference in the $\partial w_2(t+24)/\partial w_2(t)$ across all grids and across grids with SWI less than 0.25 (mean 0.96 for both), while $\partial w_1(t+24)/\partial w_2(t)$ is slightly reduced in the drier locations (mean 0.56, compared to 0.59 for all sparsely vegetated cells). The role of vegetation can be seen in Figure 3. Over France and the UK, where the vegetation fraction is greater than 0.75, $\partial w_1(t+24)/\partial w_2(t)$ is generally elevated (>0.8) where the SWI in Figure 4 is low (<0.25), yet in sparsely vegetated Spain (vegetation fraction <0.5), where the SWI is similarly low there is no such relationship.

[25] The 6-h model Jacobians used in the observation operator are plotted in Figure 5. While the 24-h Jacobian terms reflect ISBA's force component, for the descending pass AMSR-E data used here the 6-h Jacobians are estimated during the night, when the forcing is weak (excepting

regions of rain). In the absence of strong forcing, ISBA restores w_1 toward w_2 to achieve a balance between capillary rise and gravitational drainage. This introduces a weak nighttime sensitivity of w_1 to w_2 resulting in a mean $\partial w_1(t+6)/\partial w_2(t)$ of 0.20 (as already noted, this is much lower than the corresponding value over 24 h). There is also less spatial variability in the 6-h nighttime Jacobians (for $\partial w_1/\partial w_2$ the variance over 6 h is 0.007, compared to 0.1 over 24 h). Owing to the absence of strong forcing and the slow timescale of the model restore term, $\partial w_1(t+6)/\partial w_1(t)$ is reasonably high, with a mean of 0.80. Comparison of Figures 4 and 5 suggests a tendency for decreased $\partial w_1(t+6)/\partial w_2(t)$ where the SWI is lower ($\partial w_1(t+6)/\partial w_1(t)$ is also slightly increased in these regions, although this is not evident at the plotted scale), suggesting that capillary rise increases nonlinearly with increasing surface water availability. There is an additional influence from the soil type in ISBA (not shown), with lower clay content giving more rapid flow through the soil, corresponding to increased $\partial w_1(t+6)/\partial w_2(t)$ and decreased $\partial w_1(t+6)/\partial w_1(t)$.

3.4. Kalman Gain

[26] Figure 6 shows the Kalman gain for w_2 (\mathbf{k}_2) for the EKF and the SEKF on 2 July. The EKF \mathbf{k}_2 is between 0.2 and 0.4 across most of Europe, with a mean of 0.27. The SEKF gain is smaller, due to the slightly larger background errors used in this experiment, and is generally less than 0.2, with a mean of 0.12. The spatial patterns for the two gain terms differ, since they are determined by different processes. For the SEKF, the (static) \mathbf{B} is evolved 6 h by \mathbf{H} , and there is

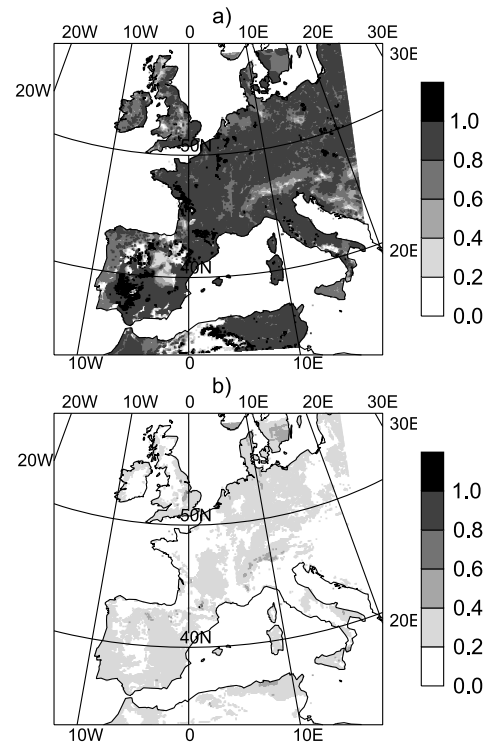


Figure 5. Jacobian terms for the observation operator (\mathbf{H}): (a) $\partial w_1(t+6)/\partial w_1(t)$ and (b) $\partial w_1(t+6)/\partial w_2(t)$, from 1800 to 2400 UTC on 2 July 2006.

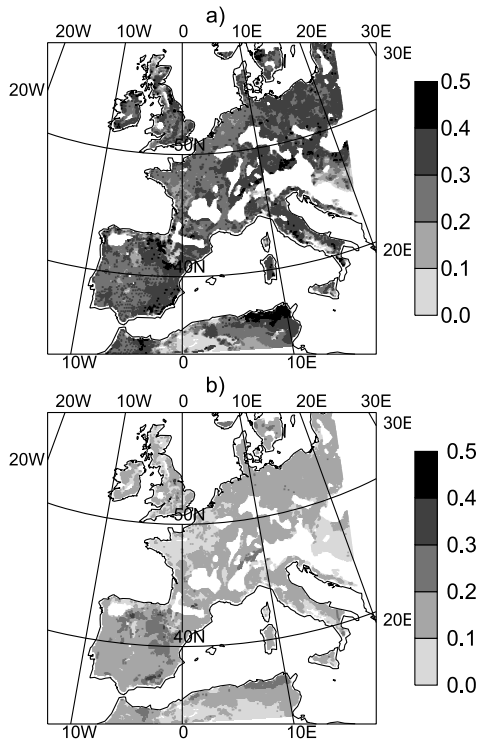


Figure 6. Kalman gain ($\text{m}^3 \text{ m}^{-3}/\text{m}^3 \text{ m}^{-3}$) for w_2 at 1800 UTC on 2 July 2006 for the (a) EKF and (b) SEKF.

a qualitative correspondence in Figure 6b (SEKF gain) and Figure 5 (H), with the gain being reduced where $\partial w_1(t+6)/\partial w_2(t)$ is lower (e.g., northern France and North Africa). In contrast, the EKF gain is determined by a combination of the Jacobian terms over 6 h (H) and 24 h (M), and it shows a combination of features from both. Even after a single assimilation cycle, the 24-h M has introduced much more fine-scale spatial heterogeneity into the EKF gain than is present in the SEKF gain.

3.5. Soil Moisture Analyses

[27] The time series in Figure 7 shows the soil moisture states for the EKF and the SEKF at four locations with contrasting conditions, together with an open-loop simulation, in which the model surface is allowed to evolve without data assimilation. The grid cells in Slovakia and France are in vegetated regions where transpiration links w_2 to w_1 ($\partial w_1(t+24)/\partial w_2(t) \sim 0.6$ in Figure 3 for both). In both cases the observations are consistently higher than the model forecast w_1 , and the assimilation improves the fit between the model w_1 and the observations by adding moisture to w_2 . The EKF and SEKF produce similar results, except for a few isolated large increments generated by the EKF in Slovakia. The observation increments in Slovakia are unusually large, particularly in the first part of the month where the observations (if correct) suggest a precipitation event not present in the model forcing. As a result, a large volume of water is added in Slovakia, and the difference between the w_2 SWI for the analyses and the open loop approaches 0.5 at times. The large observation increment on day 18 in France does not

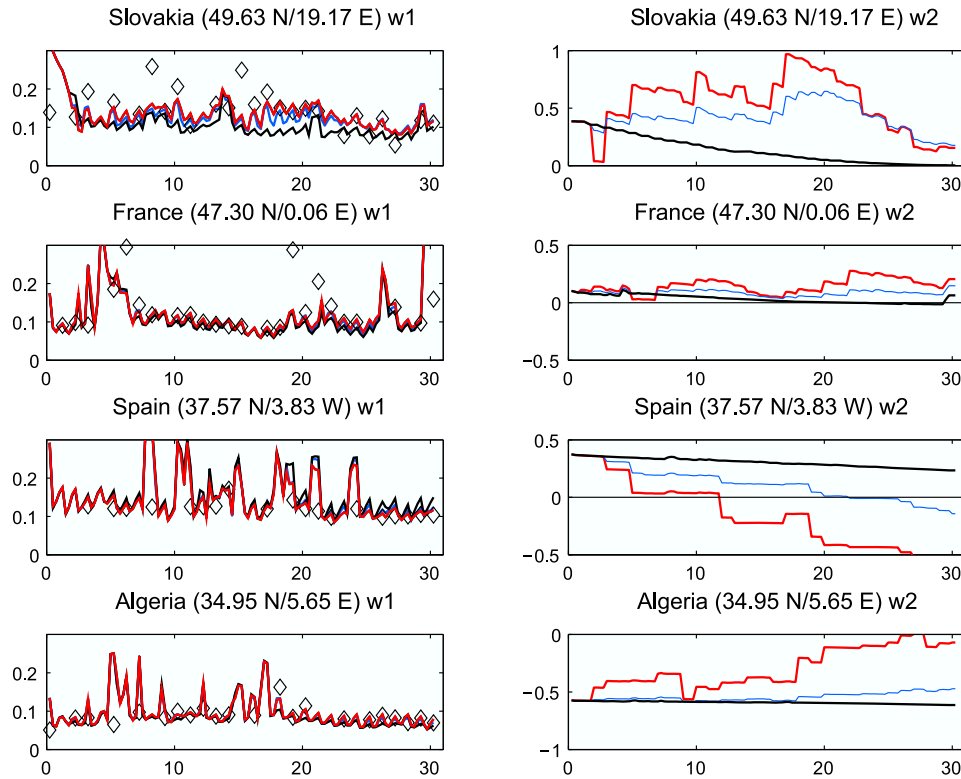


Figure 7. Time evolution of (left) w_1 ($\text{m}^3 \text{ m}^{-3}$) and (right) w_2 (as SWI) through July 2006, from the open loop (black), EKF (red), and SEKF (blue). Observations of w_1 are indicated as diamonds.

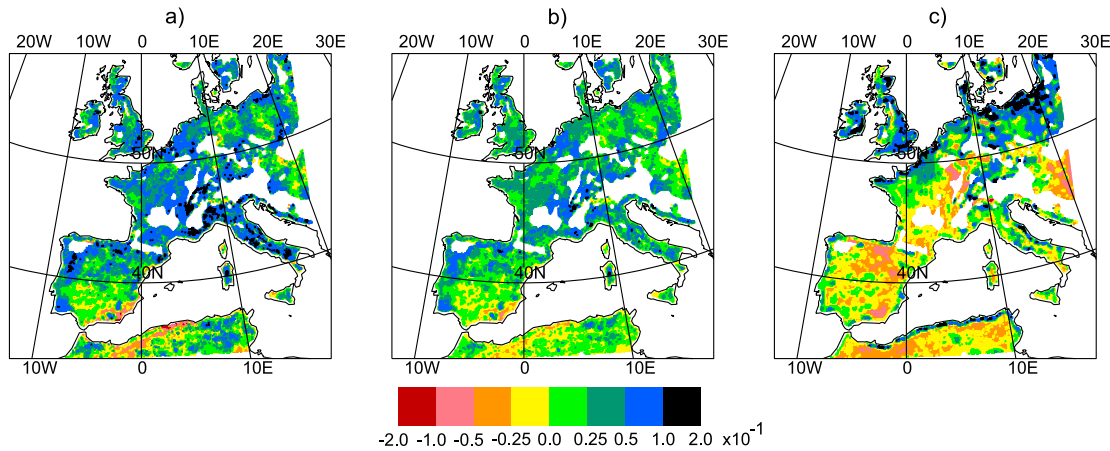


Figure 8. The net monthly w_2 increments ($\text{m}^3 \text{m}^{-3}$) over July 2006, from the (a) EKF and (b) SEKF for assimilation of the seasonal-bias corrected AMSR-E data, and (c) EKF assimilation of the nonseasonal-bias corrected AMSR-E data.

translate into an analysis increment as it is screened out by the observation quality control, which discards data more than $0.1 \text{ m}^3 \text{m}^{-3}$ away from the model w_1 .

[28] The grid cells in Spain and Algeria both have sparse vegetation cover, and with limited transpiration the dependence of w_1 on w_2 is weaker, particularly in Algeria, where w_2 is below the wilting point, and $\partial w_1(t+24)/\partial w_2(t)$ in Figure 3 is ~ 0.05 (compared to ~ 0.2 in Spain). In Spain the observations are generally lower than the model w_1 , and the analysis consistently decreases w_2 . In the first half of the month, despite the SWI having been decreased by 0.5, neither analysis generates substantial changes in w_1 , and the analysis continues to deplete w_2 until a very large net increment is evolved. It is only after the w_2 SWI has been decreased by nearly one that the analysis generates a slight reduction in w_1 (which does give a better fit to the data). A similar situation occurs in Algeria where the observations are consistently above the model, and the analysis makes a series of positive increments to w_2 , which do not affect w_1 until a large net change is accumulated. By the end of the month w_2 is approaching a SWI of 0, and w_1 shows an enhanced diurnal cycle, with greater nighttime increases. In both of these cases, since there is little transpiration to expose w_1 to the w_2 increments (compare the ratio of the net change in w_2 and w_1 to that from the previous examples), the analysis continues to make monotonic corrections to w_2 until a large (and likely erroneous) net increment has been imposed on w_2 . Initially, the SEKF and EKF increments are similar in magnitude, however as the month progresses the EKF increments become larger. This is due to an inflation of the background error where transpiration is limited. Recall from Figure 3 that $\partial w_2(t+24)/\partial w_2(t)$ approaches one in regions with little transpiration. As a result the \mathbf{b}_{22} element of \mathbf{B} is not decreased during the forecast step (equation (5)), and \mathbf{B} gradually increases with each addition of \mathbf{Q} (which is generally greater than the analysis reduction). While the \mathbf{Q} used here does not generate a discernible trend in \mathbf{B} across the remainder of the domain, \mathbf{b}_{22} in north Africa and Spain is almost doubled within two weeks.

[29] Figure 8 shows the net soil moisture increments added by the EKF and the SEKF over July 2006. Even though the

Kalman gain terms for each depend on different aspects of the model physics, the resultant analyses are similar. For both the EKF and the SEKF moisture has been added across most of the domain, except for areas in southern Spain and central North Africa (as well as some smaller isolated in northern and eastern Europe). The mean monthly net increment is $0.025 \text{ m}^3 \text{m}^{-3}$ for the EKF and $0.018 \text{ m}^3 \text{m}^{-3}$ for the SEKF. The EKF has a greater spread of increments, with more extreme values (both positive and negative), resulting in a larger standard deviation of the net monthly increment ($0.037 \text{ m}^3 \text{m}^{-3}$) than for the SEKF ($0.023 \text{ m}^3 \text{m}^{-3}$). Some of the very large ($>0.1 \text{ m}^3 \text{m}^{-3}$) increments for the EKF surrounding the Alps correspond to the high RMSD between the CDF-matched AMSR-E and ALADIN soil moisture in Figure 2, where there are known errors in both the model and observations (section 3.1). The increments are also large over Italy, where the coverage of AMSR-E data is limited by RFI, and the quality of the remaining data is questionable. In general, the analysis increments are relatively large compared to the dynamics of w_2 , being approximately the same magnitude as the mean range of w_2 throughout July 2006 ($0.02 \text{ m}^3 \text{m}^{-3}$). In terms of the net volume of water added to the surface, the EKF added a monthly mean volume of 55 mm, while a mean of 41 mm was added by the SEKF (with a total range of approximately ± 200 mm for both). This represents a substantial component of the monthly water balance, and is similar to the mean monthly volume added by precipitation (50 mm). Similarly large increments were obtained by Mahfouf *et al.* [2009] for the assimilation of screen-level observations. The large volume of water being added (or removed) is partly due to the two-layer structure of ISBA, since increments to w_2 must be applied across the total soil depth, leading to large net increments, as discussed by Mahfouf *et al.* [2009].

[30] Owing to the strong seasonal cycle in the AMSR-E–ALADIN bias, it was necessary to bias-correct the seasonal cycle in the AMSR-E data before performing the CDF matching (section 3.1). To highlight the central role of data quality to the analysis, the EKF assimilation has been repeated with the original CDF-matched (no seasonal bias correction) AMSR-E data. In this case, the resultant analyses

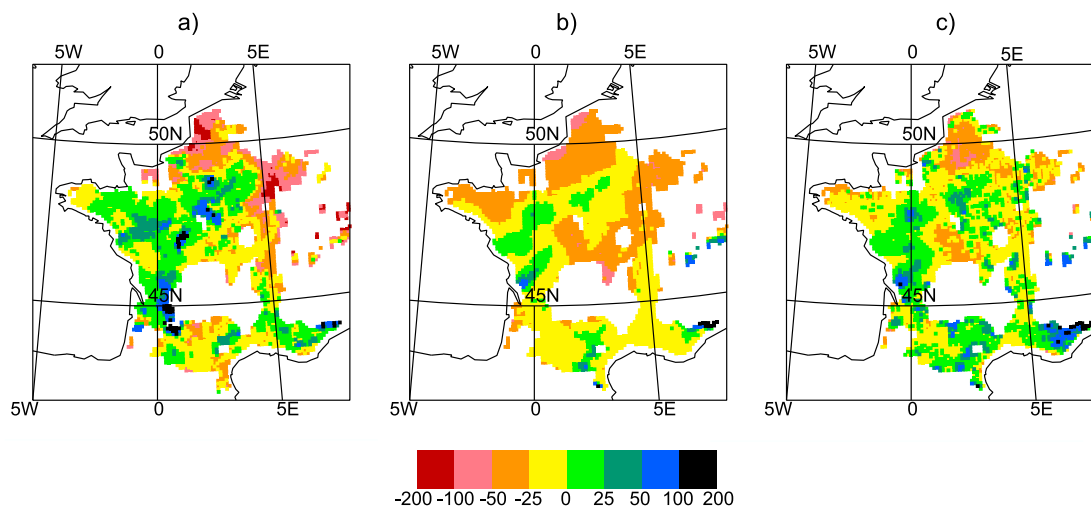


Figure 9. Change in total soil moisture storage (mm) from 1 to 31 July 2006, from (a) SIM, (b) open loop, and (c) the EKF.

differed substantially from the results obtained with the seasonally bias-corrected data. The mean absolute difference between the net monthly increments produced by assimilating the AMSR-E data with and without seasonal bias correction was $0.039 \text{ m}^3 \text{ m}^{-3}$, compared to a mean difference between the EKF and SEKF assimilation (for the seasonally bias-corrected data) of $0.014 \text{ m}^3 \text{ m}^{-3}$. Figure 8c shows the net monthly soil moisture increments for the assimilation of the nonseasonally bias-corrected data. In this case the quality control (removal of all data greater than $0.01 \text{ m}^3 \text{ m}^{-3}$ from the model w_1) was not applied as this resulted in most data being removed. Figure 8c is quite different from the previous two panels, and has net positive increments in northern Europe and net negative increments elsewhere, consistent with the strong negative bias over July 2006 in the nonseasonally bias corrected AMSR-E data.

[31] The current operational analysis also includes analysis of the soil temperature. While it is beyond the scope of this paper, an additional experiment was carried out with soil temperature included in the control variable. This experiment showed that total (deep layer) soil temperature analysis increments can also be obtained from near-surface soil moisture observations, and also that the inclusion of temperature has a slight effect on the soil moisture analysis, but does not alter the main findings presented here.

3.6. Comparison to SIM Water Balance

[32] While the focus here is on the mechanics of the assimilation, the resultant analyses have been reality checked by comparison to simulations from SAFRAN-ISBA-MODCOU (SIM [Habets et al., 2008]). SIM is a three-layer version of ISBA forced with high-quality data [Quintana-Seguí et al., 2008] over France. The soil moisture from SIM compares favorably to other estimates of soil moisture [Rüdiger et al., 2009], and it can be regarded as the best available estimate of the true surface state over France. As outlined by Mahfouf et al. [2009], the total change in column soil moisture over a time period gives an integration of the surface-moisture inputs (precipitation), outputs (evapotranspiration, runoff), and soil moisture increments

(where an assimilation is performed). Figure 9 shows the change in the total-column soil moisture over July 2006 from SIM (Figure 9a), the open loop (Figure 9b), and the EKF (Figure 9c) (the SEKF is not included, since its results are very similar to the EKF). The open loop is forced with the same ALADIN forecasts used in the EKF experiment, and the difference between the change in soil moisture from SIM and from the open-loop simulation will be predominantly due to errors in these forecasts. It is hoped that the assimilation can correct for some of the forcing errors, bringing the total change in soil moisture closer to that from SIM.

[33] In comparison to SIM (Figure 9a), the open loop (Figure 9b) has a tendency toward excessive drying and insufficient wetting, resulting in a mean monthly change in soil moisture for the open loop of -20 mm , compared to -11 mm for SIM. The open loop has generated incorrect drying (in spatial extent and magnitude) along the English Channel coast and in central France, with a region of insufficient moistening in between, associated with a low bias in the ALADIN precipitation forcing. Also, the open loop did not moisten the regions along the Atlantic Coast and south of the Alps indicated by SIM. The EKF (Figure 9c) has added moisture across most of France, increasing the mean monthly increment to -7 mm (overshooting the SIM mean). The EKF shows a general improvement in the correspondence to SIM. It has corrected the band of insufficient moistening in the north, as well as the lack of moistening along the Atlantic coast and in the southeast. However, in the east it has degraded the open loop by adding moisture where drying was correctly identified in the open loop.

4. Discussion

[34] This experiment has demonstrated that the total soil moisture in ISBA can be analyzed using an EKF from remotely sensed near-surface soil moisture observations, in this case from AMSR-E. Assuming that the background and observation errors are (approximately) equal, the EKF Kalman gain over July 2006 was typically around 20–30%, giving a mean net monthly increment of $0.025 \text{ m}^3 \text{ m}^{-3}$

(equivalent to 55 mm of water added to the soil column). While the EKF increments are large compared to the model dynamics and water balance, they are similar in magnitude to the increments generated by Météo-France's operational OI scheme over the same period [Mahfouf *et al.*, 2009]. Comparison of the monthly water balance generated by the EKF analysis to that from SIM over France showed a general improvement compared to an open loop, although some areas were degraded. The EKF requires the linearization of the forecast model in order to propagate background errors through time. The inaccuracy introduced by the linearization has been estimated by comparing the model Jacobians (calculated using a perturbation small enough that the model is approximately linear) to the Jacobians generated by applying a perturbation of the approximate size of the expected background errors. This test indicated that the linearization provides a good approximation of the model Jacobians for use in the EKF in most instances.

[35] Since w_1 does not directly influence w_2 in ISBA, the analysis of w_2 from w_1 observations must utilize the sensitivity of w_1 to changes in w_2 . The effectiveness of the assimilation is then limited by the strength of $\partial w_1(t+24)/\partial w_2(t)$. Over the diurnal cycle $\partial w_1(t+24)/\partial w_2(t)$ is dominated by daytime radiative forcing and the influence of w_2 on w_1 is principally determined by the transpiration physics (enhanced w_2 causes enhanced transpiration, causing decreased superficial soil temperature, giving decreased bare soil evaporation, and a relative increase in w_1). The greatest sensitivity, and hence most effective analysis of w_2 , occurs where transpiration is most sensitive to w_2 : in reasonably vegetated regions when w_2 is close to, but above, w_{wilt} . Conversely, where w_2 is less than w_{wilt} , or is very high (so that transpiration is not moisture limited), or where there is little vegetation, w_2 does not substantially influence w_1 , and so cannot be effectively analyzed from w_1 observations.

[36] A major motivation for using remotely sensed near-surface soil moisture in NWP is the expectation that it will provide a more direct observation of total soil moisture than screen-level observations do, since the latter rely on the model flux parameterizations to link the surface state to the screen-level atmosphere. However it has been shown here that for ISBA the link between the near-surface soil moisture observations and the deeper soil moisture is still provided by transpiration. This result is derived from the model physics, and it is expected that many other models, such as multilayer models with more substantial surface layers and more explicit drainage, will provide a more direct relationship between the near-surface and deeper soil moisture.

[37] During the nighttime, when the surface forcing is weak, the sensitivity of w_1 to w_2 in ISBA is determined by the model restore term, representing the balance between capillary rise and gravitational drainage. Since the descending AMSR-E data used here are observed at night, the gain terms are influenced by the nighttime dynamics through the observation operator (this also occurs for the alternate EKF formulation discussed in Appendix A, due to a diurnal cycle in **B**). While a nighttime assimilation has the theoretical advantage of utilizing a more direct physical link between w_2 and w_1 , it leads to problems where the nighttime model Jacobians differ from those across the full diurnal cycle. For example, due to the absence of transpiring vegetation in Spain and Algeria in Figure 7, w_1 is only very weakly

influenced by w_2 over the full diurnal cycle, however, there is a short-lived stronger sensitivity at night, which generates w_2 analysis increments from the w_1 observations. These w_2 increments do not influence the w_1 forecast for the next day (this suggests that the w_1 observation increments were not caused by w_2 errors), and since the w_1 observation increments are mostly monotonic, a large and likely erroneous net w_2 increment is generated over time. For the EKF this situation was exacerbated in this study by the inflation of **B** in these areas, and the (simplistic) error correlations used here will be refined in the future.

[38] It was assumed here that the background error standard deviations for w_1 and w_2 were both equal to the observation error standard deviation. However, in reality **b**₂₂ will be lower than **b**₁₁, since w_1 has more rapid dynamics and is more susceptible to forcing errors. Muñoz Sabater *et al.* [2007] compared soil moisture from ISBA forced with observations to in situ data from 2001 to 2004, and obtained a RMS error of 0.07 v/v for w_1 and 0.03 for w_2 (the contrast in the errors would likely be enhanced by the use of NWP forcing). The use of a smaller background error for w_2 would reduce some of the excessive model increments obtained in this study. Additionally, **Q** was chosen here in an attempt to generate stationary **B** (within the assumed structure of **Q**, proportional to $(w_{fc} - w_{wilt})$), however such a value could not be found across the entire European domain. With the chosen **Q**, the background error grew rapidly in dry and sparsely vegetated regions, including much of North Africa and Spain (resulting in large increments after several weeks in Figure 7). This suggests that **Q** should be lower in these regions. Intuitively, this is sensible: since there is little transpiration, w_2 does not vary greatly, and excepting a precipitation forcing error, the additive forecast error (**Q**) should be low compared to locations with substantial transpiration.

5. Conclusion

[39] This work is the first continental scale study to assimilate remotely sensed near-surface soil moisture into the ISBA model, and it is also the first study to contrast the assimilation of remotely sensed soil moisture using dynamic and static model error covariances. It is demonstrated that useful increments to the total soil layer in ISBA can be generated from near-surface soil moisture observations, in this case derived from AMSR-E. The spatially averaged net monthly increment for the EKF over ALADIN's European domain was $0.025 \text{ m}^3 \text{ m}^{-3}$, using approximately equal model and observation soil moisture errors. The assimilation was performed over July 2006, using both an EKF and a SEKF (in which the background error at the time of each analysis was assumed constant). While the Kalman gain terms for the SEKF and the EKF are determined by different physical processes, their resultant soil moisture analyses are similar (recall that horizontal error correlations were neglected in this study). Since performing the analysis in an off-line environment makes it computationally feasible, the EKF is suggested for future work with ISBA, although this study suggests that the SEKF provides an acceptable approximation (which is cheaper to compute and easier to implement). The difference between the two may well increase over a longer time period, particularly since the one-month period used in this experi-

ment is only three times the 10-day timescale of w_2 , and subsequently \mathbf{b}_{22} . The model itself is used as the observation operator, which in combination with the increased height of the atmospheric forcing in SURFEX enables the assimilation of screen-level observations, as demonstrated by Mahfouf *et al.* [2009]. The next stage of this work will be to investigate whether the EKF assimilation of remotely sensed soil moisture can be usefully combined with that of screen-level observations.

[40] While the focus here is on the design of the assimilation, the analyses are ultimately limited by the quality of the ingested data. In this experiment it was necessary to remove the seasonal bias between the AMSR-E and ALADIN soil moisture before using CDF-matching to rescale the AMSR-E data to the model's climatology. The profound difference in the soil analyses generated by including and excluding this seasonal correction was far greater than the difference between the EKF and the SEKF. This highlights the importance of the observation rescaling technique to soil moisture data assimilation. The relatively short period of available data (1 year), combined with the nonstationarity of the model-observation bias presented particular difficulties in rescaling the observations in this study. The issue of obtaining sufficient data to sample the model-observation climatology for rescaling presents a serious challenge for land-surface assimilation, particularly within NWP modeling, where frequent model changes are made, but also for the broader land surface community when using data from the early years of satellite missions.

Appendix A: Forecast Model as the Observation Operator

[41] The classic formulation for the EKF assimilation of near-surface soil moisture is to make the model update at the observation time, and use an observation operator of $\mathcal{H} = (1 \ 0)$ (for the state vector, $\mathbf{x} = (w_1 \ w_2)^T$, used here). To allow for the eventual assimilation of both near-surface soil moisture and screen-level atmospheric observations, a 6-h model forecast of the observation equivalent has been used for the observation operator in this study, with the analysis made 6 h before the observation time. It is shown below that for the assimilation of near-surface soil moisture this approach differs from the classic EKF only in the timing of the addition of \mathbf{Q} .

[42] For an observation at $t = 24$, the classic observation operator is written $\mathcal{H}(\mathbf{x}_{24}^b) = (1 \ 0) (\mathbf{x}_{24}^b)$, while the present version is $\mathcal{H}(\mathbf{x}_{18}^b) = (1 \ 0) \mathcal{M}_{18 \rightarrow 24}(\mathbf{x}_{18}^b)$. In both cases the result is the forecast w_1 at $t = 24$, $w_{1,24}$. \mathbf{Q} is neglected for the time being, and \mathbf{B} is expressed as a function of its value at time 0. This gives equations (A1) and (A2) for the classic and current EKF, respectively:

$$\mathbf{x}_{24}^a - \mathbf{x}_{24}^b = \mathbf{M}_{0 \rightarrow 24} \mathbf{B}_0 \mathbf{M}_{0 \rightarrow 24}^T \hat{\mathbf{H}}^T \cdot \left(\hat{\mathbf{H}} \mathbf{M}_{0 \rightarrow 24} \mathbf{B}_0 \mathbf{M}_{0 \rightarrow 24}^T \hat{\mathbf{H}}^T + \mathbf{R} \right)^{-1} (y_{24}^o - w_{1,24}) \quad (\text{A1})$$

$$\mathbf{x}_{18}^a - \mathbf{x}_{18}^b = \mathbf{M}_{0 \rightarrow 18} \mathbf{B}_0 \mathbf{M}_{0 \rightarrow 18}^T \mathbf{H}^T \cdot \left(\mathbf{H} \mathbf{M}_{0 \rightarrow 18} \mathbf{B}_0 \mathbf{M}_{0 \rightarrow 18}^T \mathbf{H}^T + \mathbf{R} \right)^{-1} (y_{24}^o - w_{1,24}) \quad (\text{A2})$$

Applying $\mathbf{M}_{18 \rightarrow 24}$ to equation (A2) carries it forward 6 h, giving:

$$\begin{aligned} \mathbf{M}_{18 \rightarrow 24} (\mathbf{x}_{18}^a - \mathbf{x}_{18}^b) &\simeq \mathbf{x}_{24}^a - \mathbf{x}_{24}^b \\ &= \mathbf{M}_{0 \rightarrow 24} \mathbf{B}_0 \mathbf{M}_{0 \rightarrow 18}^T \mathbf{H}^T (\mathbf{H} \mathbf{M}_{0 \rightarrow 18} \mathbf{B}_0 \mathbf{M}_{0 \rightarrow 18}^T \mathbf{H}^T + \mathbf{R})^{-1} (y_{24}^o - w_{1,24}) \end{aligned} \quad (\text{A3})$$

Substituting $\mathbf{H} = \hat{\mathbf{H}} \mathbf{M}_{18 \rightarrow 24}$ into equation (A3) produces equation (A1), hence the two forms of the EKF are equivalent if \mathbf{Q} is neglected.

[43] If \mathbf{Q} is included, equations (A1) and (A3) become, respectively:

$$\begin{aligned} \mathbf{x}_{24}^a - \mathbf{x}_{24}^b &= \left(\mathbf{M}_{0 \rightarrow 24} \mathbf{B}_0 \mathbf{M}_{0 \rightarrow 24}^T \hat{\mathbf{H}}^T + \mathbf{Q} \hat{\mathbf{H}}^T \right) \\ &\quad \times \left(\hat{\mathbf{H}} \mathbf{M}_{0 \rightarrow 24} \mathbf{B}_0 \mathbf{M}_{0 \rightarrow 24}^T \hat{\mathbf{H}}^T + \hat{\mathbf{H}} \mathbf{Q} \hat{\mathbf{H}}^T + \mathbf{R} \right)^{-1} \\ &\quad \cdot (y_{24}^o - w_{1,24}) \end{aligned} \quad (\text{A4})$$

and

$$\begin{aligned} \mathbf{x}_{24}^a - \mathbf{x}_{24}^b &= \left(\mathbf{M}_{0 \rightarrow 24} \mathbf{B}_0 \mathbf{M}_{0 \rightarrow 18}^T \mathbf{H}^T + \mathbf{M}_{18 \rightarrow 24} \mathbf{Q} \mathbf{H}^T \right) \\ &\quad \times \left(\mathbf{H} \mathbf{M}_{0 \rightarrow 18} \mathbf{B}_0 \mathbf{M}_{0 \rightarrow 18}^T \mathbf{H}^T + \mathbf{H} \mathbf{Q} \mathbf{H}^T + \mathbf{R} \right)^{-1} (y_{24}^o - w_{1,24}) \\ &= \left(\mathbf{M}_{0 \rightarrow 24} \mathbf{B}_0 \mathbf{M}_{0 \rightarrow 24}^T \hat{\mathbf{H}}^T + \mathbf{M}_{18 \rightarrow 24} \mathbf{Q} \mathbf{M}_{18 \rightarrow 24}^T \hat{\mathbf{H}}^T \right) \\ &\quad \times \left(\hat{\mathbf{H}} \mathbf{M}_{0 \rightarrow 24} \mathbf{B}_0 \mathbf{M}_{0 \rightarrow 24}^T \hat{\mathbf{H}}^T + \hat{\mathbf{H}} \mathbf{M}_{18 \rightarrow 24} \mathbf{Q} \mathbf{M}_{18 \rightarrow 24}^T \hat{\mathbf{H}}^T + \mathbf{R} \right)^{-1} \\ &\quad \cdot (y_{24}^o - w_{1,24}) \end{aligned} \quad (\text{A5})$$

[44] The difference between equations (A4) and (A5) is in the timing of the addition of \mathbf{Q} . In the latter \mathbf{Q} is added to \mathbf{B} at $t = 18$, and \mathbf{B} is evolved forward 6 h before the Kalman gain is calculated. This result has been confirmed by comparing the analyses generated by the two methods (using the same \mathbf{Q} , \mathbf{R} , and initial \mathbf{B}). The differences are limited to the magnitude of the analysis increments, with the increments being larger (yet showing the same spatial pattern) when the model is used as the observation operator.

[45] **Acknowledgments.** The authors thank Patrick Le Moigne and Valéry Masson for assistance with technical aspects of SURFEX, Karim Bergaoui for coding the first version of the assimilation system, and Thomas Holmes for providing the AMSR-E soil moisture data, together with advice regarding its use. We also acknowledge three anonymous reviewers for their comments and insights. Clara Draper is funded by an Australian Postgraduate Scholarship, and an eWater CRC top-up scholarship, and this work was accomplished with support from the Fondation de Coopération Scientifique Sciences et Technologies pour l'Aéronautique et l'Espace, and a University of Melbourne PORES Scholarship.

References

- Baker, R., B. Lynn, A. Boone, W.-K. Tao, and J. Simpson (2001), The influence of soil moisture, coastline curvature, and land-breeze circulations on sea-breeze initiated precipitation, *J. Hydrometeorol.*, 2(2), 193–211.
- Balsamo, G., F. Bouyssel, and J. Noilhan (2004), A simplified bi-dimensional variational analysis of soil moisture from screen-level observations in a mesoscale numerical weather-prediction model, *Q. J. R. Meteorol. Soc.*, 130(598A), 895–915, doi:10.1256/qj.02.215.
- Balsamo, G., J.-F. Mahfouf, S. Bélair, and G. Deblonde (2007), A land data assimilation system for soil moisture and temperature: An information content study, *J. Hydrometeorol.*, 8(6), 1225–1242, doi:10.1175/2007JHM819.1.
- Bélair, S., L. Crevier, J. Mailhot, B. Bilodeau, and Y. Delage (2003), Operational implementation of the ISBA land surface scheme in the

- Canadian regional weather forecast model. Part I: Warm season results, *J. Hydrometeorol.*, 4(2), 352–370.
- Bouttier, F., J.-F. Mahfouf, and J. Noilhan (1993), Sequential assimilation of soil moisture from atmospheric low-level parameters. Part II: Implementation in a mesoscale model, *J. Appl. Meteorol.*, 32, 1352–1364.
- Crow, W., and X. Zhan (2007), Continental-scale evaluation of remotely sensed soil moisture products, *IEEE Geosci. Remote Sens. Lett.*, 4(3), 451–455, doi:10.1109/LGRS.2007.896533.
- Deardorff, J. (1977), A parameterization of the ground surface moisture content for use in atmospheric prediction models, *J. Appl. Meteorol.*, 16(11), 1182–1185.
- de Jeu, R., W. Wagner, T. Holmes, A. Dolman, N. van de Giesen, and J. Friesen (2008), Global soil moisture patterns observed by space borne radiometers and scatterometers, *Surv. Geophys.*, 29, 399–420.
- Draper, C., and G. Mills (2008), The atmospheric water balance over the semiarid Murray-Darling River basin, *J. Hydrometeorol.*, 9(3), 521–534, doi:10.1175/2007JHM889.1.
- Draper, C., J. Walker, P. Steinle, R. de Jeu, and T. Holmes (2009), An evaluation of AMSR-E derived soil moisture over Australia, *Remote Sens. Environ.*, 113, 703–710, doi:10.1016/j.rse.2008.11.011.
- Drusch, M. (2007), Initializing numerical weather prediction models with satellite-derived surface soil moisture: Data assimilation experiments with ECMWF's Integrated Forecast System and the TMI soil moisture data set, *J. Geophys. Res.*, 112, D03102, doi:10.1029/2006JD007478.
- Drusch, M., and P. Viterbo (2007), Assimilation of screen-level variables in ECMWF's Integrated Forecast System: A study on the impact on the forecast quality and analyzed soil moisture, *Mon. Weather Rev.*, 135(2), 300–314, doi:10.1175/MWR3309.1.
- Drusch, M., E. Wood, and H. Gao (2005), Observation operators for the direct assimilation of TRMM Microwave Imager retrieved soil moisture, *Geophys. Res. Lett.*, 32, L15403, doi:10.1029/2005GL023623.
- Entekhabi, D., et al. (2004), The hydrosphere state (Hydros) satellite mission: An Earth system pathfinder for global mapping of soil moisture and land freeze/thaw, *IEEE Trans. Geosci. Remote Sens.*, 42(10), 2184–2195, doi:10.1109/TGRS.2004.834631.
- Fischer, E., S. Seneviratne, P. Vidale, D. Luthi, and C. Schär (2007), Soil moisture—atmosphere interactions during the 2003 European summer heat wave, *J. Clim.*, 20(20), 5081–5099, doi:10.1175/JCLI4288.1.
- Giard, D., and E. Bazile (2000), Implementation of a new assimilation scheme for soil and surface variables in a global NWP model, *Mon. Weather Rev.*, 128(4), 997–1015.
- Habets, F., et al. (2008), The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France, *J. Geophys. Res.*, 113, D06113, doi:10.1029/2007JD008548.
- Hess, H. (2001), Assimilation of screen-level observations by variational soil moisture analysis, *Meteorol. Atmos. Phys.*, 77, 145–154, doi:10.1007/s007030170023.
- Kerr, Y., P. Waldteufel, J.-P. Wigneron, J. Martinuzzi, J. Font, and M. Berger (2001), Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission, *IEEE Trans. Geosci. Remote Sens.*, 39(8), 1729–1735, doi:10.1109/36.942551.
- Li, L., E. Njoku, E. Im, P. Chang, and K. Germain (2004), A preliminary survey of radio-frequency interference over the US in Aqua AMSR-E data, *IEEE Trans. Geosci. Remote Sens.*, 42(2), 380–390, doi:10.1109/TGRS.2003.817195.
- Mahfouf, J.-F., K. Bergaoui, C. Draper, C. Bouysse, F. Taillefer, and L. Taseva (2009), A comparison of two off-line soil analysis schemes for assimilation of screen-level observations, *J. Geophys. Res.*, 114, D08105, doi:10.1029/2008JD011077.
- McCabe, M., E. Wood, and H. Gao (2005), Initial soil moisture retrievals from AMSR-E: Multiscale comparison using in situ data and rainfall patterns over Iowa, *Geophys. Res. Lett.*, 32, L06403, doi:10.1029/2004GL021222.
- Muñoz Sabater, J., L. Jarlan, J.-C. Calvet, F. Bouysse, and P. de Rosnay (2007), From near-surface to root-zone soil moisture using different assimilation techniques, *J. Hydrometeorol.*, 8(2), 194–206, doi:10.1175/JHM571.1.
- Ni-Meister, W., P. Houser, and J. Walker (2006), Soil moisture initialization for climate prediction: Assimilation of Scanning Multifrequency Microwave Radiometer soil moisture data into a land surface model, *J. Geophys. Res.*, 111, D20102, doi:10.1029/2006JD007190.
- Njoku, E., T. Jackson, V. Lakshmi, T. Chan, and S. Nghiem (2003), Soil moisture retrieval from AMSR-E, *IEEE Trans. Geosci. Remote Sens.*, 41(2), 215–229.
- Njoku, E., P. Ashcroft, T. Chan, and L. Li (2005), Global survey of statistics of radio-frequency interference in AMSR-E land observations, *IEEE Trans. Geosci. Remote Sens.*, 43(5), 938–947, doi:10.1109/TGRS.2004.837507.
- Noilhan, J., and J.-F. Mahfouf (1996), The ISBA land surface parameterisation scheme, *Global Planet. Change*, 13, 145–159.
- Owe, M., R. de Jeu, and J. Walker (2001), A methodology for surface soil moisture and vegetation optical depth retrieval using the Microwave Polarization Difference Index, *IEEE Trans. Geosci. Remote Sens.*, 39(8), 1643–1654, doi:10.1109/36.942542.
- Owe, M., R. de Jeu, and T. Holmes (2007), Multisensor historical climatology of satellite-derived global land surface moisture, *J. Geophys. Res.*, 113, F01002, doi:10.1029/2007JF000769.
- Quintana-Seguí, P., P. Le Moigne, Y. Durand, E. Martin, F. Habets, M. Baillon, C. Canellas, L. Franchisteguy, and S. Morel (2008), Analysis of near-surface atmospheric variables: Validation of the SAFRAN analysis over France, *J. Appl. Meteorol. Climatol.*, 47(1), 92–107, doi:10.1175/2007JAMC1636.1.
- Reichle, R., and R. Koster (2004), Bias reduction in short records of satellite soil moisture, *Geophys. Res. Lett.*, 31, L19501, doi:10.1029/2004GL020938.
- Reichle, R., D. McLaughlin, and D. Entekhabi (2001), Variational data assimilation of microwave radiobrightness observations for land surface hydrology applications, *IEEE Trans. Geosci. Remote Sens.*, 39(8), 1708–1718, doi:10.1109/36.942549.
- Reichle, R., J. Walker, R. Koster, and P. Houser (2002), Extended versus ensemble Kalman filtering for land data assimilation, *J. Hydrometeorol.*, 3(6), 728–740.
- Reichle, R., R. Koster, J. Dong, and A. Berg (2004), Global soil moisture from satellite observations, land surface models, and ground data: Implications for data assimilation, *J. Hydrometeorol.*, 5(3), 430–442.
- Reichle, R., R. Koster, P. Liu, S. Mahanama, E. Njoku, and M. Owe (2007), Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR), *J. Geophys. Res.*, 112, D09108, doi:10.1029/2006JD008033.
- Rüdiger, C., J.-C. Calvet, J.-F. Mahfouf, L. Jarlan, G. Balsamo, and J. Muñoz Sabater (2007), Assimilation of land surface variables for model initialisation at Météo-France, in *Proceedings of MODSIM 2007 International Congress on Modelling and Simulation*, pp. 1716–1722, Modell. and Simul. Soc. of Aust. and N. Z., Canberra, Australia.
- Rüdiger, C., J.-C. Calvet, C. Gruhier, T. Holmes, R. de Jeu, and W. Wagner (2009), An intercomparison of ERS-Scat and AMSR-E soil moisture observations with model simulations over France, *J. Hydrometeorol.*, 10(2), 431–447, doi:10.1175/2008JHM997.1.
- Scipal, K., M. Drusch, and W. Wagner (2008), Assimilation of a ERS scatterometer derived soil moisture index in the ECMWF numerical weather prediction system, *Adv. Water Resour.*, 31(8), 1101–1112, doi:10.1016/j.advwatres.2008.04.013.
- Seuffert, G., G. Wilker, P. Viterbo, M. Drusch, and J.-F. Mahfouf (2004), The usage of screen-level parameters and microwave brightness temperature for soil moisture analysis, *J. Hydrometeorol.*, 5(3), 516–531.
- Wagner, W., V. Naeimi, K. Scipal, R. de Jeu, and J. Martínez-Fernández (2007), Soil moisture from operational meteorological satellites, *Hydrogeol. J.*, 15(1), 121–131, doi:10.1007/s10040-006-0104-6.
- Walker, J., and P. Houser (2001), A methodology for initializing soil moisture in a global climate model: Assimilation of near surface soil moisture observations, *J. Geophys. Res.*, 106(D11), 11,761–11,774.
- Zhang, H., and C. S. Frederiksen (2003), Local and nonlocal impacts of soil moisture initialization on AGCM seasonal forecasts: A model sensitivity study, *J. Clim.*, 16(13), 2117–2137.

C. S. Draper and J. P. Walker, Department of Civil and Environmental Engineering, University of Melbourne, Parkville, Vic 3010, Australia. (c.draper@civenv.unimelb.edu.au; j.walker@unimelb.edu.au)
J.-F. Mahfouf, Météo France, CNRS, CNRM/GAME, 42 avenue Gaspard Coriolis, F-31057 Toulouse CEDEX, France. (jean-francois.mahfouf@

Other Key Reference Papers for Prof. Jeffrey's Lecture:

- [1] Walker, J. P., P. R. Houser (2001), A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations, *Journal of Geophysical Research*, 106: 11761 - 11744
- [2] Wenge, N., J. P. Walker, and P. R. Houser (2005), Soil moisture initialization for climate prediction: Characterization of model and observation errors, *Journal of Geophysical Research*, doi:10.1029/2004JD005745
- [3] S. Peischl, J.P. Walker, et al. (2009), Towards validation of SMOS using airborne and ground data over the Murrumbidgee Catchment, 18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July.
- [4] Mladenova, I., et al. (2010), Validation of the ASAR global monitoring mode soil moisture product using the NAFE'05 data set, *IEEE Transactions on Geoscience and Remote Sensing*
- [5] Walker, J. P., P. R. Houser (2004), Requirements of a global near-surface soil moisture satellite
- [6] mission: accuracy, repeat time, and spatial resolution, *Advances in Water Resources* 27: 785 -801
- [7] Merlin O J P Walker et al (2008) The NAFE'06 data set: Towards soil moisture retrieval at intermediate resolution *Advances in Water Resources* 31: 1444

Corn-yield estimation through assimilation of remotely sensed data into the CSM-CERES-Maize model

HONGLIANG FANG*†, SHUNLIN LIANG†, GERRIT HOOGENBOOM‡, JOHN TEASDALE§ and MICHEL CAVIGELLI§

†Department of Geography, University of Maryland, College Park, MD 20742, USA

‡Department of Biological and Agricultural Engineering, University of Georgia, Griffin, GA 30223, USA

§Sustainable Agricultural Systems Lab, USDA-ARS, Beltsville, MD 20705, USA

(Received 6 January 2006; in final form 19 April 2007)

One of the applications of crop simulation models is to estimate crop yield during the current growing season. Several studies have tried to integrate crop simulation models with remotely sensed data through data-assimilation methods. This approach has the advantage of allowing reinitialization of model parameters with remotely sensed observations to improve model performance. In this study, the Cropping System Model-CERES-Maize was integrated with the Moderate Resolution Imaging Spectroradiometer (MODIS) leaf area index (LAI) products for estimating corn yield in the state of Indiana, USA. This procedure, inversion of crop simulation model, facilitates several different user input modes and outputs a series of agronomic and biophysical parameters, including crop yield. The estimated corn yield in 2000 compared reasonably well with the US Department of Agriculture National Agricultural Statistics Service statistics for most counties. Using the seasonal LAI in the optimization procedure produced the best results compared with only the green-up LAIs or the highest LAI values. Planting, emergence and maturation dates, and N fertilizer application rates were also estimated at a regional level. Further studies will include investigating model uncertainties and using other MODIS products, such as the enhanced vegetation index.

1. Introduction

Advance information on crop yield during the crop growing season is vital for effective crop management and for national food security policy. Agricultural survey is a reliable way to estimate regional crop yields by sampling field measurements of standing crops. However, this method is time-consuming and costly, and the results are not available until after harvest. In the last three decades, satellite remote sensing data have been used to estimate crop yields over large areas, as these methods are more cost-effective and more timely than traditional survey procedures (e.g. LACIE, MacDonald and Hall 1980).

Earlier studies were mostly based on empirical regression methods that relate crop yield to remotely sensed surface reflectance and their combinations (i.e. vegetation indices). These relationships could be described with linear, cubic polynomial, or exponential regression (Jiang *et al.* 2003). For example, the Normalized Difference

*Corresponding author. Email: hfang@geog.umd.edu

Vegetation Index (NDVI) is strongly linked to vegetation condition and plant biomass, and thus has been used to estimate wheat yield during the growing season (Tucker *et al.* 1980). This method is essentially a statistical model and cannot predict the time-dependent processes of crop growth. Besides, the relationship between yield and NDVI may not be adequate under extreme weather conditions.

Mathematical crop growth models simulate fundamental processes such as photosynthesis, respiration, biomass partitioning, and water and nitrogen transfers (Baret *et al.* 2000). They have the advantage of enabling researchers to evaluate a wide array of alternatives and to assemble processes in an integrated package. They can help farmers make crop-management decisions (Tsuji *et al.* 1998). The mechanical crop growth models can simulate the dynamics of Leaf Area Index (LAI) and other structural properties of the crop fields (e.g. height and biomass). The combination of remote sensing and crop growth simulation models has become increasingly recognized as a promising approach for monitoring growth and estimating yield (Bouman 1992, Moulin *et al.* 1998, Baret *et al.* 2000, Plummer 2000, Doraiswamy *et al.* 2003).

The use of crop models is often limited by uncertainties in their input parameters such as soil conditions, sowing date, planting density and initial field conditions. Except in some controlled experimental fields, many of these parameters are poorly known. Remote sensing can play a critical role in helping identify the field and crop status from estimated biophysical parameters (Clevers and Leeuwen 1996). Remote sensing data, therefore, can be assimilated with crop growth models to improve their overall performance (Maas 1988a, Delécolle *et al.* 1992, Moulin *et al.* 1998, Baret *et al.* 2000, Plummer 2000, Doraiswamy *et al.* 2004).

In this study, we developed a practical procedure using the data assimilation method to predict crop yield at the regional scale from Moderate Resolution Imaging Spectroradiometer (MODIS) data. This method, Inversion of Crop Simulation Model (ICSM), contains several different user input modes that allow different input choices. ICSM outputs include agronomic variables (yield, planting, emergence, and maturation dates) and biophysical parameters (e.g. LAI). It also provides the flexibility to incorporate different Earth Observing System (EOS) data products and is capable of producing other diagnostic variables if necessary.

2. Existing studies and our new data assimilation schemes

2.1 Existing studies

Crop simulation models (CSMs) directly estimate biomass products, water and nitrogen balances, and crop yield through a deterministic scheme using ancillary data such as soil characteristics and climate variables (Williams *et al.* 1984, Tsuji *et al.* 1998, Engel *et al.* 1997, Baret *et al.* 2000). However, CSMs require many parameters and variables as inputs. Some of these factors are poorly known, and the model must be calibrated for particular crop conditions. Nevertheless, CSMs are also able to describe the dynamics of LAI and may be adapted to describe other structural and optical properties. They can thus be coupled with radiative transfer models (RTM) to simulate the radiometric field response. A comparison between simulated and observed radiometric data can be applied to correct CSM variables and parameters and force the CSM to run in a more realistic way.

Several assimilation schemes, of various degrees of complexity and integration, have been developed during the last 10 years (Moulin *et al.* 1998, Baret *et al.* 2000,

Plummer 2000, Doraiswamy *et al.* 2004). Various methods for integrating a crop growth model with remote sensing data were described by Maas (1988a, b) and were also reviewed by (Fischer *et al.* 1997, Moulin *et al.* 1998). Maas (1993) conducted a study to compare the results of calibrating a crop simulation model using LAI observations obtained from either field sampling or remote sensing. Winter wheat yield was more accurately predicted using remotely sensed LAI observations than using field-sampled LAI observations (Maas 1993). This difference appeared to result from the apparent ability of the remotely sensed LAI observations to better represent the photosynthetically active plant area in the crop canopy.

Bach *et al.* (2003) experimented with the coupling of a raster-based PROMET-V model with the radiative transfer model GeoSAIL to predict biomass and yield. In their study, LAI, fraction of brown leaves, and surface soil moisture were used as free variables; surface reflectance was used as the control variable. Their assimilation procedure produced improved biomass and yield results. Guérif and Duke (2000) combined the SUCROS crop model with the SAIL canopy reflectance model to obtain a more accurate estimation of sugar-beet yield. Ground measured reflectance was used to match the predicted reflectance. One limitation of their study was that many crop and soil parameters had to be obtained from field measurements. The SAIL model has also been integrated with the EPIC crop model to estimate the yield of spring wheat in North Dakota (Doraiswamy *et al.* 2003). Planting date is the only adjustable variable in their model. The estimated yields with the combined EPIC and SAIL models were mostly within 10% of the National Agricultural Statistical Service (NASS) reports. In their study, climate data were based on interpolation of weather station measurements, while the crop area mask was based on the 1-km AVHRR classification (Doraiswamy *et al.* 2003). NDVI was also calculated from AVHRR. Although the AVHRR data set is easily accessible, use of the set will compromise the precision of analyses owing to their outdated calibrations and the application of partial atmospheric corrections. This is one of the reasons why the latest MODIS products were used in our study.

In a more recent study, Doraiswamy *et al.* (2004) used a look-up table (LUT) method to estimate LAI from 250-m MODIS reflectance data. The crop-modelled LAI was adjusted to fit the MODIS simulated LAI by changing the data of planting, the time when maximum LAI was attained, and the beginning of leaf senescence. If MODIS LAI is used, this step, which involves intensive fieldwork and indoor retrieval, can be avoided.

Doraiswamy *et al.* (2004) used weather-station data, which are accurate and easily available. However, these data are spatially discrete, and spatial interpolation is indispensable for regional application. This motivates us to look into alternative weather data sources, such as the land data assimilation systems (LDAS) data (Mitchell *et al.* 2004). One of the primary goals of LDAS is to provide optimal estimates of land surface state initial conditions such as soil moisture and temperature at spatial resolutions down to 1 km. Therefore, LDAS data can be incorporated into our assimilation strategy.

In general, two main approaches have been used to couple or integrate satellite data with a crop model (Guérif and Duke 1998). The first approach uses remotely sensed estimates such as LAI, FPAR, surface soil moisture, or evapotranspiration to force or recalibrate parameters of the CSM using optimization techniques. The second approach couples CSM and RTM to simulate whole processes from canopy

functioning to above-canopy radiometric data. Some parameters of the coupled model are then directly recalibrated by finding the best agreement between simulated and observed remote sensing data. Our ICSM provides the option to trigger either approach. This study will mainly focus on the first approach. Details about the second approach are the focus of another study.

2.2 Our data-assimilation schemes

A general outline of our data assimilation strategy is given in figure 1. LAI simulated by the crop model is compared with MODIS LAI products in the optimization process. The ICSM reinitiates the input parameters based on the root-mean-square-error (RMSE) of the LAIs. The model produces crop yield and other parameters when a minimization threshold is satisfied. Several crop growth models (e.g. DSSAT, WOFOST, and EPIC) were examined for their ability to predict yields for regional assessments. The Decision Support System for Agrotechnology Transfer (DSSAT) model was selected because of its easiness for adaptation and reliable support. DSSAT consists of three components (Hoogenboom *et al.* 1999, Hoogenboom *et al.* 2004): (1) a Data Base Management System (DBMS) to enter, store, and retrieve the 'minimum data set' needed to calibrate, evaluate and use the crop models for solving problems; (2) a set of evaluated crop models for simulating processes and outcomes of genotype by environment interactions; and (3) an application program for analysing and displaying outcomes of long-term simulated agronomic experiments.

The Cropping System Model (CSM)-CERES (Crop-Environment Resource Synthesis)-Maize is one set of modes under the DSSAT shell that has different modules for different kinds of crops (Jones *et al.* 2003). The CSM-CERES-Maize model simulates growth, development, and yield of corns, taking into account the effects of genetics, weather, and soil conditions and crop management. The model offers the ability to evaluate options for increasing yield and water and nitrogen-use efficiency. This model enables users to input the biological and physical parameters

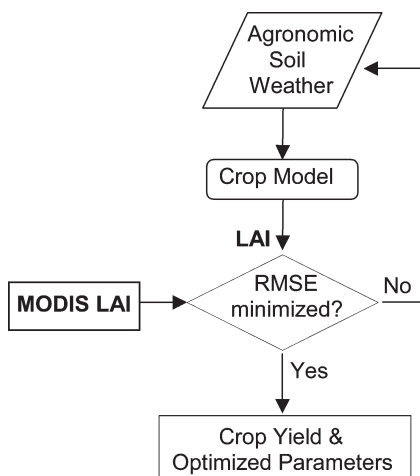


Figure 1. Flow chart of the inversion of crop-simulation models (ICSM) using MODIS LAI data.

and obtain user-specified objectives (Tsuji *et al.* 1994, Tsuji *et al.* 1998). The CSM-CERES-Maize model was modified in an inversion looping structure in ICSM to dynamically use both remotely sensed and simulated LAI data. ICSM evaluates the modelled LAI dynamics and recalibrates the crop model input parameters with the remotely sensed LAI. Besides the LAI dynamics, the CSM-CERES-Maize model also simulates the growth and development for a maize canopy as well as the water and nitrogen balances of the soil–vegetation–atmosphere system using a daily step.

The inversion models (ICSM) were first run with management records and treatments obtained from either agricultural statistics or local agronomists. Upon completion of the run, these dates when both simulated LAI and MODIS LAI were collected were identified. The simulated and MODIS LAI values were fitted to the model outputs. The residuals were analysed to determine certain input values (weather, genetics, soil, etc.) that were traced back to the input files. After removing ‘error’ input values and adjusting the parameters (determined through sensitivity studies), the final set of model input parameters was derived. The model was tuned again and was run to determine the final LAI and grain yield.

3. Data preparation

Our study area was the state of Indiana, USA (figure 2). The topography of Indiana is characterized by vast flat plains in the northern two-thirds of the state. In the

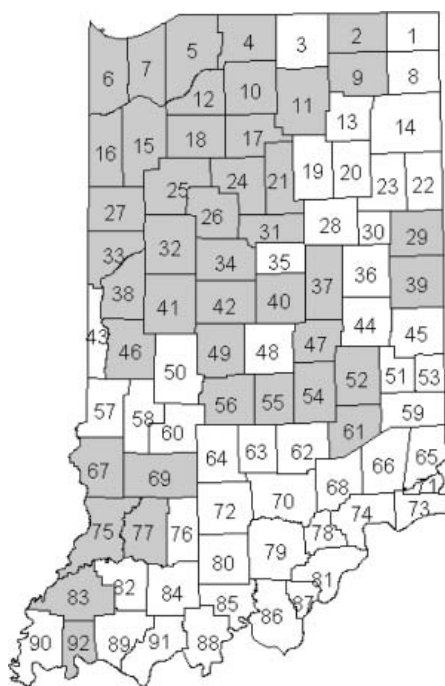


Figure 2. Administrative counties for Indiana (92 in total). The county number is created automatically from approximately north to south. The shaded counties (43 in total) are used in a corn-yield comparison with statistical data.

south, hills, ridges, and knolls abound. Land elevations range from 100 m above sea level in the south-west corner of the state to 382 m in east central Indiana (Indiana Office of the Commissioner of Agriculture 2003). Corn and soybean are the two dominant crops.

Crop-growth models need many ancillary inputs such as daily weather variables and soil characteristics in addition to crop genetics and management conditions. To run ICSM, a set of data was prepared in standard input files with specific data formats and conventions. Some are free variables, such as the planting date and nutrient applied. Their initial values were pre-determined but were optimized in the program. The environmental parameters, such as geographic positions, crop type, soil and weather conditions discussed below, are critical for obtaining accurate outputs and were determined before running the program.

3.1 *Crop-management variables*

Spatial information about the specific corn cultivar types was not obtained for the entire state, but Pioneer cultivars represent the overwhelming majority of corn cultivars planted in Indiana (personal communication with Dr WD Batchelor). In the CSM-CERES-Maize corn cultivar files (DSSAT 4.0), there are several generic corn hybrids labelled by growing degree days to maturity (PC0001, PC0002, PC0003, and PC0004). In the *a priori* sensitivity test (see §4.2), PC0003 was found to work well in Indiana.

Some variables are more critical for effective use of CSM than others, such as the initial conditions, planting date, and planting population and row spacing. They are treated as free variables that can be estimated during the data-assimilation process. They cover the likely variation over corn-fields in the study area. In practice, the number of free variables to be estimated is adjustable in ICSM. More experiments can be conducted to decide the best set of free variables. For corn, both plant population and row spacing may vary considerably (and be intercorrelated) over a large spatial domain, depending on the field conditions. Moreover, sowing dates may also vary greatly according to weather and farmer decisions. These variations have major effects on subsequent growth, as was shown by a CSM-CERES-Maize sensitivity analysis.

USDA NASS reports provide the average crop planting dates at the state level, but not specific to a county. Crop calendar models simulate crop growth incrementally, based on growing degree-days (or thermal units) for different types of crops and crop varieties. The current crop calendar models are initialized by the average start of season data derived from national crop reports (<http://www.pecad.fas.usda.gov/cropeplorer/datasources.cfm>). However, determining the start of the cropping season is still very empirical at the regional level. A more practical way is to initialize the crop models from start-of-season data derived from agrometeorological data or vegetation index (VI) data (Bethel and Doorn 1998). In this study, the initial planting date was set to the 7 May, DOY 128 when half of the corn has been planted over the state (1998–2002 average) (<http://www.nass.usda.gov/in/annbul/0304/04weather.html>), with one month variation. The crop emergence date was computed from the crop growth model. Due to spatial variability in the soil surface characteristics, the duration between planting and emergence was adjustable.

The appropriate supply of nutrients is essential for modern crop production. However, it is very difficult to generalize on fertilizers because of the diversity of

fertilizer sources used, the exact amount of nutrients that are applied, and the dates of application. The simplest nutrient to generalize would be nitrogen (N) because it is used by nearly all growers. Some growers apply fertilizers only prior to planting. Some farmers apply split applications with some prior to planting and some after initial crop establishment or after the crop has emerged. At the regional level, it is impossible to find a typical time of nitrogen application. In this study, the date of N application was set to the same date as the planting date, and hence had the same variation range. The amount of N applied varies for each individual field, and there is no information about how it is distributed across the state. Some of our initial experiments displayed a fertilization median at 170–185 kg ha⁻¹, which provides some basic information about fertilization practices in the area. We have also referred to the CERES default values for the initial nitrogen amount. In the automatic management situation, DSSAT uses a default value of 25 kg ha⁻¹ in 50% N deficit. The actual amounts applied by growers could be higher (Dr R. L. Nielsen, personal communication).

3.2 Environmental variables

3.2.1 Soil characteristics. Surface soil texture is one of the principal characteristics affecting the soil spectral reflectance (Guérif and Duke 2000). DSSAT soil input is mainly based on soil texture and soil depth. The information on soil texture spatial distribution is available from USDA Soil Geographic Database (STATSGO) (<http://www.ncgc.nrcs.usda.gov/branch/ssb/products/statsgo/index.html>). The spatial resolution was reprojected to 1 km. The STATSGO data were recoded based on soil texture and the depth of the main soil component. The predominant soil texture in the region is silt loam, with zones of sandy loam and sand in the northern region. Soil moisture in the surface fluctuates extensively, depending on soil texture, rainfall, runoff and soil evaporation. The soil moisture provided by STATSGO was used, since it is quite difficult to obtain the moisture information at a regional level.

3.2.2 Regional weather data. Daily weather data are required and must be available for the duration of the growing season. Three weather data sources were examined: the National Weather Service (NWS) station data (<http://lwf.ncdc.noaa.gov/oa/climate/stationlocator.html>), the National Centers for Environmental Prediction (NCEP) North American Regional Reanalysis (NARR) data (<http://www.emc.ncep.noaa.gov/mmb/rreanl/index.html>), and the North America Land Data Assimilation Systems (NLDAS) data (<http://ldas.gsfc.nasa.gov>). For the regional application of ICSM, weather data have to be in a spatially continuous grid format. More importantly, their spatial resolution must be sufficient to be able to simulate crop growth, development, and predict yield for the entire growing season. The NLDAS data were used in this study because they are spatially continuous, and their resolution is relatively higher (1/8 degree). NLDAS uses 3-hourly Eta data (Eta: North America Mesoscale model, from 1996 to the present), hourly GOES solar shortwave data, Stage II Doppler precipitation data, and daily rainfall gauge data. These data sets are combined to form hourly NLDAS forcing files. NLDAS data contain surface temperature, wind speed, downward shortwave flux, and precipitation that are required by DSSAT. They were aggregated into daily parameters for crop model inputs. An example of the NLDAS data is plotted in

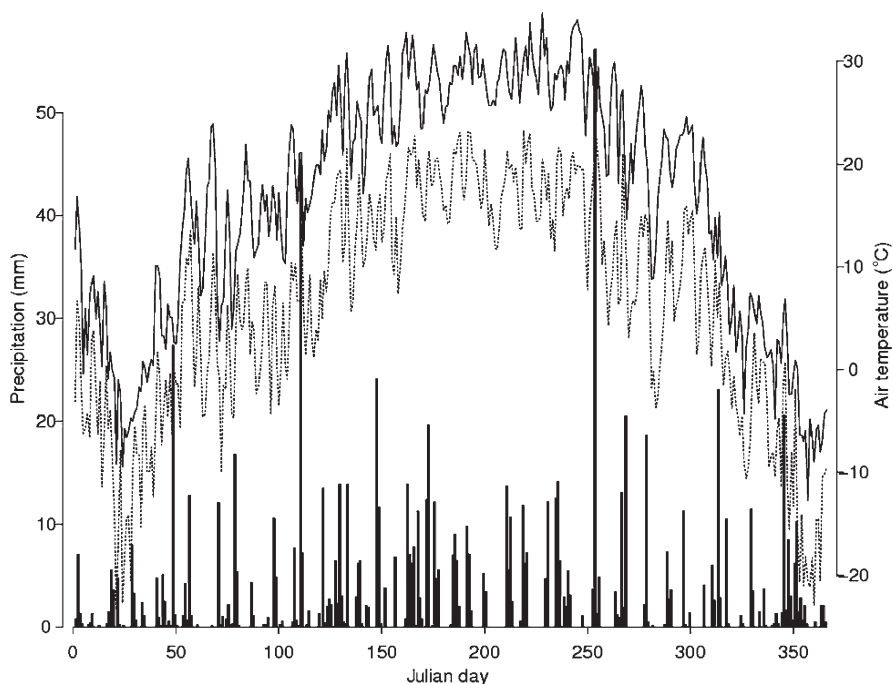


Figure 3. Example of the NLDAS daily precipitation (bars, left axis) and air temperature (lines, right axis) for Indiana in 2000. The solid and dashed lines show the maximum and minimum air temperatures, respectively.

figure 3. Normally, January is the coldest month, and July is the warmest month of the year. The daily precipitation varies widely over the year.

3.3 Remotely sensed data (LAI)

LAI is the key variable in the data-assimilation scheme. The comparison between simulated and observed LAI is applied to correct the parameters of the CSM-CERES-Maize in order to improve the performance of the model. The MODIS science team is producing LAI products globally every 8 days at a spatial resolution of 1 km (Justice *et al.* 1998, Myneni *et al.* 2002). The collection 4 MODIS LAI products are available for the general user community through the Earth Resources Observation System (EROS) Data Active Archive Center (DAAC) (<http://edcimswww.cr.usgs.gov/pub/imswelcome/>). The state of Indiana is covered by two tiles, h11v04 and h12v04.

Determination of the number of LAIs used in the optimization scheme needs to consider the LAI acquisition and data quality, the crop growth cycle, and computational demands. Guérif and Duke (2000) found that the best situation was when the data covered the whole period of LAI growth (including the highest values). Due to cloud contamination, spatially and temporally complete LAI is not available at the current stage. The retrieval index, which indicates the number of good retrievals relative to the total number of attempted pixels (Wang *et al.* 2001), was very low in the study area during the growing season due to clouds. For

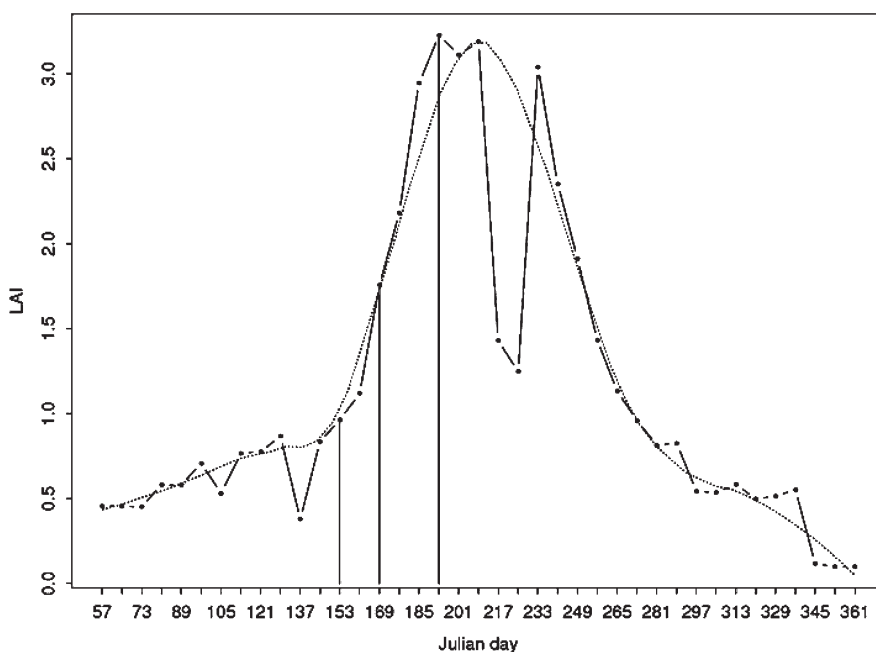


Figure 4. Mean LAI profile (solid) and curve after Savitzky–Golay filtering (dotted line) for corn grown in Indiana in 2000. The three segments show the period when LAI changes quickly (days 153, 169, and 193, respectively).

example, the retrieval indices for 4 August (DOY 217) and 12 August (DOY 225) are only 1.73% and 9.5%, respectively (figure 4). Non-vegetation and fill value pixels were excluded based on the QC (quality control) layer. LAI values derived from both the main radiative transfer method and empirical methods were used, excluding the pixels with geometry problems. The Savitzky–Golay (SG) filter (Savitzky and Golay 1964) was applied to smooth the annual LAI profile (figure 4). Figure 5 shows an example of LAI images before and after temporal filtering. Clearly, the filtered data (figure 5(b)) improved the pixels with poor retrievals and provided value-added information for crop simulation for some LAI gaps. Note that there are some forests in southern Indiana where the values for LAI are as high as 6.

MODIS LAI products contain data from before planting to after maturity. Technically, this data continuity allows a simulation to be started before planting. Nevertheless, only LAI for the growing season were used to accommodate the cyclic nature of ICSM and to avoid possible noises. The LAI series were selected according to the crop calendar (<http://www.nass.usda.gov/in/annbul/0304/04weather.html>), beginning on the date when only 5% of the crop had been planted until the date when 95% of the crop had been harvested (1998–2002 average) across the state. LAIs after maturity were not used due to the MODIS LAI uncertainties and the sensitivity of DSSAT models to LAIs of this period. For this study, three scenarios were evaluated: (1) S1, using 13 LAIs during the growing season (day 137, 145, . . . , 233); (2) S2, using 4 LAIs located at the plateau of the profile (day 185, 193, 201, and 209); and (3) S3, using three LAIs at the time when the crop was rapidly growing,

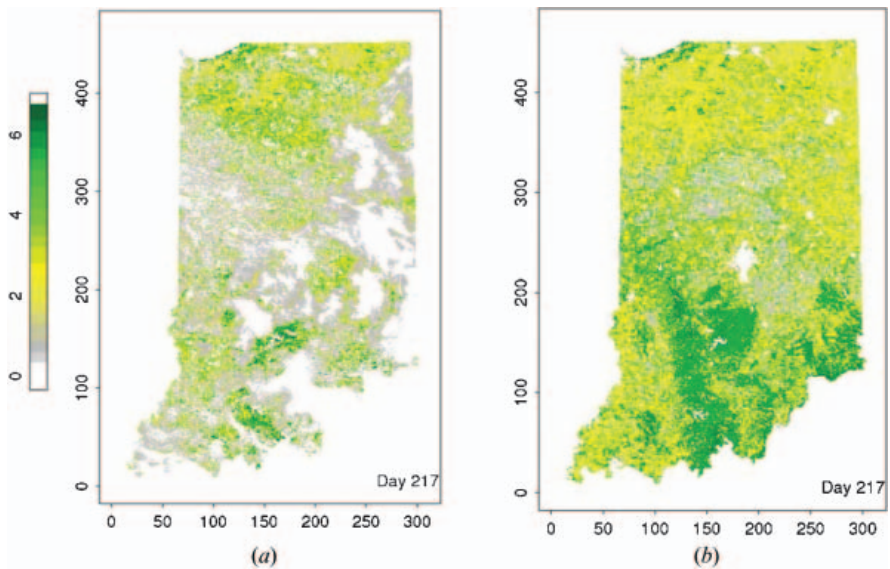


Figure 5. Image of MODIS LAI data (a) before and (b) after temporal filtering. Cloud contaminated areas in (a) have been filled in (b). Date: day 217, 2000.

and the LAI was increasing (days 153, 169, and 193). The latter two scenarios are to test the sensitivity of different LAI selections.

In addition to the above inputs, other field characteristics such as soil-analysis data, organic residue applications, environmental modifications, and harvest management were also necessary. These were determined through a series of sensitivity analyses done *a priori*. It is equally important to consult local practising agronomists. Some laboratory and field observational data obtained in similar studies (O'Neal *et al.* 2002), grown under similar conditions, were used as the initial values in a previous phase to adapt some model coefficients to the regional context. All data were transformed and reprojected to a 1-km resolution with a Universal Transverse Mercator (UTM) projection.

3.4 Pure corn pixels

The DSSAT family of models simulates grain cereal, legumes, root crops, and other crops separately. This poses a challenging task to separate different types of crops within a mixed pixel. High-spatial-resolution data (e.g. IKONOS, SPOT, and ETM+) are optimal for deriving inputs for crop growth models, but impractical for large-scale application. The USDA is producing cropland data layer (CDL) using Landsat TM and ETM+ imagery (<http://www.nass.usda.gov/research/Cropland/SARS1a.htm>) for some Mid-western states. The 30-m-resolution CDL data were upscaled to the MODIS LAI resolution (1 km). Figure 6 shows the percentage of corn in the 1-km resolution upscaled from CDL. Only pixels having more than 80% corn were processed using the CSM-CERES-Maize model. Overall, there are 871 'pure' corn pixels in 2000. If the corn percentage were lower than 80%, the results would be unacceptable due to the mixed pixels.

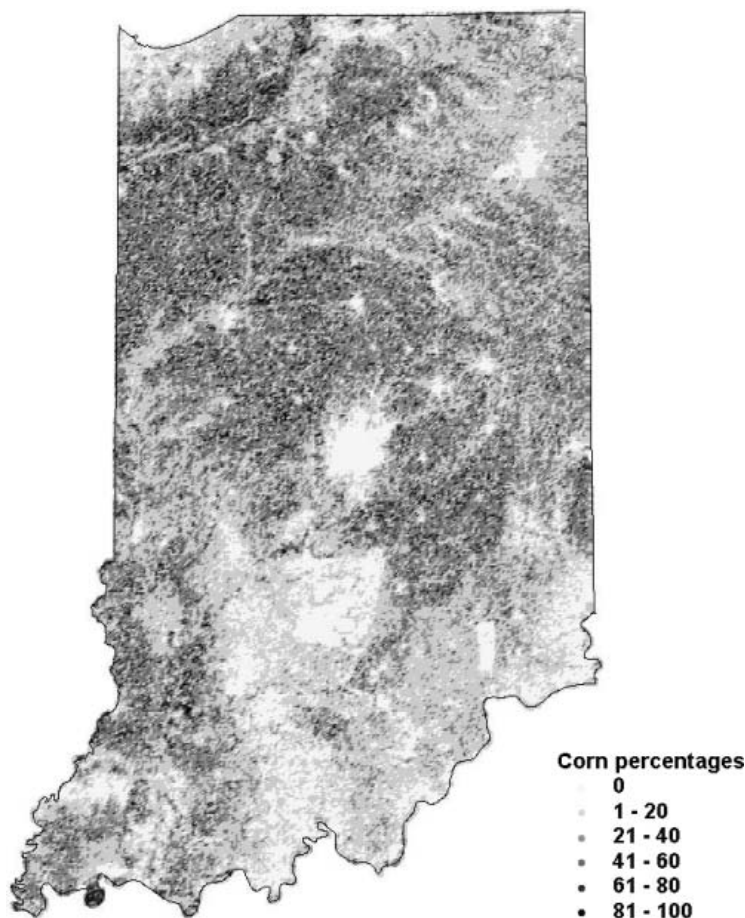


Figure 6. Corn-sowing percentage in Indiana at 1 km² resolution. Percentage values are calculated from the USDA NASS (2000) cropland data layer.

4. Inversion of the Crop Simulation Model (ICSM)

4.1 Merit function and optimization scheme

The inversion procedure of the data assimilation is similar to the 4D variational analysis in (Liang 2004). A general merit function J is constructed as:

$$J(X(t_0)) = [X(t_0) - X_b]^T B^{-1} [X(t_0) - X_b] + \sum_{i=1}^n [H_i(X(t_i)) - Y_i]^T R_i^{-1} [H_i(X(t_i)) - Y_i], \quad (1)$$

where X and Y are the input and output variables, respectively. X_b is the background value (or first guess), $X(t_0)$ is the observed value at time t_0 , H is the CSM model operator, and R and B are the observation and background error covariance matrices, respectively. The first term on the right of equation (1) is to force the optimal parameters as close as possible to input fields, and the second term is to adjust variables so that model outputs will be as close to the observations as possible.

To use equation (1) and fulfil the data-assimilation procedure, the uncertainties of input variables and output yield must be quantified. At the current stage, our objective was to minimize the difference between the measured and simulated LAI, and thus only the second term on the right of equation (1) was considered. Equation (1) takes a simplified form for this purpose:

$$J = \sum_{i=1}^n \text{abs}(\text{LAI}_S(t_i) - \text{LAI}_M(t_i)), \quad (2)$$

where $\text{LAI}_S(t_i)$, $\text{LAI}_M(t_i)$ are the simulated and measured LAI at time t_i , respectively. As stated previously, the pre-season and post-season values were excluded. Thus, the 8-day MODIS LAI was of major importance for this optimization process. Equation (2) was used to calculate the residual between the simulated and measured LAI. We tried to use the measured LAI as weights in a weighted difference equation. However, it brought some biases for lower MODIS LAI and caused higher yield deviations.

Many different types of optimization methods are available (Press *et al.* 1992). A multidimensional optimization algorithm adjusts free parameters until the merit function is minimized. In this study, the conjugate direction method (e.g. POWELL from Press *et al.* 1992) was used to force the four model parameters, planting date, planting population, row spacing, and amount of nitrogen applied. The conjugate direction method starts from a single initial position and conducts single line minimizations in each of the current p conjugate directions in order to arrive at the minimum for a given iteration. If the difference between the current and previous iterations minimum is below a user-defined tolerance, the program terminates. If the difference is greater than the tolerance, the conjugate directions are adjusted according to the vector between the two minima, and another iteration begins. Table 1 shows an example of a typical optimization process for one pixel. Column J is the merit variable calculated with different values at different iteration. The last row shows the final optimized values for that pixel when the maximum number of iteration or the minimum J is achieved.

4.2 Sensitivity studies

The current DSSAT model allows users to input different combinations of environmental variables and management practices. Thus, the sensitivity of the input parameters can be tested (Prévot *et al.* 2001). By doing this, users can determine the impact of not only crop types and management practice, but also soil quality and other environmental factors (O'Neal *et al.* 2002, Heinemann *et al.* 2002). A series of tests were designed and performed in the initial experiments of this study: (1) the number and values of initial parameters (vector X in equation (1)); (2) the number of the measured LAI values used (n in equations (1) and (2)); (3) the effect of uncertainties in the control variable (R in equation (1)); and (4) the issue of mixed pixels (analogous to R in equation (1)).

Among the many input variables, weather conditions, soil properties, crop and cultivar selection, planting date, plant spacing, and application of irrigation and fertilizer were found to be very important in determining final crop yield. Table 2 shows a sensitivity experiment for Jasper County, Indiana (No. 15 in figure 2). The initial planting date was set on 7 May (DOY 128). The plant density and row spacing were 6.5 plants m^{-2} and 80 cm, respectively. Other initial variables for the

Table 1. Example of the optimization process.

No. of iterations	Planting date	Planting population	Row spacing	Nitrogen amount	<i>J</i>
7	129	6.5	80	44	16.1331
8	129.2	6.5	80	44	16.4054
9	128.8	6.5	80	44	16.1331
23	128.6	6.5	80	44	15.9872
24	128.6	7.5	80	44	9 242 283
25	128.6	6.5	80	44	15.9872
43	128.6	7.5	80.8	44	12.5159
44	128.6	7.5	90	44	12 242 396
45	128.6	7.5	80.8	44	12.5159
63	128.6	7.5	90	48.1	12.2210
64	128.6	7.5	90	64.9	11.2910
65	128.6	7.5	90	92.1	10.4034
66	128.6	7.5	90	110.1	10.1295
67	128.6	7.5	90	139.3	9.9472
68	128.6	7.5	90	141.3	9.9410
69	128.6	7.5	90	144.4	9.9327
Optimized values	127.5	7.5	81.6	176.4	9.6731

The planting date, planting population, row spacing, and amount of nitrogen are optimized consequently based on the *J* value in equation (2). The last row shows the final optimized values.

county were set with the data discussed in §3. All 13 MODIS LAI observations over the growing season were used (scenario 1). The simulated yield (9723 kg ha⁻¹) was slightly higher than the USDA NASS data (9172 kg ha⁻¹). Various input values were changed to determine the impact on the simulated yield from the assimilated model. Changes in planting date, planting density, and MODIS LAI will all impact the final simulated yield. For example, with the same parameter setting, moving the planting date 5 days earlier caused a yield decrease of up to 57.1%. On the other hand, yield was relatively stable for different row spacings. We also experimented with some alternative management scenarios, such as cultivar selections and crop rotation. Management measures are important for crop growth, but they were kept constant and were not adjusted in the data assimilation process. It is rather difficult to select the planting date and density for crop models without knowledge of field crop type. By running the coupled inversion model, the range and uncertainties of the sowing date and density can be evaluated by choosing different input options.

The influence of MODIS LAI data was also tested. Decreasing the MODIS LAI will also lower the simulated yield, indicating that the MODIS LAI is another important parameter in this study. In this study, the number of input LAI values can

Table 2. Impact of input variables on the simulated yield in Jasper county, Indiana, 2000.

	Simulated values	Planting date (DOY 128)		Plant density (6.5 plants m ⁻²)		Row spacing (80 cm)		MODIS LAI	
Cases		-5	+5	-0.5	+0.5	-10	+10	-10%	+10%
Yield (kg ha ⁻¹)	9723	4173	11 926	10 046	9250	9745	9717	8301	10 495
Difference (%)		-57.1	22.7	3.3	-4.9	0.2	-0.1	-14.6	7.9

The first row lists parameters tested and their initial values. The last row shows the difference with the simulated yield.

be modified in the inversion. As described above, we designed three scenarios to test the effect of using all LAIs over the whole season, the highest LAIs, and the LAIs during the fastest growth period. Sensitivity studies showed that the nitrogen application, followed by the planting date, row spacing, and planting population were some of the variables that were mostly affected by the variation of the state variable (i.e. LAI). For example, a 10% LAI decrease can lower the retrieved nitrogen by up to 30%. It is acknowledged that these variables are interrelated, and compensations occur between various parameters. For example, a lower nitrogen application could be counterbalanced by an earlier sowing date, leading to a similar LAI profile.

A careful review of similar studies was conducted to identify the parameters that are critical for the ICSM tune-up process. For example, O'Neal *et al.* (2002) simulated the impact of precipitation on yield with CERES in an Indiana field. Kravchenko *et al.* (2000) used a multifractal analysis to study the relationship between topography and yield. Further, Kravchenko and Bullock (2000) analysed the yield–topography–soil relationships and found that soil properties and topography explained about 40% of yield variability. These works confirmed the importance of counting the soil and weather conditions in the ICSM model.

4.3 Various ICSM modes

In practice, ICSM was designed to cope with different input data availabilities, and thus has different running modes. The forward simulations were inherited from DSSAT and work for a single or multiple points simulations (modes 1 and 2, respectively). To run the forward simulation, the DSSAT 'minimum data set (MDS)' is required (<http://www.icasa.net/dssat/minimum.html>). To run the inversion model (ICSM) at a point level (mode 3), some observational LAI data are necessary in addition to the DSSAT MDS. With additional information available, such as the regional LDAS weather data in this study, another mode could be initiated to run ICSM at a regional level (mode 4). The ICSM has an additional option to integrate crop models with a canopy radiative transfer model (mode 5).

5. Results

In ICSM, corn LAI was used to adjust crop model input parameters by minimizing the difference between MODIS derived LAI and crop model simulated LAI. The model simulates the corn yield, assuming that the crop has no water stress and is free from any insect, pest, and disease effects.

Figure 7 compares the simulated corn yield and NASS yield data for selected counties in Indiana, 2000. Each of the 43 counties had at least 5 km² planted to corn (≥ 5 pixels) based on figure 2. Three different scenarios were compared. Within the state, the measured corn yields ranged from 7533 kg ha⁻¹ to 10 365 kg ha⁻¹. The best results were obtained from the first scenario (using 13 LAIs), when the simulated yield was about 7.9% higher than the NASS data (table 3). This phenomenon demonstrates that if more LAIs were used in the ICSM (higher n in equations (1) and (2)), the simulated LAIs could better capture the seasonal variation than if fewer LAIs were used. This finding is consistent with Guérif and Duke (2000). Therefore, the results from the first scenario (13 LAIs) were mainly discussed in the following parts. Results from the other two scenarios using four and three LAIs were 16.6% and 17.9% higher than the NASS yield, respectively (table 3). The standard

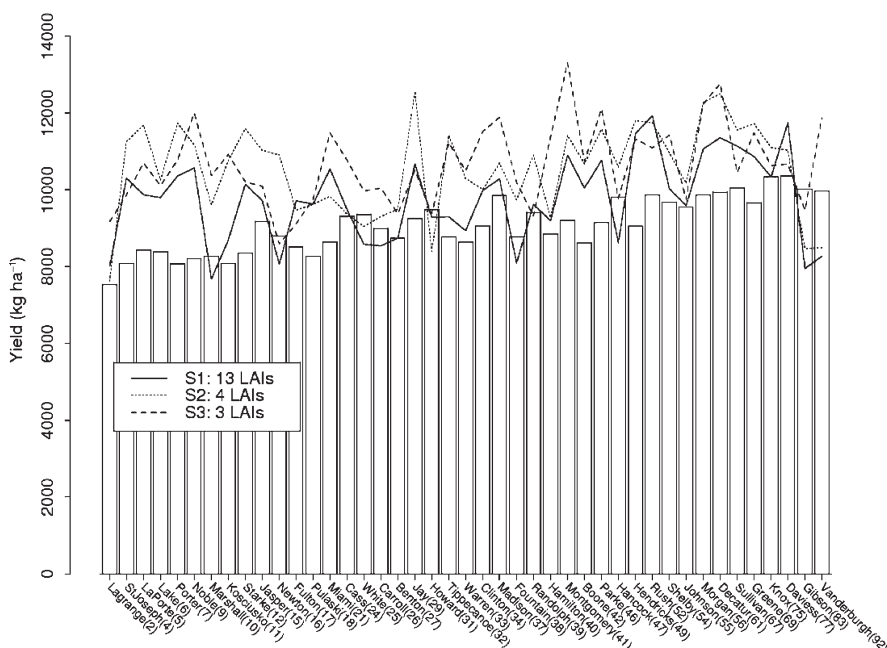


Figure 7. Comparison of simulated corn yield (lines) and NASS data (bars) for Indiana counties with ≥ 5 pixels planted to corn in 2000. Simulated results are from three different scenarios, using 13 LAIs (S1), the four highest LAIs (S2), and three LAIs during the maximum growth phase (S3). See figure 2 for the county numbers.

deviations of the simulated results from these two scenarios were also higher than the NASS data. This was not unexpected considering the variation of the input data used in this study (see LAI example in figure 10).

Figure 8 shows the spatial distribution of corn yield for all corn pixels (total of 871 in 67 counties) for the first scenario (S1: 13 LAIs). For all the 871 pixels, the

Table 3. Comparison of the estimated corn yield and NASS data for Indiana in 2000 (kg ha^{-1}).

	NASS data	Simulated yield					
		S1	Difference (%)	S2	Difference (%)	S3	Difference (%)
<i>Selected counties (43)</i>							
Mean	9080	9765	7.9	10 540	16.6	10 660	17.9
Standard deviation	700	1107		1186		1041	
<i>All 871 pixels in 67 counties</i>							
Mean	9189	9477	3.1	10 264	11.7	10 340	12.5
Standard deviation	644	1669		1878		1933	

The simulated mean values and standard deviations are calculated for the selected 43 counties (figure 2) and all 871 pixels (figure 8), respectively. Simulated yields are from three different schemes: S1, 13 LAIs; S2, the four highest LAIs; and S3: three LAIs during the maximum growth phase. Their differences from the NASS yield are also shown.

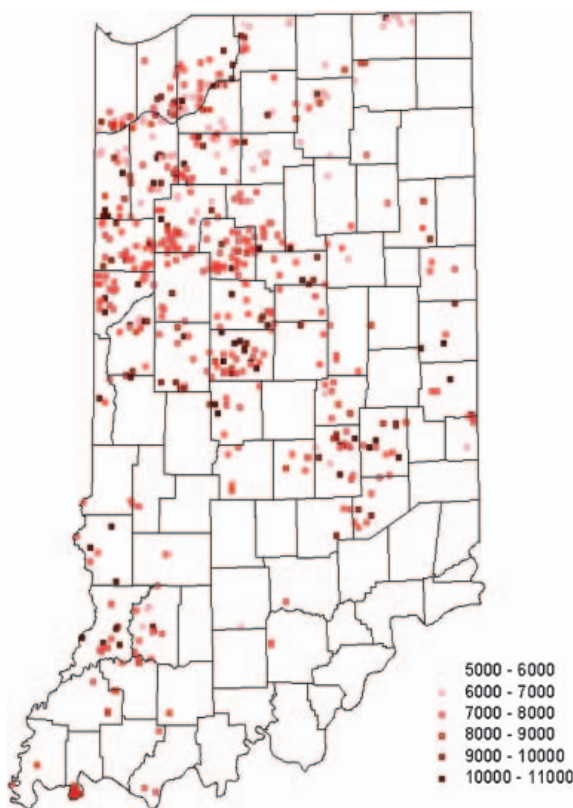


Figure 8. Simulated corn yield (871 pixels in total) with S1 (13 LAIs) in 2000 (unit: kg ha^{-1}).

simulated mean corn potential yields ranged from 6705 to 15 525 kg ha^{-1} , while the average yield was 9477 kg ha^{-1} (table 3). Results from the other two scenarios (S2: four LAIs and S3: three LAIs) were more than 11% higher than the NASS data. However, the standard deviations from all three scenarios were much higher than the NASS data. This was attributed to the fact that some counties that had only a few scattered corn pixels and yields.

The NASS data showed a gradient of increasing yields from north to south in Indiana. However, this pattern was not so obvious for the simulated corn yields (figures 7 and 8). The weakness of spatial gradient was attributed to the uncertainties in MODIS products and the small number of 'pure' corn pixels used in the study. The overestimation in figure 7 may be caused by several factors. For example, drought stress, pests, diseases, and harvest loss were not accounted for in our simulation. In addition, it was also very difficult to determine the exact management practices that farmers use. A few underestimations happened for some north and south counties. This was mainly due to the inadequacy of soil and topography information described in the previous section.

In addition to the yield, the ICSM also simulated LAI itself, which can be compared with the raw MODIS LAI products. Figure 9 shows the spatial distribution of simulated LAI for day 217, 2000. Figure 10 compares the MODIS

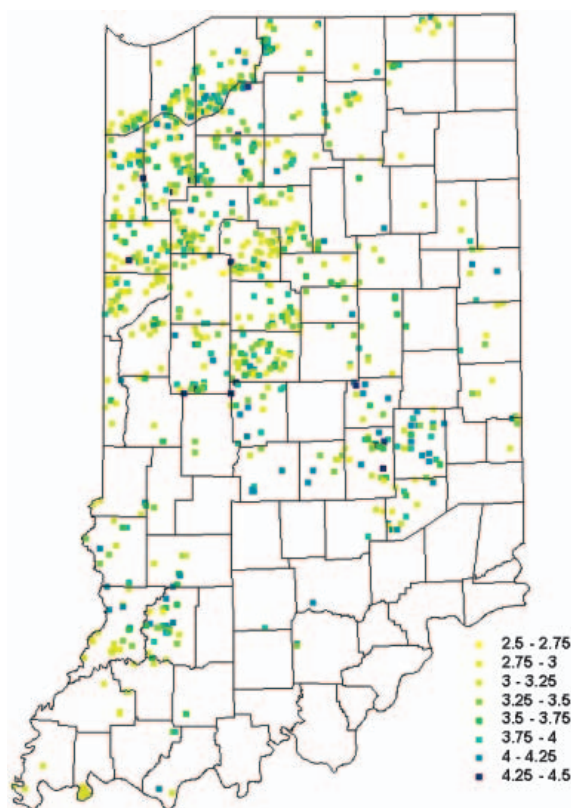


Figure 9. Simulated LAI for day 217, 2000 for the same pixels in figure 8.

LAI and the simulated LAI for day 217 of 2000. The ordinate shows the probability density of the data. For comparison, the temporally filtered LAI (LAI-SG) is also displayed for this specific day. Obviously, most of the MODIS LAI, located between 1.0 and 2.0, is underestimated due to the cloud effect (figure 10). After the SG filtering, the LAI is more evenly distributed between 0.5 and 4.0, while there is a density peak around 2.9. The peak of the simulated LAI (3.1) is very close to LAI-SG. There is another peak (4.2) for the simulated LAI. The range of simulated LAI is within those of the MODIS LAI and LAI-SG, but more pixels are located at the higher end of the histogram compared with the LAI-SG (figure 10). Therefore, the mean of the simulated LAI (3.5) was larger than the LAI-SG (2.3). Their different distributions are attributed to the mixed pixels.

This method has one significant merit to generate important agronomic parameters through inversion. The estimated dates of planting, emergence, and maturation are shown in figure 11. In 2000, 90% of corn was planted between 19 April (DOY 110) and 19 May (DOY 140) (figure 11). In general, this was very similar to the field observations (<http://www.nass.usda.gov/in/annbul/0304/04weather.html>). Corn emerges around 10 days after planting. The distribution profile of the planting date and emergence date were very similar (figure 11). Most corn (>92%) matured between 27 August (DOY 240) and 26 September (DOY 270).

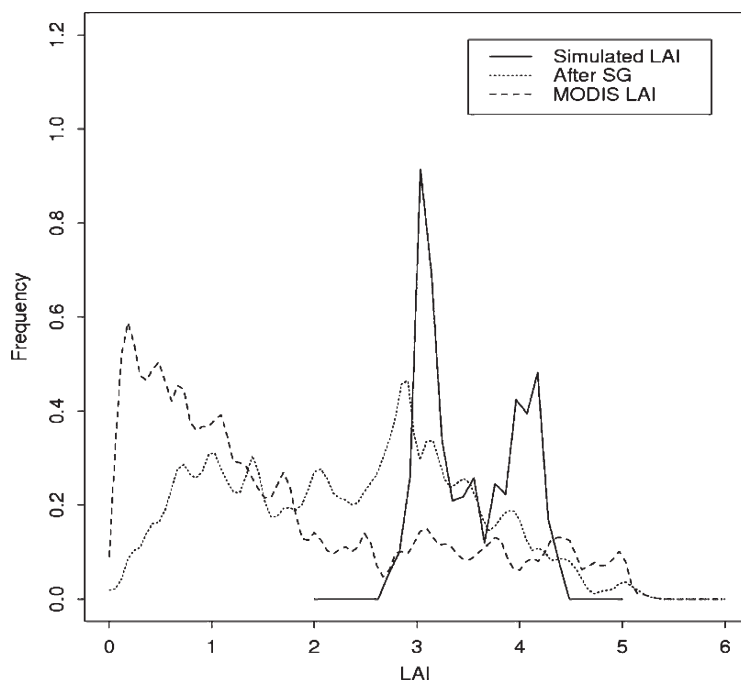


Figure 10. Comparison of simulated LAI from ICSM with raw MODIS LAI data for day 217, 2000. The intermediate LAI after Savitzky–Golay filtering is also shown. The ordinate shows the probability density of the data.

6. Discussion and summary

Crop simulation models simulate crop growth status and predict crop yield and its associated uncertainties at harvest maturity. Remote sensing is a powerful tool for estimating crop biophysical parameters. Based on the comparison between modelled and estimated LAI, the input variables, such as the planting date, planting population, row spacing and the nitrogen amount, were adjusted. Quality LAI products are critical in such data assimilation studies. MODIS LAI data for crops have been shown to agree well with field measurements by several validation studies (Cohen *et al.* 2003, Fang and Liang 2005). Nevertheless, continuous measurements of LAI and other biophysical parameters are needed to fulfil the reparameterization needs. This study applied the 8-day MODIS LAI product; however, a higher revisit frequency (e.g. daily) would be better for the simulation and inversion process in ICSM. Field LAI measurements could also be used as a supplement to the remotely sensed data during the reparameterization process.

This study focused on using the 1-km MODIS data for regional application. Crop-type information derived from 250-m MODIS data (Lobell and Asner 2004) may reduce biases brought in by mixed pixels. Moreover, other satellite products such as the vegetation indices can be used. The current ICSM has included the function to integrate remotely sensed vegetation indices with a crop-growth model. If the model were run at a local field, higher-spatial-resolution sensors like ETM+ would be advantageous. Application of the methodology from this study to precision farming is therefore very attractive. Given the relatively coarse resolution

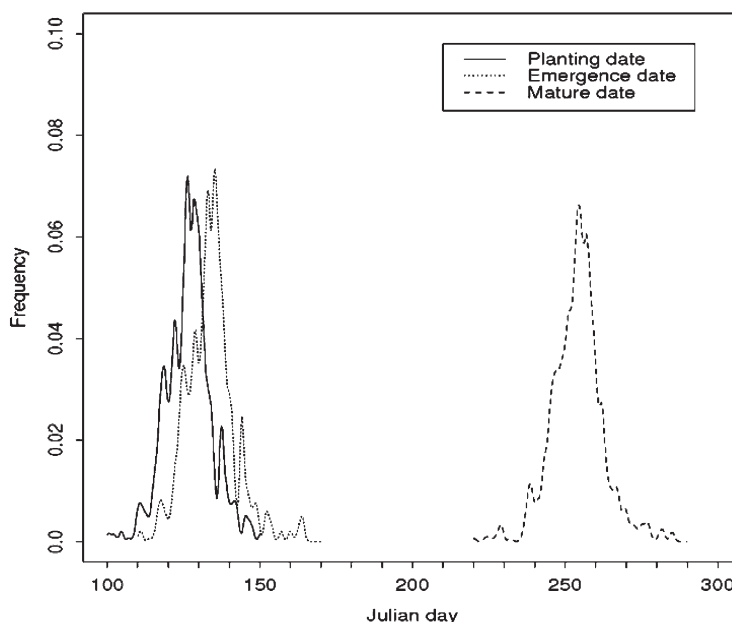


Figure 11. Histogram distribution of planting, emergence, and maturation dates retrieved for year 2000. The ordinate shows the probability density of the data.

of the MODIS data, we suggest that field campaigns be conducted and ETM+ data be used to develop regional calibrations.

Soil and weather data are two crucial environmental parameters. A general conclusion cannot be made without considering the soil and weather data used in the study. The inversion in this study was based on the STATSGO soils database. The inversion approach presented in this paper could certainly benefit from collecting higher-quality soil data such as the soil survey geographic database (SSURGO) (<http://www.ncgc.nrcs.usda.gov/branch/ssb/products/ssurgo/index.html>). With respect to the weather data, discrepancies probably arose due to uncertainties in both downward solar radiation and precipitation (Luo *et al.* 2003). Since we focused on the application of the LDAS data, we did not attempt to resolve these discrepancies, but it leaves some uncertainties in the results. However, the temporal dimension of the LDAS has not been fully explored. More follow-up studies are necessary, involving the use of improved forcing data and incorporating the hourly data into ICSM.

Currently, the reparameterization was performed on a pixel basis because almost all crop models are developed for point-based simulations. On a larger scale, for the real-world applications, this method will inevitably require considerable computation resources. Similarly, the remote-sensing inversion model (equation (2)) must also be executed pixel by pixel, and so it is very time- and computer-intensive when applied to a large area. An alternative is to use the look-up table method (Kimes *et al.* 2000). The look-up table method may be applied in the data assimilation and reparameterization process to increase computing efficiency. This approach will be very useful when a new generation of crop growth models for the regional scale becomes available (Hansen and Jones 2000).

This paper discusses the application of MODIS LAI data to predict crop yield through the inversion of a crop simulation model. The agreement between the estimated yield and NASS is fairly good. STATSGO soil parameters and NLDAS weather data offer reasonable inputs for simulating crop yield over a regional scale. Simulated corn yields were compared with yields reported by the National Agricultural Statistical Service (NASS). The simulated corn-yield at the 1 km² resolution captured the yield variability within the state and showed a reasonable agreement with the NASS yield reports. There is some room for future improvement, such as the selection of free variables and their initialization, and alternative optimization algorithms. This method can be used to predict regional or national crop-yield data before harvest. The outputs from ICSM are worthy of further investigation, for example, for studying regional or continental water and energy dynamics.

Acknowledgements

Dr R. L. (Bob) Nielsen, Professor of Agronomy, Purdue University, is acknowledged for his critical assistance in Indiana corn management. Dr William D. Batchelor, Professor of Agricultural and Biosystems Engineering, Iowa State University, is acknowledged for his help in CERES parameterization. This work was supported by a US Department of Agriculture grant SCA58-1275-9-096.

References

- BACH, H., MAUSER, W. and SCHNEIDER, K., 2003, The use of radiative transfer models for remote sensing data assimilation in crop growth models. In *Precision agriculture: Papers from the 4th European Conference on Precision Agriculture*, 15–19 June 2003, Berlin.
- BARET, F., WEISS, M., TROUFLEAU, D., PREVOT, L. and COMBAL, B., 2000, Maximum information exploitation for canopy characterisation by remote sensing. *Aspects of Applied Biology*, **60**, pp. 71–82.
- BETHEL, G. and DOORN, B., 1998, USDA remote sensing technical and systems support for operational worldwide agricultural analysis. In *First International Conference: Geospatial Information in Agricultural and Forestry*, Disney's Coronado Springs Resort, Lake Buena Vista, FL.
- BOUMAN, B.A.M., 1992, Linking physical remote sensing models with crop growth simulation models, applied for sugar beet. *International Journal of Remote Sensing*, **13**, pp. 2565–2581.
- CLEVERS, J.P.G.W. and LEEUWEN, V., 1996, Combined use of optical and microwave remote sensing data for crop growth monitoring. *Remote Sensing of Environment*, **56**, pp. 42–51.
- COHEN, W.B., MAIERSPERGER, T.K., YANG, Z., GOWER, S.T., TURNER, D.P., RITTS, W.D., BERTERRETICHE, M. and RUNNING, S.W., 2003, Comparisons of land cover and LAI estimates derived from ETM+ and MODIS for four sites in North America: a quality assessment of 2000/2001 provisional MODIS products. *Remote Sensing of Environment*, **88**, pp. 233–255.
- DELÉCOLLE, R., MAAS, S.J., GUÉRIF, M. and BARET, F., 1992, Remote sensing and crop production models: present trends. *ISPRS Journal of Photogrammetry and Remote Sensing*, **47**, pp. 145–161.
- DORAISWAMY, P.C., HATFIELD, J.L., JACKSON, T.J., AKHMEDOV, B., PRUEGER, J. and STERN, A., 2004, Crop condition and yield simulation using Landsat and MODIS. *Remote Sensing of Environment*, **92**, pp. 548–559.

- DORAISWAMY, P.C., MOULIN, S., COOK, P.W. and STERN, A., 2003, Crop yield assessment from remote sensing. *Photogrammetric Engineering and Remote Sensing*, **69**, pp. 665–674.
- ENGEL, T., HOOGENBOOM, G., JONES, J.W. and WILKENS, P.W., 1997, AEGIS/WIN: A computer program for the application of crop simulation models across geographical areas. *Agronomy Journal*, **89**, pp. 919–928.
- FANG, H. and LIANG, S., 2005, A hybrid inversion method for mapping leaf area index from MODIS data: Experiments and application to broadleaf and needleleaf canopies. *Remote Sensing of Environment*, **94**, pp. 405–424.
- FISCHER, A., KERGOAT, L. and DEDIEU, G., 1997, Coupling satellite data with vegetation functional models: review of different approaches and perspectives suggested by the assimilation strategy. *Remote Sensing Reviews*, **15**, pp. 283–303.
- GUÉRIF, M. and DUKE, C., 1998, Calibration of the SUCROS emergence and early growth module for sugar beet using optical remote sensing data assimilation. *European Journal of Agronomy*, **9**, pp. 127–136.
- GUÉRIF, M. and DUKE, C.L., 2000, Adjustment procedures of a crop model to the site specific characteristics of soil and crop using remote sensing data assimilation. *Agriculture, Ecosystems & Environment*, **81**, pp. 57–69.
- HANSEN, J.W. and JONES, J.W., 2000, Scaling-up crop models for climate variability applications. *Agricultural Systems*, **65**, pp. 43–72.
- HOOGENBOOM, G., JONES, J.W., WILKENS, P.W., PORTER, C.H., BATCHELOR, W.D., HUNT, L.A., BOOTE, K.J., SINGH, U., URYASEV, O., BOWEN, W.T., GIJSMAN, A., TOIT, J., WHITE, J.W. and TSUJI, G.Y. (Eds), 1999, *Decision Support System for Agrotechnology Transfer, Version 4.0*, CD-Rom (Honolulu, HI: University of Hawaii).
- INDIANA OFFICE OF THE COMMISSIONER OF AGRICULTURE, 2003, *A Profile of Indiana's Agricultural Industry, 2003–2004*.
- JIANG, D., WANG, N.B., YANG, X.H. and WANG, J.H., 2003, Study on the interaction between NDVI profile and the growing status of crops. *Chinese Geographical Science*, **13**, pp. 62–65.
- JONES, J.W., HOOGENBOOM, G., PORTER, C.H., BOOTE, K.J., BATCHELOR, W.D., HUNT, L.A., WILKINS, P.W., SINGH, U., GIJSMAN, A.J. and RITCHIE, J.T., 2003, The DSSAT cropping system model. *European Journal of Agronomy*, **18**, pp. 235–265.
- JUSTICE, C.O., VERMOTE, E., TOWNSHEND, J.R.G., DEFRIES, R., ROY, D.P., HALL, D.K., SALOMONSON, V.V., PRIVETTE, J.L., RIGGS, G., STRAHLER, A., LUCHT, W., MYNENI, R.B., KNYAZIKHIN, Y., RUNNING, S.W., NEMANI, R.R., WAN, Z., HUETE, A.R. and VAN LEEUWEN, W., 1998, The Moderate Resolution Imaging Spectroradiometer (MODIS): land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing*, **36**, pp. 1228–1249.
- KIMES, D.S., KNYAZIKHIN, Y., PRIVETTE, J.L., ABUELGASIM, A.A. and GAO, F., 2000, Inversion methods for physically-based models. *Remote Sensing Review*, **18**, pp. 381–440.
- KRAVCHENKO, A.N. and BULLOCK, D.G., 2000, Correlation of corn and soybean grain yield with topography and soil properties. *Agronomy Journal*, **92**, pp. 75–83.
- KRAVCHENKO, A.N., BULLOCK, D.G. and BOAST, C.W., 2000, Joint multifractal analysis of crop yield and terrain slope. *Agronomy Journal*, **92**, pp. 1279–1290.
- LIANG, S., 2004, *Quantitative Remote Sensing of Land Surfaces* (New York: Wiley).
- LOBELL, D.B. and ASNER, G.P., 2004, Cropland distribution from temporal unmixing of MODIS data. *Remote Sensing of Environment*, **93**, pp. 412–422.
- LUO, L., ROBOCK, A., MITCHELL, K.E., HOUSER, P.R., WOOD, E.F., SCHAAKE, J.C., LOHMANN, D., COSGROVE, B., WEN, F., SHEFFIELD, J., DUAN, Q., HIGGINS, R.W., PINKER, R.T. and TARPLEY, J.D., 2003, Validation of the North American Land Data Assimilation System (NLDAS) retrospective forcing over the southern Great Plain. *Journal of Geophysical Research*, **108**, p. 8843, doi: 10.1029/2002JD003246.

- MAAS, S., 1988a, Using satellite data to improve model estimates of crop yield. *Agronomy Journal*, **80**, pp. 662–665.
- MAAS, S.J., 1988b, Use of remotely-sensed information in agricultural crop growth models. *Ecological Modeling*, **41**, pp. 247–268.
- MAAS, S.J., 1993, Within-season calibration of modeled wheat growth using remote sensing and field sampling. *Agronomy Journal*, **85**, pp. 669–672.
- MACDONALD, R. and HALL, F., 1980, Global crop forecasting. *Science*, **208**, pp. 670–679.
- MITCHELL, K.E., LOHMANN, D., HOUSER, P.R., WOOD, E.F., SCHAAKE, J.C., ROBOCK, A., COSGROVE, B.A., SHEFFIELD, J., DUAN, Q., LUO, L., HIGGINS, R.W., PINKER, R.T., TARPLEY, J.D., LETTENMAIER, D.P., MARSHALL, C.H., ENTIN, J.K., PAN, M., SHI, W., KOREN, V., MENG, J., RAMSAY, B.H. and BAILEY, A.A., 2004, The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *Journal of Geophysical Research*, **109**, doi:10.1029/2003JD003823.
- MOULIN, S., BONDEAU, A. and DELÉCOLLE, R., 1998, Combining agricultural crop models and satellite observations: from field to regional scale. *International Journal of Remote Sensing*, **19**, pp. 1021–1036.
- MYNENI, R.B., HOFFMAN, S., KNYAZIKHIN, Y., PRIVETTE, J.L., GLASSY, J., TIAN, Y., WANG, Y., SONG, X., ZHANG, Y., SMITH, G.R., LOTSCH, A., FRIEDL, M., MORISSETTE, J.T., VOTAVA, P., NEMANI, R.R. and RUNNING, S.W., 2002, Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sensing of Environment*, **83**, pp. 214–231.
- O'NEAL, M.R., FRANKENBERGER, J.R. and ESS, D.R., 2002, Use of CERES-Maize to study effect of spatial precipitation variability on yield. *Agricultural Systems*, **73**, pp. 205–225.
- PLUMMER, S.E., 2000, Perspectives on combining ecological process models and remotely sensed data. *Ecological Modelling*, **129**, pp. 169–186.
- PRESS, W.H., TEUKOLSKY, S.A., VETTERLING, W.T. and FLANNERY, B.P., 1992, *Numerical Recipes in Fortran 77: The Art of Scientific Computing* (New York: Cambridge University Press).
- PRÉVOT, L., CHAUKI, H., TROUFLEAU, D., WEISS, M., BARET, F. and BRISSON, N., 2001, Assimilating optical and radar data into the STICS crop model for wheat. In *8th International Symposium Physical Measurements & Signatures in Remote Sensing*, Centre Paul Langevin, Aussois, France.
- SAVITZKY, A. and GOLAY, M.J.E., 1964, Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, **36**, pp. 1627–1639.
- TSUJI, G.Y., HOOGENBOOM, G. and THORNTON, P.K., 1998, *Understanding Options for Agricultural Production. System Approaches for Sustainable Agricultural Development* (Dordrecht: Kluwer Academic Publishers).
- TSUJI, G.Y., UEHARA, G. and BALAS, S., 1994, *Decision Support System for Agrotechnology Transfer, Version 3* (Honolulu, HI: University of Hawaii).
- TUCKER, C.J., HOLBEN, B., ELGIN, G. and MCMURTREY, J., 1980, Relationship of spectral data to grain yield variation. *Photogrammetric Engineering and Remote Sensing*, **46**, pp. 657–666.
- WANG, Y., TIAN, Y., ZHANG, Y., EL-SALEOUS, N., KNYAZIKHIN, Y., VERMOTE, E. and MYNENI, R.B., 2001, Investigation of product accuracy as a function of input and model uncertainties: case study with SeaWiFS and MODIS LAI/FPAR algorithm. *Remote Sensing of Environment*, **78**, pp. 296–311.
- WILLIAMS, J.R., JONES, C.A. and DYKE, P.T., 1984, A modeling approach to determining the relationship between erosion and soil productivity. *Transactions of the ASAE*, **27**, pp. 129–144.

Other Key Reference Papers for Prof. Fang's Lecture:

Fang, H., S. Liang, G. Hooogenboom, 2010. Integration of MODIS products and a crop simulation model for crop yield estimation. *International Journal of Remote Sensing* (in press).

The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system

Kenneth E. Mitchell,¹ Dag Lohmann,¹ Paul R. Houser,² Eric F. Wood,³ John C. Schaake,⁴ Alan Robock,⁵ Brian A. Cosgrove,² Justin Sheffield,³ Qingyun Duan,⁴ Lifeng Luo,^{5,6} R. Wayne Higgins,⁷ Rachel T. Pinker,⁸ J. Dan Tarpley,⁹ Dennis P. Lettenmaier,¹⁰ Curtis H. Marshall,^{1,11} Jared K. Entin,² Ming Pan,³ Wei Shi,⁷ Victor Koren,⁴ Jesse Meng,^{1,2} Bruce H. Ramsay,⁹ and Andrew A. Bailey⁹

Received 1 June 2003; revised 10 October 2003; accepted 28 October 2003; published 9 April 2004.

[1] Results are presented from the multi-institution partnership to develop a real-time and retrospective North American Land Data Assimilation System (NLDAS). NLDAS consists of (1) four land models executing in parallel in uncoupled mode, (2) common hourly surface forcing, and (3) common streamflow routing: all using a $1/8^\circ$ grid over the continental United States. The initiative is largely sponsored by the Global Energy and Water Cycle Experiment (GEWEX) Continental-Scale International Project (GCIP). As the overview for nine NLDAS papers, this paper describes and evaluates the 3-year NLDAS execution of 1 October 1996 to 30 September 1999, a period rich in observations for validation. The validation emphasizes (1) the land states, fluxes, and input forcing of four land models, (2) the application of new GCIP-sponsored products, and (3) a multiscale approach. The validation includes (1) mesoscale observing networks of land surface forcing, fluxes, and states, (2) regional snowpack measurements, (3) daily streamflow measurements, and (4) satellite-based retrievals of snow cover, land surface skin temperature (LST), and surface insolation. The results show substantial intermodel differences in surface evaporation and runoff (especially over nonsparse vegetation), soil moisture storage, snowpack, and LST. Owing to surprisingly large intermodel differences in aerodynamic conductance, intermodel differences in midday summer LST were unlike those expected from the intermodel differences in Bowen ratio. Last, anticipating future assimilation of LST, an NLDAS effort unique to this overview paper assesses geostationary-satellite-derived LST, determines the latter to be of good quality, and applies the latter to validate modeled LST. **INDEX TERMS:** 1878 Hydrology: Water/energy interactions; 1836 Hydrology: Hydrologic budget (1655); 1860 Hydrology: Runoff and streamflow; 1818 Hydrology: Evapotranspiration; 3337 Meteorology and Atmospheric Dynamics: Numerical modeling and data assimilation; **KEYWORDS:** land modeling, land data assimilation, surface energy budget

Citation: Mitchell, K. E., et al. (2004), The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *J. Geophys. Res.*, 109, D07S90, doi:10.1029/2003JD003823.

¹Environmental Modeling Center, National Centers for Environmental Prediction, National Oceanic and Atmospheric Administration—National Weather Service, Camp Springs, Maryland, USA.

²Hydrological Sciences Branch and Data Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

³Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA.

⁴Office of Hydrologic Development, National Oceanic and Atmospheric Administration—National Weather Service, Silver Spring, Maryland, USA.

⁵Department of Environmental Sciences, Rutgers University, New Brunswick, New Jersey, USA.

⁶Now at Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA.

⁷Climate Prediction Center, National Centers for Environmental Prediction, National Oceanic and Atmospheric Administration—National Weather Service, Camp Springs, Maryland, USA.

⁸Department of Meteorology, University of Maryland, College Park, Maryland, USA.

⁹Office of Research and Applications, National Environmental Satellite Data and Information Service, Camp Springs, Maryland, USA.

¹⁰Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington, USA.

¹¹Now at Department of Atmospheric Sciences, Colorado State University, Fort Collins, Colorado, USA.

1. Introduction

[2] Improving weather and seasonal climate prediction by dynamical models requires multidisciplinary advances in providing reliable initial states for the atmosphere, ocean and land components of the Earth system. For two decades, advances in providing atmospheric initial states via four-dimensional data assimilation (4DDA) have paved the way for emerging 4DDA systems for the ocean and land. The backbone of any 4DDA system is the geophysical model whose execution provides temporally and spatially continuous background states, into which generally discontinuous observations are assimilated from various observing platforms (in situ, satellite, radar). For example, present space-based microwave estimates of soil moisture sense only the top 1–5 cm of soil, far short of the root-zone depths needed for land-state initialization.

[3] Thus a land data assimilation system (LDAS) is needed to blend sparse land observations with the background fields of a land surface model (LSM). The accuracy of the LSM background field (and companion surface and subsurface water/energy fluxes) is crucial to LDAS viability. The chief objective of the North American Land Data Assimilation System (NLDAS) study here is to generate and validate, over a 3-year period over the continental U.S. (CONUS) domain, the background land states and surface fluxes of four LSMs: Noah, Mosaic, VIC, and Sacramento (denoted SAC) (hereinafter, all acronyms are defined in the Notation). Future NLDAS papers will address actual data assimilation experiments using such methods as adjoint models and Kalman filtering. As one step to assimilation of satellite land surface skin temperature (LST), this paper assesses geostationary-satellite-derived LST and uses it to validate NLDAS LST.

[4] It is instructive to consider the infancy of real-time large-scale land 4DDA. Global atmospheric 4DDA has been a mainstay of operational NWP centers since the late 1970s. Real-time ocean 4DDA on large-scale ocean basins followed in the middle to late 1980s [Ji *et al.*, 1994] on the heels of the TOGA program. Yet until the mid-1990s, initiatives in real-time continental or global land 4DDA were virtually nonexistent. The first viable examples of real-time land 4DDA on continental or global scales were the coupled land-atmosphere 4DDA systems at major NWP centers such as NCEP [Kalnay *et al.*, 1996] and the European Centre for Medium-Range Weather Forecasts [Gibson *et al.*, 1997]. Such coupled land-atmosphere 4DDA systems (including global reanalysis) often yield significant errors and drift in soil moisture/temperature and surface energy/water fluxes, owing to substantial biases in the surface forcing from the parent atmospheric models. To constrain such errors and drift, coupled land-atmosphere 4DDA systems temporally nudge the soil moisture by such means as (1) a climatology of soil moisture [Kalnay *et al.*, 1996], (2) differences between the observed and 4DDA background fields of precipitation [Kanamitsu *et al.*, 2002], or (3) screen-level air temperature and dew point [Douville *et al.*, 2000]. Such nudging methods, however, do not reduce the main error source, namely, large bias in the land surface forcing (especially precipitation and solar insolation) of the parent atmospheric model.

[5] Substantial biases in atmospheric model surface forcing also plague ocean 4DDA. To improve these surface fluxes, “flux corrections” are applied in ocean 4DDA [Ji *et al.*, 1994]. NLDAS here also applies surface flux corrections. As a pathfinder for this, the GEWEX Global Soil Wetness Project (GSWP) [Dirmeyer *et al.*, 1999] retrospectively demonstrated the viability of using nonmodel, observation-based precipitation analyses and nonmodel, satellite-based surface insolation fields (with all other surface forcing from atmospheric 4DDA) to drive uncoupled, land surface models over a global domain. However, the monthly satellite retrievals of precipitation and insolation used in GSWP are not conducive to the daily/weekly updates of land states needed to initialize operational prediction models. Hence the NLDAS project set and achieved the following key objectives: (1) develop and execute the first real-time operational prototype of a continental-scale uncoupled land 4DDA backbone (continuously cycled land-model states) executed daily at NCEP using real-time streams of hourly to daily data and (2) a companion retrospective mode for research. The NLDAS generates hourly surface forcing (using model-independent, observation-based precipitation and insolation fields) that drives four LSMs running in parallel to produce hourly output on a $1/8^\circ$ grid over a CONUS domain.

[6] The retrospective NLDAS spans October 1996 to September 1999 and uses GCIP-supported archives of NOAA operational data streams. NLDAS thus provides a land 4DDA counterpart from the GEWEX community to complement the ocean 4DDA thrusts that followed TOGA. Moreover, a core objective of GCIP is the infusion of GCIP research into NOAA operational practice. The NLDAS partnership of operational and research investigators in both meteorology and hydrology is a flagship of GCIP success in such infusion. This paper gives an overview of the methodology and results of the initial development and evaluation of NLDAS, providing an overview of the nine papers by NLDAS partners given in Table 1 (hereinafter, each paper is cited with the label given in Table 1, denoting the last initials of first two authors and N for NLDAS).

[7] These papers and the sections that follow illustrate that a pillar feature of NLDAS is the integrated application of a multitude of GCIP-sponsored products, as listed in Table 2. Section 2 describes the NLDAS configuration, surface forcing, land models, and streamflow routing. Sections 3 and 4 evaluate the NLDAS surface water budget and surface energy budget, respectively. Section 5 presents conclusions and future plans.

2. NLDAS Configuration

2.1. General Configuration

[8] Pilot studies of ocean 4DDA began on relatively data-rich subglobal domains [Ji *et al.*, 1994], i.e., the tropical Pacific Ocean, which included the TOGA observing network. For the NLDAS domain, we also chose a relatively data-rich subglobal domain, thereby heeding a lesson from GSWP, namely that uncoupled land surface simulation is notably less viable over regions lacking moderately dense precipitation gages to anchor the precipitation forcing [Oki *et al.*, 1999]. Outside such regions, global precipitation analyses are dominated by satellite-based precipitation,

Table 1. List and Topics of the Nine Companion Papers by NLDAS Partners in the GCIP-3 Special Issue

	Label	Subject
<i>Cosgrove et al.</i> [2003a]	CL-N	generation of land surface forcing
<i>Luo et al.</i> [2003]	LR-N	validation of land surface forcing
<i>Pinker et al.</i> [2003]	PT-N	production/validation of GOES-based solar insolation
<i>Lohmann et al.</i> [2004]	LM-N	production/validation of streamflow and water budget
<i>Robock et al.</i> [2003]	RL-N	validation of energy budget, soil moisture/temperature
<i>Schaake et al.</i> [2004]	SD-N	evaluation of soil moisture storage and range
<i>Sheffield et al.</i> [2003]	SP-N	validation of simulated snow cover
<i>Pan et al.</i> [2003]	PS-N	validation of simulated snowpack content
<i>Cosgrove et al.</i> [2003b]	CM-N	evaluation and testing of spin-up

which may be only marginally better (or even worse) than model-based precipitation, especially in the extratropics and in winter. Hence we limit the NLDAS domain (shown in Figure 1) essentially to the CONUS, thereby benefiting from relatively dense precipitation gages and the CONUS-oriented GCIP-supported products in Table 2.

[9] On this domain, NLDAS applies the following in common across the four LSMs: a $1/8^\circ$ regular latitude/longitude grid, land mask and terrain elevation, hourly input surface forcing, soil texture and vegetation classes, streamflow network and routing model, and content, frequency (hourly) and format (GRIB) of model input and output. The elevation was derived by averaging, in each $1/8^\circ$ grid cell, the 30 arc-second (~ 1 km) digital elevation of the GTOPO30 database of *Verdin and Greenlee* [1996]. Of the four LSMs, VIC alone also employs subgrid elevation tiles (see section 3.4). The vegetation classification was derived from the global, 1-km, AVHRR-based, 13-class vegetation database of UMD [*Hansen et al.*, 2000]. For each $1/8^\circ$ cell, the vegetation field includes the percent of each class based on its 1-km frequency. Mosaic and VIC use subgrid vegetation tiles, whose weights correspond to the percent of the classes. Noah uses the most predominant vegetation class. SAC omits explicit treatment of vegetation.

[10] The soil texture database over CONUS was derived from the 1-km STATSGO database of *Miller and White* [1998], which carries 16 texture classes by layer over 11 layers to 2-m depth. For each $1/8^\circ$ grid cell, the NLDAS soil database carries the percent of each class by layer, based on the original 1-km frequency. Noah, Mosaic, and VIC assume a vertically uniform soil class based on the predominant soil texture of the top 5-cm layer. The exception is Mosaic's soil porosity, derived for each Mosaic soil layer based on weighted averages from the 11-layer soil textures. Though SAC uses conceptual soil-water storage reservoirs rather than explicit soil parameters of an explicit soil column, many a priori parameters for the NLDAS SAC were derived using all 11-layer soil types. Outside the CONUS, the NLDAS soil database applies the same 16 texture classes, but carries only a single, vertically uniform class at each cell, derived from the 5-min ARS FAO global data of *Reynolds et al.* [2000]. Spatial maps depicting the NLDAS vegetation and soils databases, as well as tables defining NLDAS vegetation and soil classes, may be viewed under the NLDAS tab at <http://ldas.gsfc.nasa.gov>, maintained by NASA.

[11] Although NLDAS control runs employ common fields of vegetation and soil class, the NLDAS partners chose NOT to impose additional commonality in the vegetation and soil properties, such as (1) parameter values,

(2) configuration of a vegetation class (root depth and density) or the soil column (number and thickness of layers) and (3) seasonal cycle of vegetation. The desire was to avoid negating the legacy of calibration or tuning invested over the past decade in the LSMs. No extra calibration of LSMs was carried out for the control runs. (Additionally, the NLDAS project conducted various sensitivity tests, discussed throughout this overview.)

Table 2. GCIP-Supported Products Applied by the NLDAS Project

	Product
	<i>A: For Producing NLDAS Surface Forcing</i>
1	daily $1/8^\circ$ gage-only CONUS precipitation analysis by NCEP
2	hourly 4-km radar-dominated (WSR-88D) CONUS precipitation analysis by NCEP and OHD
3	hourly $1/2^\circ$ GOES-based CONUS surface insolation by NESDIS and UMD
4	3-hourly 40-km Eta-based 4DDA analyses of near-surface meteorology by NCEP
	<i>B: For Validating NLDAS Surface Forcing</i>
5	OU Mesonet surface meteorology observations
6	SURFRAD network of surface solar insolation observations (receives support from GCIP sister program in NOAA/OGP)
	<i>C: For Validating NLDAS Land Model Output (States/Fluxes)</i>
7	Oklahoma/Kansas ARM/CART surface flux stations (DOE)
8	Oklahoma Mesonet soil moisture/temperature observations (OU Climate Survey)
9	CONUS-wide GOES-based satellite LST (NESDIS and UMD)
10	Northern Hemisphere 23-km IMS daily snow cover analysis by NESDIS
11	Illinois Water Survey network of 18 soil moisture measuring stations
12	western U.S. network of SNOTEL observations (not GCIP supported)
13	USGS streamflow observations (not GCIP supported)
	<i>D: For NLDAS Land Surface Characteristics</i>
14	1-km CONUS soil texture database by Pennsylvania State University
15	NESDIS 0.144° global monthly NDVI-based vegetation greenness by NESDIS
	<i>E: Improved Land/Hydrology Models (LSMs)</i>
16	Noah LSM improvements (including in coupled EDAS) by NCEP, OHD and collaborators
17	VIC LSM improvements (Princeton University, University of Washington and collaborators)
18	SAC LSM improvements by OHD, NCEP and collaborators
19	Mosaic LSM improvements by NASA GSFC and collaborators
20	streamflow connectivity network and routing model by NCEP, University of Washington, Princeton University and OHD

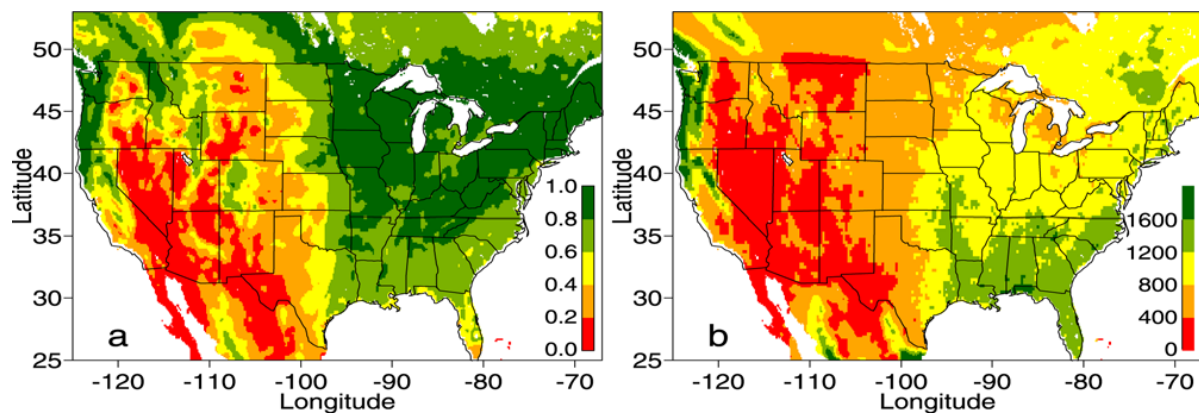


Figure 1. Depiction of NLDAS domain, showing (a) July mean green vegetation fraction from Gutman and Ignatov [1998] and (b) mean annual NLDAS precipitation (mm) for 1 October 1997 to 30 September 1999.

[12] For example, NLDAS does not impose a common treatment for the seasonality of vegetation, as the latter is central to a given model's canopy resistance formulation. In broad terms, Noah, VIC, and Mosaic runs here all use a satellite-derived, AVHRR-based, monthly seasonality of vegetation and all interpolate their respective monthly values to daily. Yet significant nuances exist between LSMs regarding whether the monthly values are for the given year (Mosaic) or from a multiyear climatology (Noah and VIC) and whether the seasonality is carried in LAI (VIC), or in vegetation fraction (Noah), or both (Mosaic). Noah uses the global, 0.144° (~ 15 -km), monthly 5-year climatology of the green vegetation fraction (GVF) derived by Gutman and Ignatov [1998] of NESDIS from AVHRR-based NDVI. Figure 1a depicts this GVF climatology for the NLDAS domain for July, as a reference for later sections. For Mosaic, NASA obtained monthly 16-km AVHRR-based green LAI fields from Boston University for each month of each year (not climatology). NASA then first derived dead LAI (estimated from the difference in green LAI between consecutive months, along with vegetation-class-dependent values of minimum dead LAI) and then derived monthly total LAI (sum of green and dead LAI), from which GVF was computed (as green LAI divided by total LAI). VIC applies a global, AVHRR-based, multiyear monthly climatology of total LAI, used in conjunction with a vegetation-class-dependent look-up table of fixed annual-maximum vegetation fraction (dead and green).

[13] The LSMs of NLDAS provide common hourly output of about 50 required fields, including all terms of the surface energy and water budgets, all soil and snowpack moisture and temperature states, and ancillary fields. For a given LSM, some outputs are omitted if its physics omit the relevant process. NLDAS input/output fields are viewable at the NLDAS tab of <http://ldas.gsfc.nasa.gov>. For utility in NCEP operations, the input-output format of NLDAS is GRIB: a WMO standard at NWP centers. The NLDAS may be run in a "reduced-domain" mode by reducing the land mask, say to a single or handful of points (e.g., near flux stations), for purposes of lower output volume, faster execution, and locally focused sensitivity studies.

[14] The NLDAS requires initial values of all LSM state variables for the NLDAS start time of 00 UTC on 1 October 1996. Initial snowpack was set to zero (reasonable for 1 October over the NLDAS domain at $1/8^\circ$ resolution), as was canopy interception storage. Initial states of soil moisture and temperature were derived from the soil states of the NCEP/DOE Global Reanalysis 2 [Kanamitsu *et al.*, 2002] valid at the start time. The soil moisture was provided to each LSM as a vertically uniform percent of saturation, which each LSM converted to its own absolute moisture state compatible with its parameters. The spin-up from this cold start was examined in all four LSMs by CM-N, who concluded that the practical drift in NLDAS land stores ceased within about one year. CM-N also conducted spin-up experiments in Mosaic, in which spin-up from the above initial states was found to be shorter than using saturated or dry initial states.

2.2. Surface Forcing

[15] The studies by Cosgrove *et al.* [2003a] (CL-N), Pinker *et al.* [2003] (PT-N), and Luo *et al.* [2003] (LR-N) summarized below describe the data sources, generation and validation of NLDAS forcing, produced in real-time and retrospectively on the NLDAS grid. Of the 16 fields in each forcing file (Table 3), nine fields required by Noah, Mosaic, and VIC are primary: U/V 10-m wind components, 2-m air temperature and specific humidity, surface pressure, downward longwave and shortwave radiation, and convective and total precipitation. SAC requires only total precipitation (P), air temperature and potential evaporation (PE). In NLDAS, SAC uses the PE computed in the Noah LSM. Mosaic alone requires convective precipitation.

[16] The chief source of NLDAS forcing is NCEP's Eta-model-based Data Assimilation System (EDAS) [Rogers *et al.*, 1995], a continuously cycled North American 4DDA system. It utilizes 3-hourly analysis-forecast cycles to derive atmospheric states by assimilating many types of observations, including station observations of surface pressure and screen-level atmospheric temperature, humidity and U and V wind components. EDAS 3-hourly fields of the latter five variables plus surface downward shortwave and longwave radiation and total and convective precipitation are provided

Table 3. Content and Data Sources of the Fields in the Hourly Surface Forcing Files of NLDAS^a

Content	EDAS	GOES	Gage	Radar	Real-Time	Retrospective
Primary forcing						
2-m temperature, K	X				X	X
2-m specific humidity, kg/kg	X				X	X
10-m U-wind component, m/s	X				X	X
10-m V-wind component, m/s	X				X	X
Surface pressure, mb	X				X	X
Downward longwave radiation, W/m ²	X				X	X
Downward shortwave radiation, W/m ²		X			X	X
Convective precipitation, kg/m ²	X				X	X
Total precipitation, kg/m ²			X ^b	X ^b	X	X
Backup forcing						
Downward shortwave radiation, W/m ²	X				X	X
Total precipitation, kg/m ²	X				X	X
Auxiliary forcing						
WSR-88D precipitation, kg/m ²			X ^c	X ^c		
PAR, W/m ²		X			X	X
Downward diffuse radiation, ^d W/m ²		X			X	
CAPE	X				X	X
For validation (plus future assimilation)						
Land surface temperature (LST), ^e K		X			X	X

^aTo date, retrospective forcing is available from 1 October 1996 through 2002. Real-time forcing is available from 16 April 1999 to present.

^bDaily total is gage-only. Radar estimate is used only to temporally partition gage-based daily into hourly.

^cWSR-88D precipitation estimate is radar dominated, but some gage data are used.

^dDiffuse radiation is present in forcing files since 15 November 1999.

^eLST is present in forcing files since 1 May 1997 (LST derivation and application in section 4.3).

on a 40-km grid to NLDAS forcing software, which interpolates the fields spatially to the NLDAS grid and temporally to one hour. Last, to account for NLDAS versus EDAS surface-elevation differences, a terrain-height adjustment is applied to the air temperature and surface pressure using a standard lapse rate (6.5 K km^{-1}), then to specific humidity (keeping original relative humidity) and downward longwave radiation (for new air temperature, specific humidity). CL-N details the spatial/temporal interpolations and terrain-height adjustment.

[17] EDAS precipitation and shortwave radiation serve only as backup (Table 3). Though Roads *et al.* [2003] found EDAS monthly precipitation to have rather smaller errors than other 4DDA systems, EDAS precipitation errors can be significant for daily events, as in summer convection (Figure 4 of CL-N). Thus NLDAS precipitation forcing over CONUS is anchored instead to NCEP's $1/4^\circ$ gage-only daily precipitation analyses of Higgins *et al.* [2000] (available at http://www.cpc.ncep.noaa.gov/research_papers/ncep_cpc_atlas/7/index.html), which utilize about 6500 (real-time) or 13000 (retrospectively) gage observations of daily precipitation. In NLDAS, this daily analysis is interpolated to $1/8^\circ$, then temporally disaggregated to hourly (details given by CL-N) by applying hourly weights derived from hourly, 4-km, radar-based (WSR-88D) precipitation fields. The latter radar-based fields (saved as auxiliary field in Table 3) are used only to derive disaggregation weights and do not change the daily total precipitation. Last, convective precipitation is estimated by multiplying NLDAS total precipitation by the ratio of EDAS convective to EDAS total precipitation. Figure 1b shows the annual mean NLDAS precipitation for the two years of the water budget analysis done later in section 3.1. CL-N shows examples of hourly and daily NLDAS precipitation fields, as well as EDAS and radar-based counterparts.

[18] Downward shortwave radiation (solar insolation) in the EDAS and Eta model typically show high bias of 10–20% [Betts *et al.*, 1997], even higher in cloudy winter conditions. At two SURFRAD sites, Figure 7 of CL-N illustrates the high bias in EDAS insolation and the far less bias in GOES-based solar insolation, which provides the primary insolation forcing for NLDAS. PT-N describes the retrieval of this $1/2^\circ$ surface insolation from GOES since January 1996, via collaboration of UMD, NESDIS/ORA and NCEP, and its hourly to monthly validation against fifty CONUS stations over a 1–2 year period. GOES insolation is not retrieved for zenith angles below 75° and so is supplemented with EDAS insolation near the day/night terminator (Figure 5 of CL-N). In validations by PT-N and LR-N, GOES insolation verifies well against flux stations, with some deterioration toward high bias at low sun angles and over snow cover (when cloud detection is difficult). CL-N shows that even in winter, the high bias of the GOES insolation is about half that of EDAS. Last from the GOES-based product suite, downward diffuse radiation, PAR and LST fields are included in the NLDAS forcing files (Table 3).

[19] NCEP originally selected the viable real-time NOAA data sources (Table 2, part A) on which to base NLDAS surface forcing fields and then developed the algorithms for their real-time production, which NCEP has sustained since April 1999. GCIP has supported the archiving back to 1996 of all NOAA data streams needed for this forcing. NASA GSFC acquired these archives and adapted NCEP software to produce retrospective forcing for October 1996 through 2002, the first 3 years of which force the NLDAS executions evaluated here and in the NLDAS papers. The retrospective forcing was created for purposes of (1) executing NLDAS over longer periods, especially those overlapping special validating observations (Table 2), such as

Table 4. Primary Attributes of the Four NLDAS Land Surface Models (LSMs)

	Mosaic	Noah	VIC	SAC
Full domain runs	yes	yes	yes	yes
Limited domain runs	yes	yes	yes	yes
Input surface forcing	seven forcing fields ^a	seven forcing fields ^a	seven forcing fields ^a	precipitation, Noah PE, 2-m air temperature
Energy balance	yes	yes	yes	n/a
Water balance	yes	yes	yes	yes
Model time step	15 min	15 min	1 hour	1 hour
Model soil layers	3	4	3	2 storages
Model soil layer depths	10, 30, 160 cm	10, 30, 60, 100 cm	10 cm, variable	variable
Tiling: Vegetation	yes	no	yes	no
Tiling: Elevation	no	no	yes	no
Snow model layers	1	1	2	1
Frozen soil: thermal	no	yes	disabled	n/a
Frozen soil: hydraulics	partial	yes	disabled	n/a
Soil thermodynamics	force-restore	heat conduction equation	heat conduction equation modified	no
Soil temperature profile	no	yes	yes	no
Soil water: drainage	yes	yes	yes	yes
Soil water: vertical diffusion	yes	yes	no	no
Snow-free albedo	vary wrt LAI, GVF, biome	monthly input background field	vary wrt LAI and biome	n/a
Diurnal albedo	yes	no	no	n/a
Explicit vegetation	yes	yes	yes	no
Canopy resistance	<i>Sellers et al.</i> [1986]	<i>Jarvis</i> [1976]	<i>Jarvis</i> [1976]	n/a
Rooting depth	0.4 m	variable (1 or 2 m)	variable (1.35–3 m)	n/a
Root density profile	constant	constant	exponential	n/a
Canopy capacity	0–1.6 mm	0.5 mm	0.1–1.0 mm	n/a

^aSeven forcing fields: precipitation, downward solar and longwave radiation, 10-m wind speed, surface pressure, 2-m air temperature, 2-m air humidity.

the soil moisture used in RL-N, (2) leveraging the near doubling of gage observations (about 13000 versus 6500) of daily precipitation applied in the retrospective versus real-time CONUS precipitation analyses of *Higgins et al.* [2000], and (3) applying added quality control checks to and (when such checks warrant) reprocessing of the forcing that is not feasible in real time.

[20] The LR-N study assesses NLDAS retrospective forcing of January 1998 to September 1999 against hourly ARM/CART and OU Mesonet stations (yielding independent observations, not assimilated in EDAS). Except for precipitation, differences between NLDAS forcing and these observations were small at hourly to monthly time-scales. For precipitation the agreement was marginal at hourly periods, but better at daily and rather good at 5-day and monthly periods. In net radiation, a small low bias in downward longwave partially offset the modest high bias in solar insolation. To investigate how these differences impacted NLDAS simulations, LR-N compared control simulations using standard NLDAS forcing with test simulations using site-specific, station-observed forcing. Simulation differences in soil moisture and temperature for each LSM were small: much smaller than differences between the LSMs and between LSMs and observations. Thus NLDAS provides quality forcing for land modeling, at least over the nonwestern CONUS. In section 3.4, the study by PS-N uncovers a 50% low bias in NLDAS precipitation at mountain SNOTEL sites at high elevations in western CONUS. In the final section, we discuss future remedies that will apply an adjustment to PRISM [*Daly et al.*, 1994].

2.3. Land Models

[21] Table 4 compares the attributes of the LSMs in NLDAS. Of the many LSMs, these four give a good

cross-section of different early legacies, including small scale versus large scale, coupled versus uncoupled, distributed versus lumped, with and without explicit vegetation, tiled and nontiled, and significant versus minimal calibration. Mosaic and Noah emerged from the surface-vegetation-atmosphere transfer (SVAT) setting of coupled atmospheric modeling with little calibration. VIC and SAC grew from the hydrology community as uncoupled hydrology models with considerable calibration. Mosaic was developed for use in the NASA global climate model [*Koster and Suarez*, 1994, 1996; *Koster et al.*, 2000]. Noah was developed for use in the NCEP mesoscale Eta model [*F. Chen et al.*, 1997; *Betts et al.*, 1997; *Ek et al.*, 2003]. VIC was developed as a macroscale semi-distributed model [*Liang et al.*, 1994; *Wood et al.*, 1997]. SAC was developed as a lumped conceptual hydrology model [*Burnash et al.*, 1973], calibrated for small catchments and used operationally in NWS RFCs.

[22] Subsequent to their early heritage, Mosaic, Noah, and VIC have been widely executed coupled and uncoupled from small to large scales. Now all three models can be considered as both SVATs and semi-distributed hydrological models. All three have undergone testing on local and regional scales in the PILPS project [*T. H. Chen et al.*, 1997; *Wood et al.*, 1998; *Schlosser et al.*, 2000; *Bowling et al.*, 2003] and on the global scale in GSWP [*Dirmeyer et al.*, 1999]. We use “semi”-distributed to mean applied on a gridded basis with gridded state variables and gridded parameters, but no horizontal interaction between model grid cells, except for routing of gridded runoff into streamflow. Seeking SAC suitability over a broad range of scales, OHD of NWS recently developed a semi-distributed (non-lumped) version of SAC [*Koren et al.*, 2000] with a priori uncalibrated parameters. This SAC version is intended for testing from small basins to entire continents. NLDAS

provides the first tests of the semi-distributed SAC at continental scales. These are pilot tests, as SAC lacks the legacy of continental testing of the other LSMs.

[23] The three SVAT models simulate LST, the surface energy and water balance, snowpack, and soil moisture in several soil layers, though the number and thickness of the layers differ. Only Noah simulates soil freeze-thaw and its impact on soil heating or cooling and transpiration, after *Koren et al.* [1999]. The snowpack physics in the LSMs are described in section 3.4. In all three SVATs, the surface infiltration schemes account for subgrid variability in soil moisture and precipitation, but the treatments differ, as do the drainage approaches. All three SVATs include direct evaporation from soil, transpiration from vegetation, evaporation of interception, and snow sublimation; and all explicitly model canopy resistance, though their formulations (see Table 4) and parameters differ, as does their vegetation phenology (section 2.1) and root profiles. The aerodynamic conductance in the SVATs also differs, a focus of section 4.2.

[24] SAC is a conceptual rainfall-runoff, storage-type model [*Burnash et al.*, 1973]. It treats only the surface water budget, omitting the surface energy budget, and uses the snowpack model of *Anderson* [1973], called SNOW-17. Hereafter SAC means the SAC-SNOW-17 pair. SAC outputs evaporation E and runoff, with E being a fraction of input PE. SAC uses a “two-reservoir” soil water storage structure (a shallow upper reservoir and a deeper lower reservoir) and utilizes 28 parameters, 16 primary and 12 in SNOW-17. At the RFCs, the parameters are calibrated by catchment. Calibration was omitted in NLDAS SAC runs. Rather, the primary parameters are specified a priori, after *Koren et al.* [2000], as a function of the STATSGO-based soils of section 2.1 and other data sets. SNOW-17 parameters are prescribed uniformly over the domain.

[25] Mosaic was developed by *Koster and Suarez* [1994, 1996] to account for subgrid vegetation variability with a tile approach. Each vegetation tile carries its own energy and water balance and soil moisture and temperature. Each tile has three soil layers and the first two are the root zone. In NLDAS, Mosaic is configured to support a maximum of 10 tiles per grid cell with a 5% cutoff that ignores vegetation classes covering less than 5% of the cell. Additionally in NLDAS, all tiles of Mosaic in a grid cell have the predominant soil type of section 2.1 and three soil layers with fixed thickness values of 10, 30, and 160 cm (hence constant rooting depth of 40 cm and constant total column depth of 200 cm). This Mosaic configuration in NLDAS departs from the standard Mosaic configuration, for the purpose of easier comparison in NLDAS [*Robock et al.*, 2003] with the soil moisture observation levels of the OU Mesonet and the soil layers of the VIC and Noah models (e.g., their 10 cm top layer). Although never executed before with fixed layer thickness, Mosaic performed well in the PILPS experiments when configured in the standard way [*T. H. Chen et al.*, 1997; *Lohmann et al.*, 1998; *Wood et al.*, 1998]. The standard Mosaic configuration varies the soil type and layer thickness tile by tile by vegetation type and yields top-down layer thickness ranges of 1–2 cm, 1–150 cm, and 30–200 cm, total column depth ranges of 32–350 cm, and root depths of 2–49 cm for nonforest and 150 cm for forests.

[26] The Noah LSM [*Chen et al.*, 1996; *Koren et al.*, 1999; *Ek et al.*, 2003] is targeted for moderate complexity and computational efficiency for operational NWP and climate models. Thus it omits tiling and uses a single-layer snowpack, plus a linearized (noniterative) solution to the surface energy balance. Originating from the LSM of *Pan and Mahrt* [1987], Noah benefits from improvements arising from year-round assessment in the NCEP Eta model over North America by NCEP and collaborators [*Ek et al.*, 2003]. The Noah version here in NLDAS is that implemented in the NCEP Eta/EDAS suite on 19 June 2002 and includes four layers of fixed thickness (Table 4), of which the first three (nonforest) or four (forest) span the root zone. Virtually this same version of Noah was executed in NCEP’s 24-year Regional Reanalysis. *Berbery et al.* [2003] examines the large-scale hydrology of the coupled Eta/Noah model over the Mississippi Basin for the period June 1995 to May 2002.

[27] The variable infiltration capacity (VIC) model was developed at the University of Washington and Princeton University [*Liang et al.*, 1994, 1996a, 1996b; *Cherkauer and Lettenmaier*, 1999]. In NLDAS, VIC executes with one-hour time step and uses three soil layers, with 10 cm top layer and varying depth for bottom two layers, partly determined from calibration. The root zone can span all three layers, depending on vegetation class. Like Mosaic, the VIC model carries subgrid vegetation tiles. Additionally, VIC is the only LSM of the four to apply subgrid elevation bands or tiles (see section 3.4). VIC has been tested over large river basins, such as the Columbia [*Nijssen et al.*, 1997] and Arkansas-Red [*Abdulla et al.*, 1996; *Wood et al.*, 1997], and over continental scales [*Maurer et al.*, 2002; *Roads et al.*, 2003] and global scales [*Nijssen et al.*, 2001]. Traditional executions of VIC [e.g., *Maurer et al.*, 2002] apply a uniform disaggregation of total daily precipitation to VIC time steps (typically 3-hourly). In NLDAS, VIC is executed with one-hour time steps and nonuniform, radar-based disaggregation of daily precipitation (see section 2.2). Implications of this departure from standard VIC are presented in section 3.2.

2.4. Streamflow Simulation and Assessment

[28] *Lohmann et al.* [2004] (LM-N) present the formulation of the streamflow modeling in NLDAS and the assessment of the control-run simulations of daily streamflow by the four LSMs. The latter study assesses model streamflow for 9 major and 1145 small to medium-sized CONUS basins (ranging from 23 km² to 10,000 km²) using measured daily streamflow from the USGS. LM-N describes the criteria for choosing the 1145 basins. One criterion is the absence of obvious regulation signatures in the observed streamflow record. This yields few basins for assessment in the arid southwestern CONUS (see Figure 4).

[29] The streamflow routing requires both a river network (flow-direction mask) on the NLDAS grid and a routing model. LM-N derives and displays the river network. The chosen routing model is linear and identical to that in PILPS 2c and 2e [*Lohmann et al.*, 1998; *Bowling et al.*, 2003]. It calculates the timing of the runoff reaching the

grid-cell outlet, as well as the transport of water through the river network. It operates in two modes: (1) distributed, using a-priori grid-cell specific routing parameters common to all four models and (2) “lumped”, in which constant routing parameters were separately calibrated for each of the 1145 basins for each model to minimize the least squares difference between modeled and measured daily streamflow. The lumped mode thus yielded a separate calibrated unit hydrograph for each basin for each model.

[30] LM-N shows the streamflow time series of the Nehalem River in Oregon and associates the derived lumped routing function for each LSM with the timing of the runoff produced by each LSM (Figures 10 and 11 of LM-N). In turn, a major cool season impact on the modeled runoff is the timing of snowmelt in each LSM. LM-N quantifies the delay between modeled and measured streamflow by means of the maximum of the temporal cross-correlation between measured and modeled streamflow. In general for all the LSMs, streamflow performance was degraded with increasing snowfall amounts (Figure 17 of LM-N). The worst case of peak streamflow timing occurs in the snowpack season of the mountain ranges of the north-west CONUS. There the LSMs vary by up to four months in the timing of peak streamflow (Figure 18 of LM-N). VIC showed the most realistic timing in such regions. Mosaic and SAC melted on the order of weeks too early, and Noah on the order of months too early. The SM-N and PS-N studies featured in section 3.4 further examine and elucidate the snowpack and snow cover simulations and forcing.

[31] In contrast, throughout the eastern half of CONUS, streamflow simulation skill as measured by the Nash-Sutcliffe efficiency (Figure 16 of LM-N) showed that Noah had the highest scores in general; VIC had the highest scores in the northeast, and Mosaic and SAC in the southeast Atlantic coastal states. The higher skill for Noah streamflow reflected Noah having the smallest bias in evaporation and runoff when assessed against the observed annual water budget (next in section 3.1). Finally, the validation of simulated monthly discharge for the 9 large river basins showed behavior and bias in each LSM consistent with that expected from spatial integration of the behavior and effects established in the assessment of the smaller basins.

3. Assessment of the NLDAS Water Budget

[32] Sections 3 and 4 assess the water and energy budgets of the LSMs, respectively, in the three-year NLDAS simulation with retrospective forcing for 1 October 1996 to 30 September 1999. The assessment focuses on the last two years, termed the “control” simulation, since the first year is a necessary and (mostly) sufficient spin-up year, as reported by *Cosgrove et al.* [2003b] (CM-N). Additionally, complementary sensitivity tests of the LSMs in NLDAS are assessed. Altogether, the hallmark of the assessment is its breadth, addressing continental to local scales, all four seasons and multiple types of validating observations (Table 2, part C).

3.1. Annual Water Budget: Partitioning Between Evaporation and Runoff

[33] Over one or more annual cycles, the surface water budget is well approximated by mean annual precipitation

being equal to the sum of mean annual evaporation and mean annual runoff, since mean annual storage change (in soil moisture, snowpack, etc.) is negligible by comparison. Thus analysis of observed mean annual precipitation minus observed mean annual runoff (from observed streamflow) yields crucial observation-based estimates of mean annual evaporation, as was carried out for NLDAS by *Lohmann et al.* [2004] (LM-N) and summarized here.

[34] Figure 1b depicts the observed mean annual precipitation forcing common to each LSM in the NLDAS control run. Figure 2 depicts the simulated mean annual evaporation of each control-run LSM. (See Figure 6 of LM-N for companion map of simulated mean annual runoff.) Last, Figure 3a shows the partitioning of mean annual NLDAS precipitation into mean annual evaporation and runoff by each LSM for the four CONUS quadrants of SW, NW, NE, and SE depicted in the figure inset. In Figure 3a, each diagonal denotes the mean-annual area-averaged precipitation of a given quadrant (given by the diagonal's x or y axis intercept). On each diagonal, each LSM's symbol projected onto the x axis (y axis) yields that LSM's quadrant-average mean annual runoff (evaporation). Since each LSM conserves water over the annual cycle, the tiny displacement in Figure 3a of a given LSM symbol from the diagonal represents the negligible change in that LSM's total water storage over the two years.

[35] The disparity in evaporation among LSMs in Figures 2 and 3a is striking, especially over the well-vegetated NE and SE (Figure 1a). Disparity is far less over the arid SW and NW. Over the NE and SE, Noah and VIC have notably lower evaporation and hence higher runoff than Mosaic and SAC, which give high evaporation and low runoff. Mosaic similarity to SAC in Figures 2 and 3 is surprising, as Mosaic (like Noah and VIC) includes explicit treatment of vegetation cover and canopy conductance, while SAC does not. Large disparity among LSMs in evaporation versus runoff partitioning was noted in PILPS and GSWP [*Wood et al.*, 1998; *Dirmeyer et al.*, 1999].

[36] The partitioning can be validated over subregions of the quadrants for which basin-observed streamflow is available. Figure 3b, which uses observed streamflow from the 1145 assessment basins of section 2.4, is the counterpart to Figure 3a obtained by area averaging the NLDAS observed precipitation and LSM simulated evaporation and runoff only over the quadrant subarea spanned by these basins (Figure 3b inset). For each basin, observed streamflow (m^3s^{-1}) is converted to mean-annual total discharge (m^3), in turn converted (using the basin area) to area-average mean-annual runoff (mm) for the basin. The plus symbol in Figure 3b depicts the area-mean of this observed runoff over the same quadrant subarea. Projecting the plus symbol onto the y axis yields the budget-based estimate of the area-average mean annual evaporation.

[37] The reliability of the observation-based evaporation estimate depends on the reliability of both the observed streamflow (high reliability) and the NLDAS precipitation forcing. Over NE and SE, which manifest relatively flat terrain and good density of precipitation gages, we trust the precipitation analysis and the estimates of evaporation there. For NE and SE, one sees in Figure 3b that evaporation and

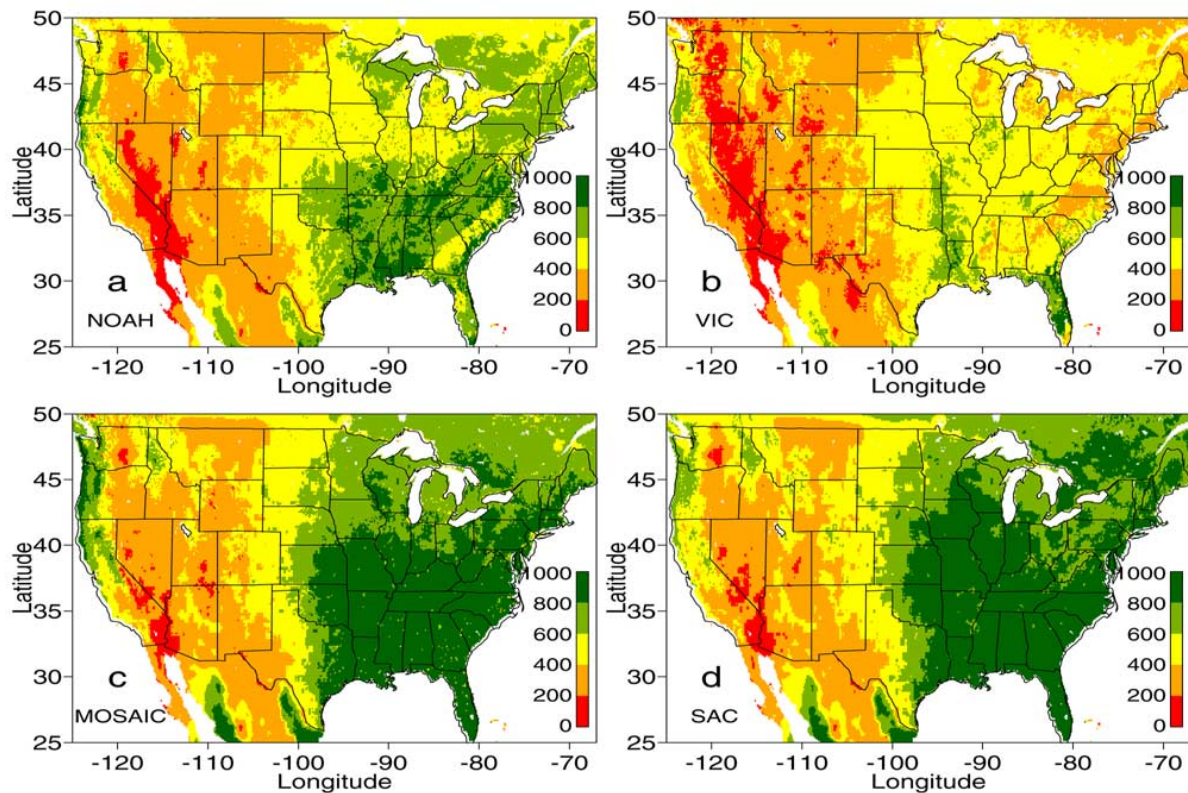


Figure 2. Mean annual evaporation (mm/year) in NLDAS from (a) Noah, (b) VIC, (c) Mosaic, and (d) SAC for 1 October 1997 to 30 September 1999.

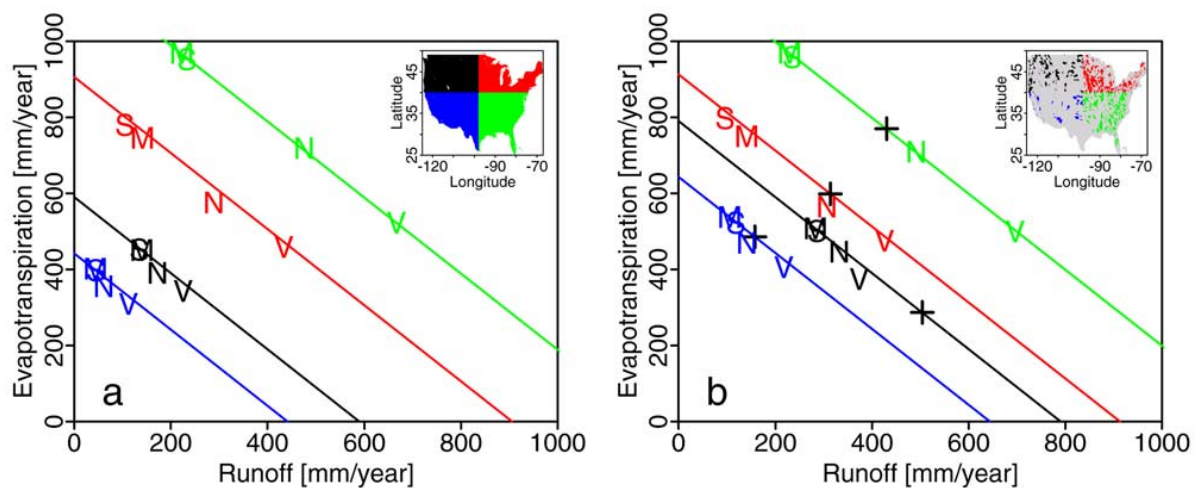


Figure 3. (a) Partitioning of mean annual area-mean precipitation (diagonal, mm/year) into mean annual area-mean runoff (x axis, mm/year) and evaporation (y axis, mm/year) for the CONUS quadrants (inset) of NW (black), NE (red), SW (blue), and SE (green) by Noah (N), VIC (V), Mosaic (M), and SAC (S) for 1 October 1997 to 30 September 1999. Model symbols below diagonal indicate (negligible) positive storage change. (b) As in Figure 3a, except area-mean is for subarea of basin set depicted in inset for each quadrant (same basins as in Figure 4) and the plus symbol depicts observed mean annual area-mean runoff for the same subarea.

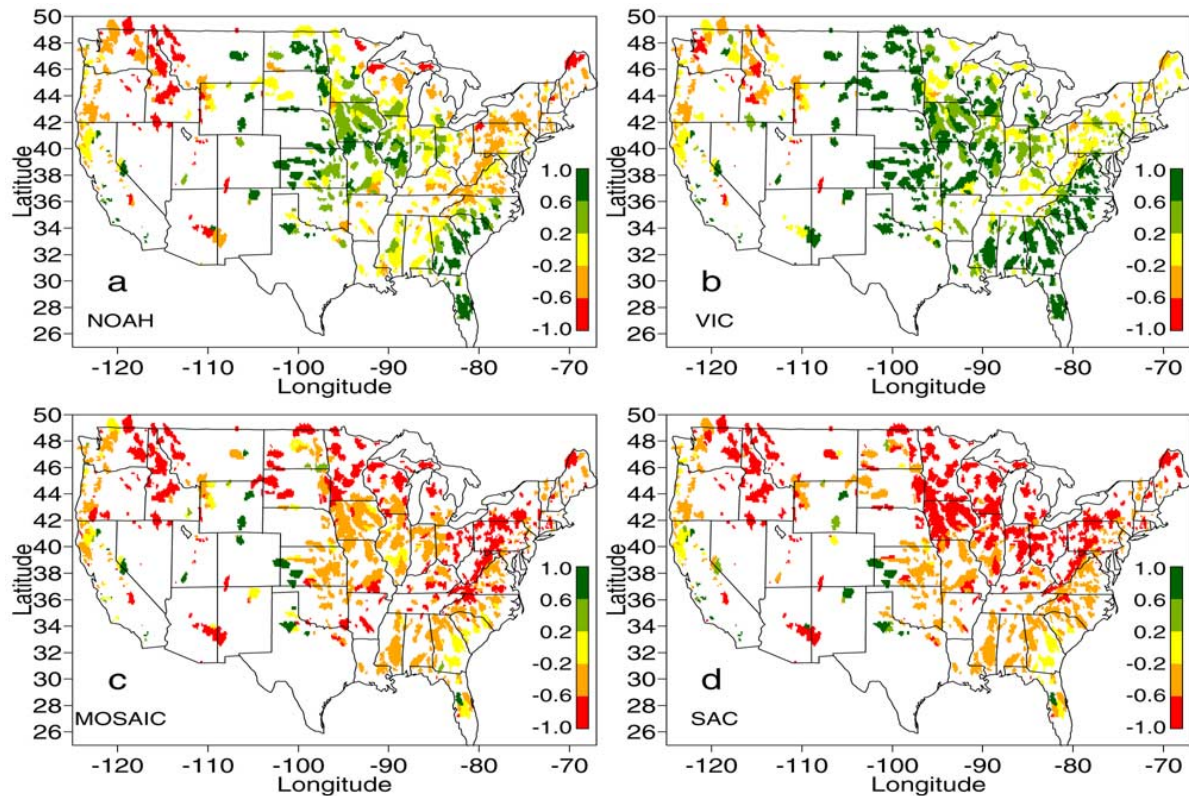


Figure 4. Relative bias [(model - observed)/observed] of mean annual runoff for selected basins for 1 October 1997 to 30 September 1999 from (a) Noah, (b) VIC, (c) Mosaic, and (d) SAC. Observed runoff for a given basin is calculated from basin area and observed basin-outlet stream discharge provided by USGS.

runoff of Noah are close to observed, while Mosaic and SAC show large biases of high evaporation and low runoff, with VIC yielding the reverse: large biases of low evaporation and high runoff. Figure 4 shows the variability of relative runoff bias [(model-observed)/observed] across the basins. Over NE and SE, Mosaic and SAC show similar patterns of pervasive underestimates of runoff, exceeding 60% (dark red) in the Appalachians and upper Midwest, while VIC runoff is highly overestimated (except for a corridor of near neutral bias west of the Appalachians), often by more than 60% (dark green) in the southeast and Midwest. Noah has the least runoff bias over NE and SE, with a more balanced likelihood of positive or negative bias, and fewer basins of (dark green) large bias.

[38] In NW, NLDAS precipitation has a large low bias (see section 3.4), owing to mountainous terrain, sparse precipitation gages, and lack of an adjustment to PRISM [Daly *et al.*, 1994] in the NLDAS precipitation analysis. This low bias thwarts the reliability of the NW budget-based evaporation estimate in Figure 3b and causes large low bias in runoff in all the LSMs over NW in Figures 3b and 4. For example, the LSMs substantially underestimate runoff in the Northern Rockies (dark red), though VIC yields much smaller bias there for reasons given in section 3.4. Last, over SW, in addition to precipitation gages being sparse, the number of assessment basins there in Figure 4 is small,

casting uncertainty on observation-based SW evaporation estimates.

3.2. Monthly Water Budget and Soil Moisture Change

[39] In studying nonannual water budgets, storage changes of soil moisture and snowpack are important. Thus, on monthly scales, we apply the full surface water budget equation given by

$$\frac{dS_1}{dt} + \frac{dS_2}{dt} + \frac{dS_n}{dt} = P - E - R_1 - R_2. \quad (1)$$

Each term is the area average, of storage change (left side) and accumulations (right side), of water mass per unit area (kg/m^2) per month, or depth of water (mm) per month. S_1 and S_2 are soil moisture stores of an upper and lower zone (defined later), respectively, and S_n is snowpack storage. P , E , R_1 and R_2 are precipitation, evaporation, surface runoff, and subsurface runoff, respectively. On monthly scales, canopy interception storage change is negligible and omitted in equation (1). Figure 5 gives the time series of area-average monthly evaporation E for each CONUS quadrant for the control runs.

[40] We focus here on the eastern quadrants, where evaporation in section 3.1 showed the most disagreement. During middle and late summer, Mosaic clearly has the

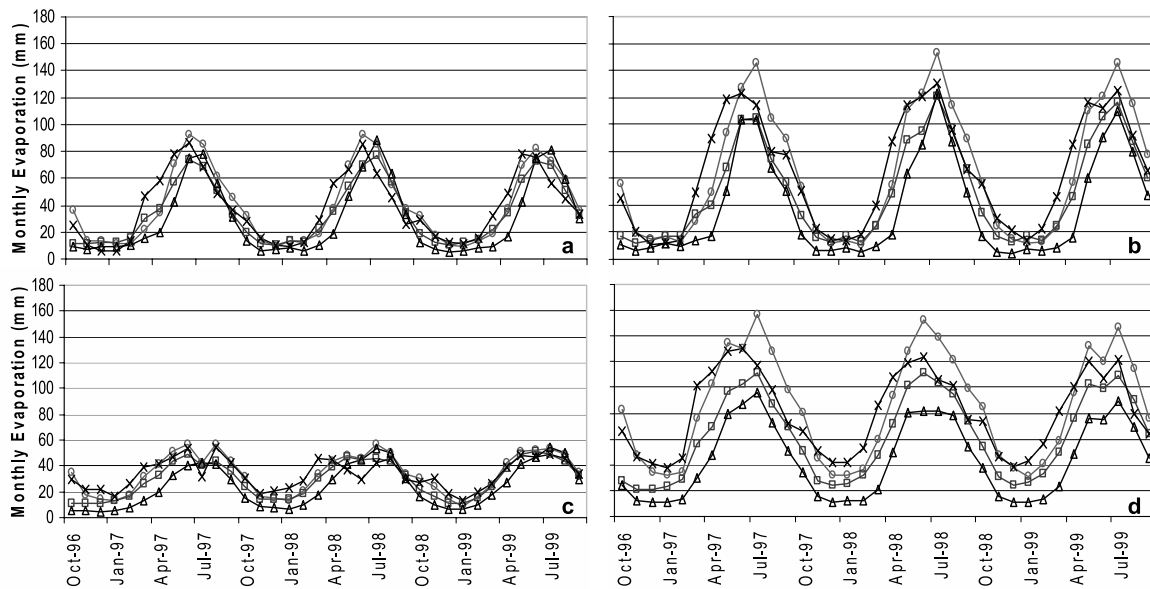


Figure 5. Time series of area-averaged monthly evaporation (mm/month) in NLDAS for Noah (squares), VIC (triangles) Mosaic (circles), and SAC (crosses) over the four CONUS quadrants of (a) NW, (b) NE, (c) SW, and (d) SE for October 1996 to September 1999.

highest evaporation, while SAC has the highest in winter and spring. Noah evaporation generally falls between that of Mosaic and VIC in the warm season. VIC has the lowest evaporation in virtually every month in the vegetated eastern quadrants, consistent with the earlier annual results. Figures 5b and 5d strongly suggest that Mosaic and VIC manifest rather different canopy conductance. This is most evident in the SE during 1998, where VIC monthly evaporation tops off at 80 mm during May–August while that of Mosaic sustains 120–150 mm. (Signatures of this middle and late summer stress over vegetation in VIC compared to Mosaic and Noah is evident again later in Figure 13b.)

[41] Such large differences in warm season evaporation imply large warm season differences in soil moisture storage change. Figure 6 shows the time series of area-average monthly mean total column soil moisture ($S_1 + S_2$) for all four quadrants. Figure 7 is the companion depiction for root zone soil moisture for the vegetated eastern quadrants. (Figure 7 excludes SAC, which has no explicit root zone.) The emphasis below is on years 2–3 of Figures 6 and 7, as nontrivial spin-up is evident in year one of Figure 6, though more so in Noah and VIC. After year one, spin-up is essentially complete (though small spin-up may still be occurring in NW in Noah).

[42] Focusing then on years 2–3, inspection of Figures 5–7 reveals the following: (1) very different levels of time-mean total soil moisture across the models, ranging in the SE from about 325 mm for VIC and SAC to 550 mm for Mosaic and 650 mm for Noah; (2) more similarity, yet important differences, in annual-cycle amplitude (seasonal change) of total soil moisture among the models, ranging in the SE from about 100–150 mm for Noah, VIC, and SAC to a notable high of 230 mm for Mosaic; (3) larger differences among the models over the wetter eastern quadrants

than the drier western quadrants, in both total soil moisture and its seasonal change or range; (4) among the three models with a root zone, the contribution of the subroot zone to the change in total soil moisture varies widely, e.g. over the SE it is very large in Mosaic (about 180 of 230 mm), moderate in Noah (about 70 of 140 mm), and small in VIC (about 10 of 100 mm); (5) the model with the highest level of total soil moisture is not the model with the largest seasonal change in soil moisture, nor the model with the largest monthly/annual evaporation; (6) a model with high annual evaporation is not necessarily a model having a high annual range of soil moisture storage (as the counterpart of SAC illustrates).

[43] Intriguingly in Figure 6, while VIC and Noah have very different levels of total soil moisture in the SE and NE, they have more similar magnitudes of soil moisture in the root zone and seasonal change of soil moisture in the root zone in Figure 7. Thus the moisture source for the higher summer evaporation in Noah than VIC is Noah's subroot zone, consistent with Noah usually having a deeper total soil column (maximum storage capacity) than VIC, as shown in SD-N.

[44] Recalling that Mosaic executions apply a shallow root depth of 0.4 m, compared to 1–2 m in Noah and 1.35–3 m in VIC, it is a paradox in Figures 5 and 6 that Mosaic in SE and NE has the highest warm season evaporation and highest warm season change in total soil moisture. Comparing Mosaic's annual cycle amplitude in Figures 6b and 6d with Figure 7 provides the answer by revealing that Mosaic's subroot zone accounts for the bulk of Mosaic's annual storage range. Later, we show that Mosaic develops vigorous upward diffusion of water to its root zone from its subroot zone during the warm season. Mosaic's diffusion is not a dominant process in

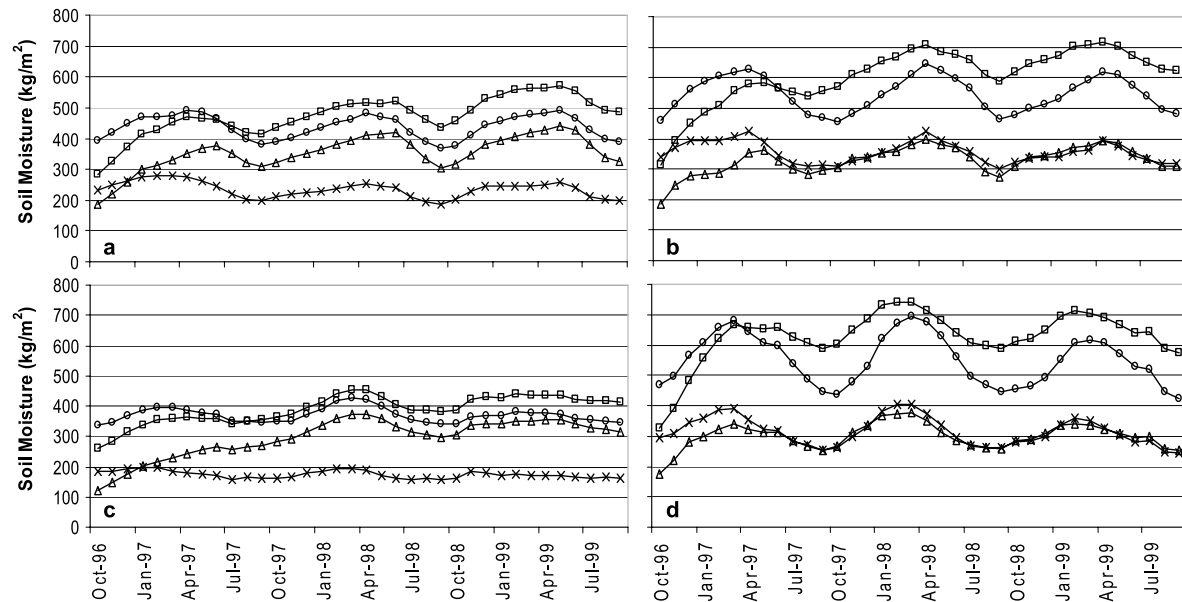


Figure 6. Time series of area-averaged monthly mean total column soil moisture (mm) in NLDAS for Noah (squares), VIC (triangles), Mosaic (circles), and SAC (crosses) for the CONUS quadrants of (a) NW, (b) NE, (c) SW, and (d) SE for October 1996 to September 1999.

the arid west, because root zone and subroot zone there are both typically dry.

[45] Figure 8 shows the May through September (nominal soil dry-down season) change in total soil moisture for 1999. Positive values denote soil drying. As an aside, the soil moistening (negative change: dark red) of Florida, south Texas, and Arizona/New Mexico is out of phase with the dry-down over the bulk of the CONUS, because their “wet season” is summer. Also, the Atlantic coastal states show moistening from two coastal hurricanes in September 1999. The hallmark of Figure 7 is vivid model differences in total soil moisture depletion over the southern Mississippi basin, which experienced significantly below normal precipitation during this period (not shown). Mosaic shows the largest depletion there and the largest evaporation (Figure 5d), followed by SAC and then Noah, with VIC showing the

least depletion and lowest evaporation (Figure 5d). Together, Figures 5d, 6d, 7b, and 8 illustrate that over vegetated areas with a warm-season precipitation shortage, Mosaic taps notably more than the other LSMs from its deep (subroot) soil moisture to sustain evaporation, though too much so (SE in Figure 3b).

[46] Figure 8 shows large intermodel variation in the east-west gradient across CONUS of seasonal range in total soil moisture storage ($S_1 + S_2$). The study by *Schaake et al.* [2004] (SD-N) depicts this intermodel difference as a function of basin climatology, in terms of the P/PE ratio for the 12 CONUS RFCs. As one can infer from east-west inspection along latitude 37° N in Figure 8, SD-N finds that VIC’s storage range is relatively invariant with respect to east-west gradient in P/PE between arid and moist regions, while SAC and Noah show more variability, and Mosaic the

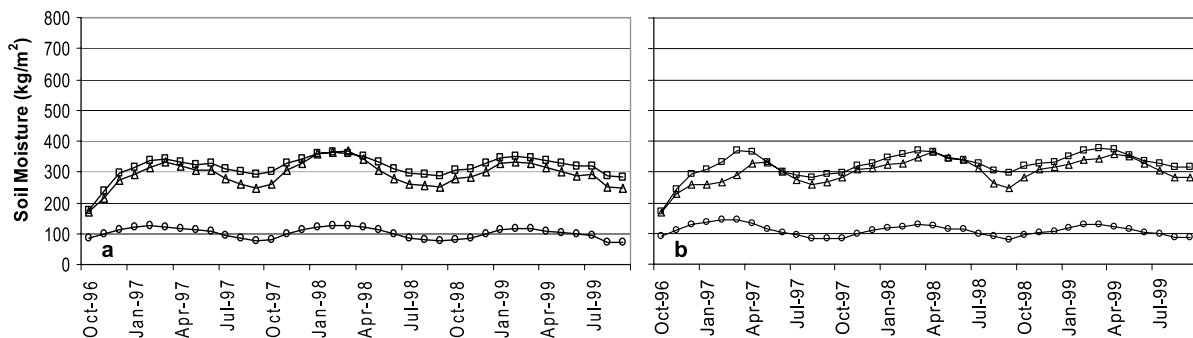


Figure 7. Time series of area-averaged monthly mean root zone soil moisture (mm) in NLDAS for Noah (squares), VIC (triangles), and Mosaic (circles) for the CONUS quadrants of (a) NE and (b) SE for October 1996 to September 1999.

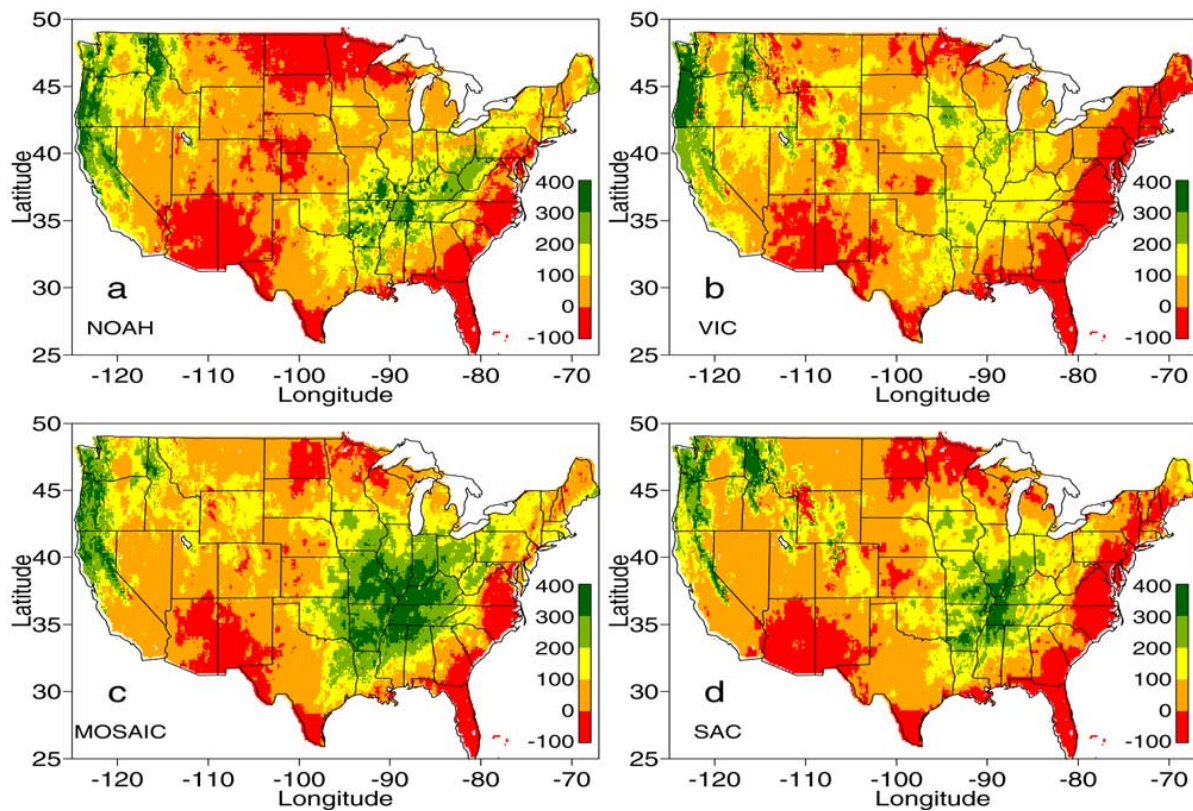


Figure 8. Warm season storage change (mm) of total column soil moisture, from difference of total column soil moisture (mm) of 30 April minus that of 30 September of 1999 at 23 UTC for (a) Noah, (b) VIC, (c) Mosaic, and (d) SAC. All colors (except dark red) are positive and denote a net drying during the period.

most variability. SD-N emphasizes that a given model's dynamic storage range in a given region is not reliably inferred from the model's maximum water holding capacity, but is rather the result of more complex interplay between a region's climatology and major facets of a model's physics. *Koster and Milly* [1997] show that a model's dynamic range of soil moisture is highly controlled by interaction between the model's runoff and evaporation formulations and the functional dependence of these formulations on the model's soil moisture state.

[47] We turn now to Figure 9, which depicts the annual cycle time series of every term in equation (1). Here storage S_1 is defined as the top two soil layers in Mosaic, Noah, and VIC, and the top storage reservoir in SAC and S_2 represents all remaining soil storage. The difference between the solid black line (total precipitation P) and dashed black line (liquid precipitation) is the snowfall. The red triangle is $P - dS_n/dt$ and the red triangle's departure below (above) the solid black line (P) equates to the monthly increase (decrease) in snowpack depth. In the snow season, the red triangle coincides with P if monthly snowfall is balanced by the sum of monthly snowmelt and sublimation, yielding zero net monthly change in snowpack.

[48] In Figure 9 we first examine region SE, which has negligible snow and rather small month-to-month changes in precipitation (in percent terms). Thus the annual cycle of

the SE water budget is driven mainly by PE. The model differences in runoff response are vivid. Mosaic's dominance in soil moisture depletion in the warm season is mirrored by Mosaic having the largest soil recharge during November–February, leaving rather less precipitation available during cool months for runoff. SAC also has less cool season runoff than Noah or VIC, though not so much from high storage recharge as in Mosaic, but rather from SAC having the highest cool season evaporation. Noah and VIC have larger total runoff than Mosaic and SAC in most every month, with VIC having the notably largest runoff (mostly subsurface) throughout the fall, winter, and spring, as VIC requires less cool-season soil recharge to replenish its smaller summer depletion. In NE in Figure 9, the above tendencies in SE continue to hold in a broad sense, but other signatures arise from less precipitation in the cool season, greater monthly variability of precipitation in the warm season, and nonnegligible snowpack processes. Lower precipitation in the cool season compared to SE results in less cool season runoff in all four models, but in general VIC still produces the most monthly runoff, followed by Noah, then Mosaic and SAC. The lack of SAC runoff all year in NE is noteworthy, as is the high SAC evaporation in spring.

[49] Given that SAC in NLDAS takes its PE forcing from Noah PE output, the higher March–April evaporation in

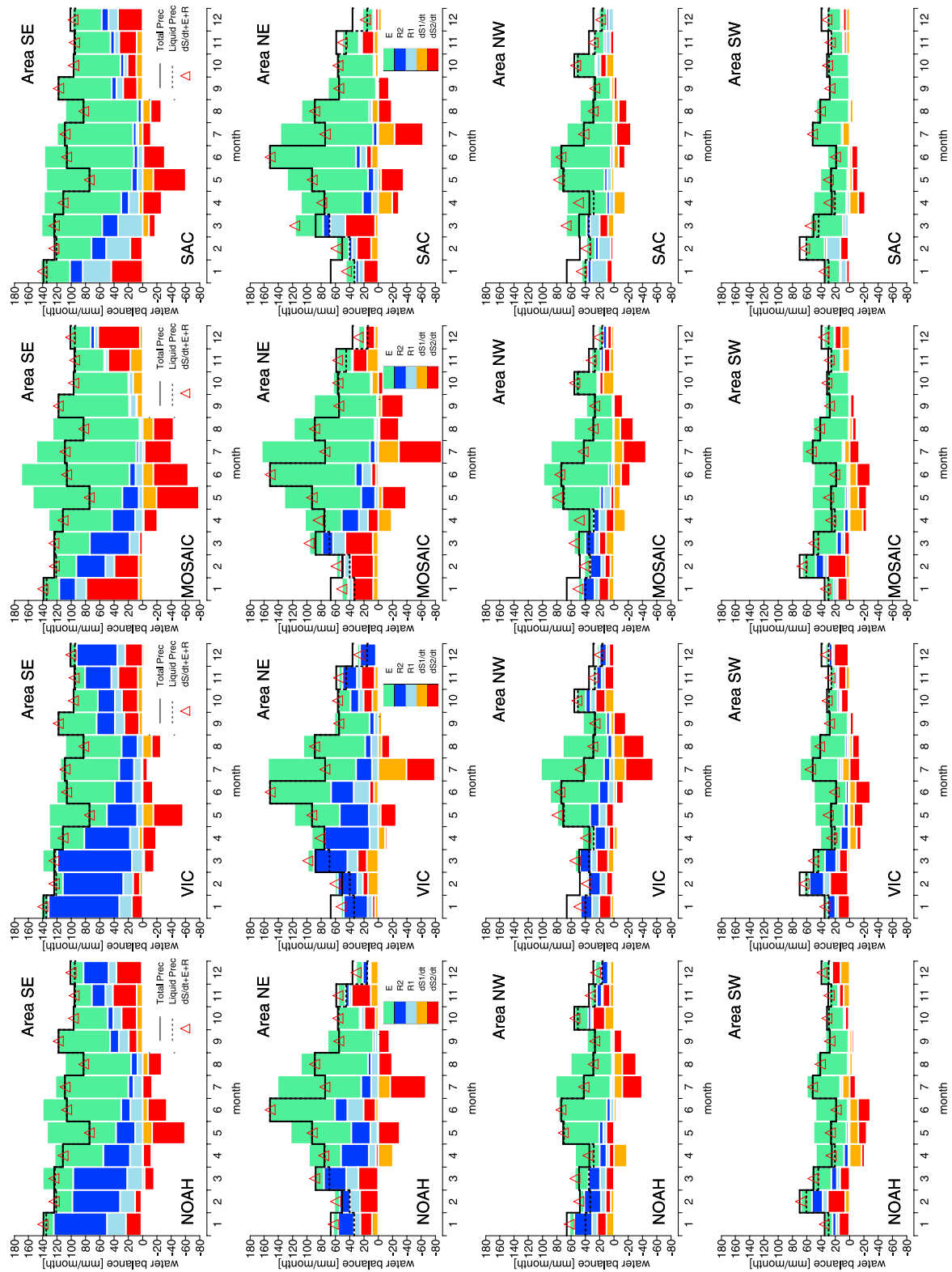


Figure 9.

SAC versus Noah in NE and NW is instructive, as vegetation greenness and hence transpiration are low then in both quadrants in Noah. Noah, Mosaic and VIC draw soil moisture for direct evaporation only from their first soil layer (0.1 m). This one layer is likely unable to sustain as high evaporation rates as SAC, which taps into both of its top two soil storages in response to PE demand. This structural difference in direct evaporation between SAC and the SVATS likely explains SAC's dominance in cool season evaporation (especially winter and spring), which contributes in turn to SAC's high mean annual evaporation. Thus, while Figures 2 and 3 show SAC and Mosaic having similar annual evaporation, the monthly water budget in Figures 5 and 9 (notably for NE and SE) shows SAC and Mosaic having rather different temporal character in evaporation, with SAC (Mosaic) having more in spring (summer).

[50] Moreover in Figure 9, the high precipitation in June in NE followed by a large drop of precipitation in July yields another vivid example of Mosaic's ability to draw upon its deep soil moisture. In the face of the steep drop in July precipitation, Mosaic is the only model in NE to yield July evaporation larger than June evaporation, and it does so via the largest July soil depletion. Following in August and September, Mosaic still sustains higher evaporation than the other models, despite its larger July storage depletion. Mosaic's layer 3 (subroot) storage change was dissected for July 1998 at one grid cell near 45°N latitude and 92°W longitude. For this month at this point, the Mosaic drainage (R_2) out the bottom of layer 3 ranged from 4 to 10 mm across its tiles there, while the upward diffusion of water to layer 2 ranged from 116 to 124 mm, confirming Mosaic's vigorous supply of subroot water to the root zone by diffusion. Though Noah also includes vertical diffusion of soil water, the magnitude is much larger in Mosaic.

[51] The contrast between models in Figure 9 is rather less in the western quadrants, wherein the warm-season water budget of the three SVAT models is quite similar, though SAC has less warm season evaporation and soil moisture depletion. The greater model similarity in SW and NW likely stems from the sparse vegetation (Figure 1a), whereby canopy conductance and root-zone processes are not dominant. In the cool season of NW and SW, VIC still tends to have the most runoff and is still dominated by subsurface runoff. Interestingly, SAC has virtually no subsurface runoff in NW and SW. One cool season contrast between models in NW (and NE) is the lower snowpack accumulation in Noah during winter. Comparison of the NW January water budget components of Noah with the other LSMs reveals model agreement in monthly snowfall amount (difference between solid and dashed lines), but Noah (unlike the accumulating snowpack in VIC, Mosaic and SAC) is melting and sublimating in the month about as much snow as it receives in snowfall, a topic revisited in section 3.4.

[52] We conclude this section with some discussion of the causes of the systematic biases noted in the models thus far.

The high bias in Mosaic evaporation over CONUS-east is most likely a result of the vigorous diffusion of water from the subroot zone to the root zone. The high bias in SAC evaporation over CONUS-east is likely a consequence of SAC's use in NLDAS of (1) uncalibrated a-priori parameters (section 2.3) and (2) PE from Noah. As a counterexample, in NWS operations at the RFCs, SAC inputs a NOAA monthly climatology of PE, derived from evaporation-pan measurements. This climatological PE is then scaled during SAC runs by a monthly fractional coefficient. This coefficient is a key calibration parameter allowed to be moderately larger or smaller than 1.0, as determined from SAC calibration runs over a catchment. No such coefficient had been derived to date for use in SAC large-scale runs over the CONUS-wide domain, either with Noah PE or NOAA climatological PE, and hence a universal coefficient of 1.0 was used SAC control runs in NLDAS. Moreover, the Noah PE is known to be higher than NOAA PE climatology. These two factors contribute to high SAC evaporation (E) in NLDAS, but not always the highest E , and typically less than the E in Mosaic in the warm season over nonsparse vegetation. Hence the SAC results here are not outliers and they represent important pathfinder runs of SAC executed over a national domain in semi-distributed mode with uncalibrated parameters. Since the SAC control runs here in NLDAS, NWS/OHD has derived a CONUS-wide field of the PE coefficient from the vegetation greenness database cited in Figure 1a. SAC experiments of this field in NLDAS are imminent.

[53] We last address the unexpectedly low bias in VIC evaporation in the results here over CONUS-east. Two separate but related VIC modeling efforts have been conducted over the NLDAS grid and terrain heights; specifically, the 3-year retrospective runs executed here with 1-hour time steps, and the 50-year retrospective runs reported by Maurer *et al.* [2002], executed with 3-hour time steps and with different sources for the surface forcing. The 3-year VIC runs here use essentially the same parameters as the VIC runs of Maurer *et al.* [2002]. Yet two significant differences were hourly temporal disaggregation of the daily precipitation and subgrid spatial disaggregation within a grid box, both used in the VIC runs here but not in those of Maurer *et al.* Not having the advantage of hourly radar-anchored precipitation analyses, Maurer *et al.* used uniform distribution of the daily precipitation throughout the day and within each 3-hour forcing interval. Maurer *et al.* [2002, Figure 2] analyzed the impact of this uniform distribution versus nonuniform disaggregation to 3-hour time steps and the results showed that the differences for the subregion analyzed (Lower Mississippi basin) were modest. Nonetheless, subsequent comparisons between the retrospective runs of Maurer *et al.* and the 3-year retrospective runs of VIC here show that the combined and interactive impact of the three factors of temporal disaggregation, spatial disaggregation, and 1-hour versus 3-hour time steps can be significantly larger than suggested by the Lower Mississippi tests of the temporal

Figure 9. Monthly water budget in NLDAS for October 1997 to September 1998. In order, columns 1–4 are Noah, VIC, Mosaic, and SAC, and rows 1–4 are quadrants SE, NE, NW, and SW. Colors depict terms (mm/month) in equation (1): dS_1/dt (orange), dS_2/dt (red), R_1 (light blue), R_2 (dark blue), and E (green). Black solid line is total precipitation P (mm/month); black dashed line is liquid precipitation (mm/month). See text for definition of red triangles.

disaggregation alone. The differences (shown for a transect across the eastern and central United States at <http://www.hydro.washington.edu/Lettenmaier/Models/VIC/VIChome.html>) are evident in portions of the country with a high fraction of convective precipitation and full canopy cover (e.g., CONUS-east summer). More study of the differences, and development of parameter transformations to account for disaggregation and time step differences will be addressed in a future paper.

3.3. Regional Validation of Soil Moisture

[54] This section presents validation of NLDAS soil moisture over (1) Illinois from *Schaake et al.* [2004] (SD-N) and (2) Oklahoma from *Robock et al.* [2003] (RL-N). SD-N evaluated NLDAS soil moisture at 17 of 18 sites of the Illinois State Water Survey [Hollinger and Isard, 1994], which measures soil moisture at 11 levels down to 2 m. Figure 10 shows the resulting two-year scatterplot (and best-fit linear line) of model versus observed, state-wide average, total-column soil moisture over 2-m at bimonthly intervals. For VIC, two best-fit lines for northwest and southwest Illinois were required, because past VIC calibration yielded rather different soil moisture storage capacities in these regions. In Figure 10, a best-fit line having slope greater than one indicates a storage range greater than observed. Mosaic yields a storage range greater than the other models (as in sections 3.1–3.2) and about 50% more than observed. Noah and SAC agree with observations in both storage range and storage magnitude. VIC also shows good storage range over its two regions, but storage magnitude lower than observed. In addition to Illinois validation, SD-N validates NLDAS soil moisture storage range in the Arkansas-Red River basin (not shown). More broadly, SD-N intercompares the soil moisture storage capacity (explicit capacity) and storage range (“active” capacity) of the four LSMs across the entire NLDAS domain, including mean statistical properties and spatial variation. The findings reveal significant LSM differences in soil moisture, as did section 3.2. Such differences challenge modelers using soil moisture from one LSM to initialize another.

[55] Similar contrasts in soil moisture between LSMs themselves and between observations and LSMs are found over Oklahoma by *Robock et al.* [2003] (RL-N), which included close scrutiny at individual stations. In-situ observations of soil moisture have been installed and calibrated at 72 Oklahoma Mesonet stations by the Oklahoma Climatological Survey. Figure 9 of RL-N (not shown) depicts a 21-month time series during 1998–1999 of observed and NLDAS-simulated daily mean, 0–40 cm total soil moisture averaged over all 72 Mesonet stations. The time series show substantial differences in soil moisture magnitude among the LSMs and between the LSMs and observations, with VIC showing the best agreement with the observations. Nonetheless, there is rather good agreement among the models and between models and observations in the soil moisture changes in time, a theme cited earlier here with Figure 6. RN-L also performs and evaluates important sensitivity tests in the three SVAT models at many of the OU Mesonet soil moisture measuring stations, wherein model soil type is set to match the station-reported soil type and the

assigned soil parameters are unified across the models. This test improved a model’s performance if the parameters were not incompatible with previous model calibration. Since Noah is not substantially calibrated, the matching with local soil type and use of unified soil parameters did improve Noah’s soil moisture performance at the stations. In VIC, which has been regionally calibrated to streamflow over several large basins, including the Arkansas-Red river basin [Abdulla et al., 1996], the use here over Oklahoma of local station-matching soil types and unified soil parameters had less consistent impact on improving model agreement with the soil moisture measurements, showing more variance of positive or negative impact depending on station. Thus soil parameter changes in an LSM to match local site characteristics may degrade LSM performance at those sites, especially if the changes disturb an LSM’s calibration legacy.

3.4. Validation of Snow Cover and Snowpack Content

[56] This section presents results from *Pan et al.* [2003] (PS-N) and *Sheffield et al.* [2003] (SP-N), which perform large-scale assessment of NLDAS snowpack water equivalent (SWE) and snow cover extent (SCE), respectively. First, we summarize how the four LSMs treat snowpack. SAC simulates only the snowpack water balance, via the SNOW17 model [Anderson, 1973], which includes snowfall and snowmelt but not sublimation. Snowmelt is determined empirically via an index-method based on maximum-minimum daily air temperature. VIC, Noah and Mosaic also simulate the snowpack water balance, including sublimation as well as snowfall and snowmelt, plus the snowpack energy balance (net radiation, sensible, latent and subsurface heat fluxes, phase-change heat sources/sinks). The treatments for processes such as snow cover fraction, snow albedo and retention/refreezing of snowmelt differ among the models. SAC, VIC, and Noah explicitly account for retention of liquid water (snowmelt or rainfall) in the snowpack, but only VIC and Noah allow refreezing. Mosaic and VIC carry explicit subgrid vegetation tiles, but only VIC carries added subgrid tiles for elevation, known as “elevation banding”, which is cited in PILPS studies as a key factor in VIC’s good snowpack simulations [Bowling et al., 2003; Boone et al., 2004]. For each subgrid tile, Mosaic and VIC carry separate water-energy balances and separate soil, vegetation and snow states. More on the model snow physics is provided by SP-N, and by *Wigmosta et al.* [1994], *Koster and Suarez* [1996], *Koren et al.* [1999], and *Anderson* [1973] for VIC, Mosaic, Noah, and SAC, respectively.

[57] There is no distinction between rainfall and snowfall in NLDAS precipitation forcing. This requires criteria to infer snowfall. The input precipitation at each model time step was assumed to be all rainfall for surface air temperature $>0^{\circ}\text{C}$ and all snowfall otherwise. This criterion does not guarantee identical snowfall in the models owing to different model time steps and VIC’s elevation banding. Noah and Mosaic interpolate hourly air temperature to their 15-min time steps, thus allowing rainfall and snowfall inside one hour, unlike the hourly steps of VIC and SAC. VIC adjusts the hourly air temperature to the elevations of its subgrid elevation bands, thus VIC

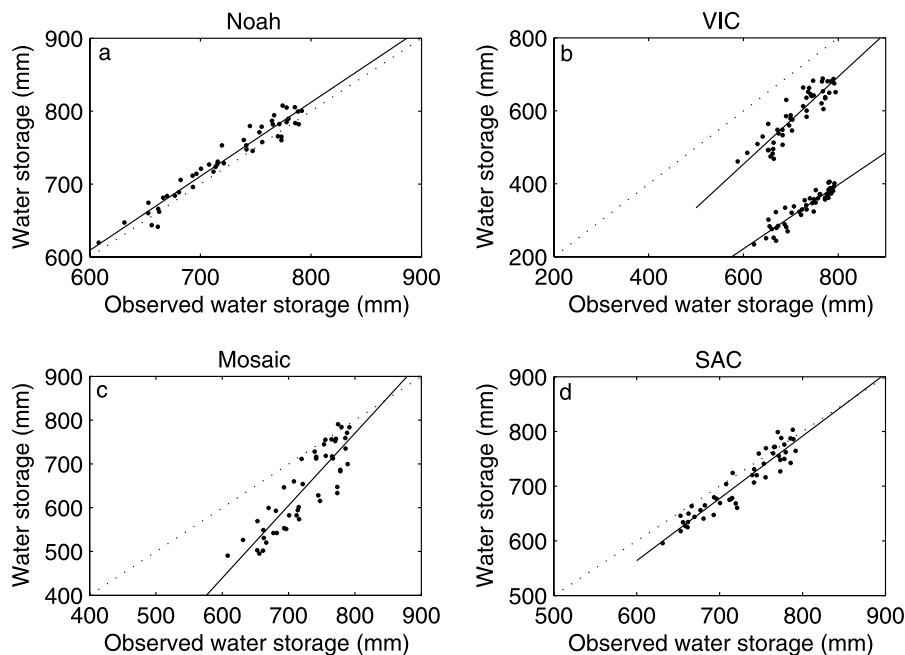


Figure 10. Comparison of NLDAS versus observed bimonthly total soil moisture (mm) in top 2 m, averaged over 17 sites throughout Illinois for October 1997 to September 1999 for (a) Noah, (b) VIC, (c) Mosaic, and (d) SAC. Note different x axis and y axis ranges. See text for discussion of two sets in Figure 10b.

allows both rainfall and snowfall inside a given grid cell for a given time step.

[58] PS-N validated NLDAS SWE simulations against NRCS SNOTEL stations west of 104°W. SNOTEL sites measure SWE, air temperature and precipitation every 15 min. The majority of SNOTEL elevations are above 1000 m, with mean elevation near 2500 m (see PS-N for references). Since NLDAS terrain resolution is 1/8°, comparing model SWE with point-wise SNOTEL is a challenge. Hence PS-N omitted use of SNOTEL sites whose elevation differs from that of the nearest NLDAS grid point by more than 50 m. This retains 110 SNOTEL sites.

[59] Figure 11 shows model versus observed mean-annual maximum SWE for the LSM control runs (and two VIC tests described later) at the SNOTEL sites by subset in four mountain ranges: 3 in Sierra Nevada (squares), 17 in Cascades (diamonds), 29 in southern Rockies (circles), 61 in northern Rockies (triangles). All the LSMs substantially underestimate maximum SWE in all four regions. Noah has the largest low bias and the lowest correlation. Mosaic also shows a rather low correlation. SAC and VIC have notably better bias and substantially higher correlation. The model with elevation tiling (VIC) yields the highest correlation, yet the simplest model (SAC) without elevation tiling or energy balance treatment is closely competitive.

[60] PS-N found the NLDAS precipitation forcing to be substantially low when compared to observed precipitation at all 110 SNOTEL sites. Sites with the highest observed precipitation are where NLDAS precipitation has the largest low bias. PS-N determined the linear regression between SNOTEL and NLDAS precipitation at the 110 sites to be $P_{\text{SNOTEL}} = 2.1693 P_{\text{NLDAS}}$, with an R^2 value of 0.64, revealing a factor-of-2 underestimation by NLDAS precip-

itation at the sites on average. The low precipitation bias is consistent with that anticipated from the low bias in annual streamflow in the LSMs in the Northwest in section 3.1 (Figure 4).

[61] PS-N executed two tests in VIC with two methods of bias-adjusted precipitation. Both tests executed VIC in the reduced-grid mode (section 2.1) at grid cells nearest the SNOTEL sites. Test 1 used NLDAS precipitation forcing scaled upward by a separate regional factor for the four mountain ranges, based on regional regression fit of the NLDAS and SNOTEL mean annual precipitation. In test 1 (Figure 11e), the model SWE bias is dramatically less, with the scatter rather evenly balanced about the 1:1 line. Yet the scatter remains substantial and the correlation is only modestly improved, as the regional scaling does not eliminate site-specific bias. VIC in test 2 (Figure 11f) was forced with site-specific adjusted NLDAS precipitation, scaled to match the observed annual total precipitation at each site. Test 2 yields a much smaller bias, substantially reduced scatter and increased correlation of $R^2 = 0.82$. PS-N also evaluated NLDAS air temperature bias at SNOTEL sites, finding it small in the cool season (exceptions at some stations), contributing much less to model SWE bias than precipitation. The high bias in NLDAS solar insolation over snow (section 2.2) also contributes to a low bias in SWE in the three SVAT models. The impact of this insolation bias on SWE has not been quantified yet.

[62] SP-N validated NLDAS simulations of areal fraction of snow cover extent (SCE) against the NESDIS operational, daily, 23-km, Northern Hemisphere snow cover product known as the Interactive Multisensor Snow (IMS) [Ramsay, 1998], viewable at <http://www.ssd.noaa.gov/PS/SNOW>. NESDIS analysts produce the IMS using an interactive

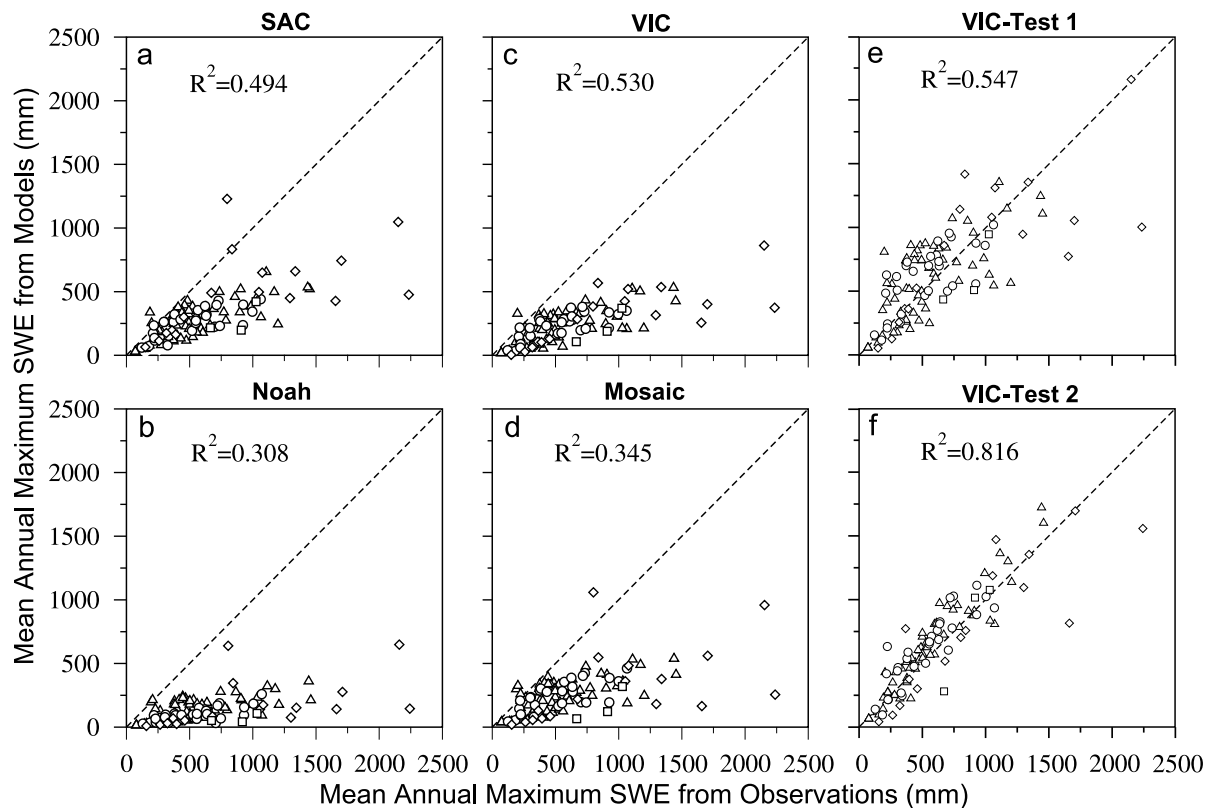


Figure 11. Comparison of mean annual maximum snow water-equivalent (SWE) during October 1996 to September 1999 between observations (x axis) and model simulations (y axis) at 110 SNOTEL sites for the control runs of (a) SAC, (b) Noah, (c) VIC, and (d) Mosaic, and two VIC tests runs forced with (e) regionally adjusted precipitation and (f) locally adjusted precipitation.

workstation to assess snow cover related visible, infrared, and microwave satellite products, as well as in situ snow depth observations. Details of mapping the 23-km IMS field to the NLDAS $1/8^\circ$ grid and mapping NLDAS snow cover fraction (0–1 range) to the IMS binary snow cover state (yes/no) are given in SP-N.

[63] All four LSMs diagnose SCE (0–1 fraction) as an empirical function of the model simulated SWE, but the treatments differ significantly, namely, the critical SWE needed to reach SCE = 1 and the form of the function relating SCE to SWE. SP-N provides details, but in broad terms SAC and Noah require relatively deep snowpack for high SCE values, while Mosaic requires notably less, and VIC very little. For a nondeep SWE value, VIC will yield the highest snow cover, followed by Mosaic, then Noah, and finally SAC. For illustration, assuming a nonforest vegetation type, a SWE value of 4 mm yields SCE = 1.0, 0.67, and 0.24 in VIC, Mosaic, and Noah, respectively. Correspondingly (shown later), VIC, Mosaic, and Noah in order yield the highest, intermediate, and lowest albedo over snow. (SAC requires no albedo.) The high SCE fractions in VIC arise in part because VIC assigns SCE = 1 over any tile with any nonzero SWE.

[64] SP-N carried out validation of SCE separately over the 8 of 12 CONUS RFCs that exhibit substantial winter snow cover. In general, all models simulate reasonably well the regional-scale spatial and seasonal dynamics of snow cover.

Yet systematic biases exist, with (on average over 8 RFCs) underestimation of SCE by Noah (–22.5%) and MOSAIC (–19.8%) and overestimation of SCE by VIC (22.3%), with SAC being essentially unbiased. The level of bias over individual RFC regions varies (see Figure 4 of SP-N). The more mountainous RFC regions (Northwest, Colorado, California-Nevada) show the largest model differences with IMS observations and between models. Here VIC further overestimates SCE, while Noah further underestimates SCE and manifests an early bias in spring snowmelt. VIC's high SCE bias in the west is surprising at first, since all four LSMs had notably low bias in SWE at SNOTEL sites in Figure 11, owing to the low bias in NLDAS precipitation. However, the number of pixels at SNOTEL sites in any RFC domain is relatively small, so the effect on the regional mean is difficult to judge. Yet we surmise that VIC's high SCE bias stems from VIC assigning SCE = 1 at any subgrid tile with any nonzero SWE, however small.

[65] The low bias in Noah SCE appears to result from not only the high SWE threshold required in Noah for high SCE, but also the low snow albedo in Noah and its positive feedback effect on the energetics of snowmelt. For one RFC domain, Figure 12 shows the time series of monthly domain-mean (1) snowmelt, (2) snow sublimation and (3) albedo from the four models (only snowmelt for SAC, as SAC excludes sublimation and albedo). There are large differences in snow albedo among the models, with Noah

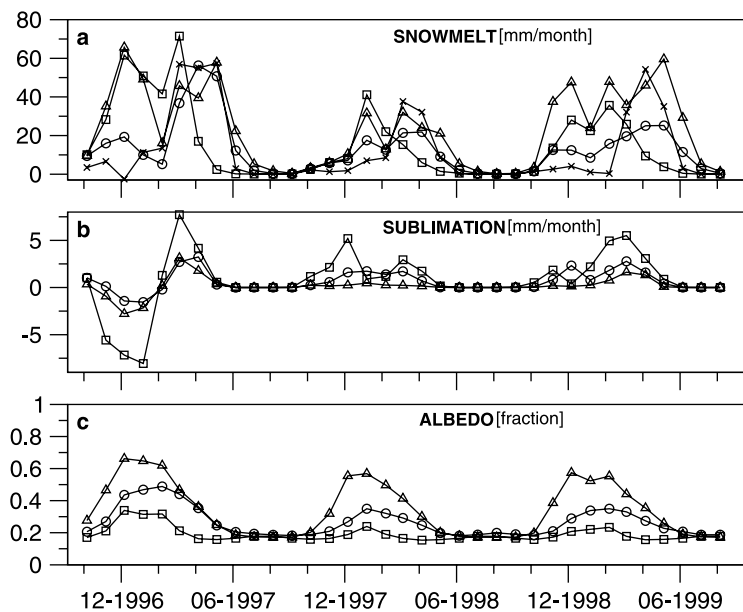


Figure 12. For the Northwest RFC domain, time series of monthly domain-mean (a) snowmelt (mm/month), (b) snow sublimation (mm/month), and (c) albedo in NLDAS for Noah (squares), VIC (triangles), Mosaic (circles), and SAC (crosses) for the time period October 1996 to September 1999. (Note different y axis range in top two panels.)

yielding the lowest (0.2–0.3), Mosaic being intermediate (0.3–0.5), and VIC yielding the highest (0.5–0.65). Not surprisingly then in Figure 12, Noah yields notably higher snowmelt and sublimation in early and midwinter, while VIC yields very little sublimation, with Mosaic in between. The larger midwinter snowpack sinks in Noah are consistent with Noah having the greatest low bias in annual maximum SWE in Figure 11. By spring, Noah's snowmelt and sublimation reduce to almost zero, as much of Noah's snowpack has already melted or sublimated earlier in the winter, while VIC and SAC have the largest spring snowmelt volumes, in part because they had the smallest sublimation sinks during early and midwinter. SAC and Mosaic tend to have higher melt in the spring than the winter months while VIC melts at a more quasi-steady rate throughout the winter and spring.

[66] A low bias in snow albedo is vulnerable to positive feedback problems in the surface energy balance. A low bias in albedo contributes to a high bias in net solar insolation, which melts more snow and reduces the snow cover, yielding still lower albedo and so on. The feedback is amplified by the high bias here in the incoming solar insolation over snow. High albedo and high snow cover, such as in VIC, is also vulnerable to positive feedback in the opposite direction, but the high albedo in VIC likely acted to offset the high bias in NLDAS incoming insolation over snow. A counterpart to such feedback risk is the simplicity of snow physics in the SAC/SNOW17 model. Its snowpack predictions perform rather well when assessed at the large regional scales of NLDAS. One reason is SAC's simple temperature-index approach to snowmelt, which avoids the feedback loops that can plague energy balance models over snow. The recent PILPS high latitude modeling experiments [Bowling *et al.*, 2003] found large differences in snow

ablation and snowmelt among 21 LSMs and also concluded that differences in model parameterizations of albedo and SCE have large effects on energy available to the snowpack.

4. Validation of NLDAS Surface Energy Fluxes and LST

4.1. In Situ Validation of Surface Energy Fluxes Over the Southern Great Plains (SGP)

[67] This section presents key results from the surface energy-validation portion of the Robock *et al.* [2003] study (RL-N), which validates energy fluxes during January 1998 to September 1999 using the 24 extended facility (EF) flux stations of the ARM/CART network in Oklahoma and Kansas. The results include the three SVATS (Noah, Mosaic, VIC) but not SAC, as SAC omits the physics of surface energy balance. Multistation spatial averaging and hourly temporal averaging are used to reduce the influence of scale differences between NLDAS grid cells (~ 12 km) and pointwise flux stations. Radiation fluxes were averaged over the 22 of 24 EF stations using Solar and Infrared Radiation Station instruments (SIRS). Heat fluxes were averaged over the 14 of 24 stations using Energy Balance Bowen Ratio systems (EBBR). Though the energy budget is not exactly closed in the averaging, discrepancy is less than 20 W m^{-2} in most months.

[68] The ARM network spans central and northern Oklahoma and southern Kansas (see map in RL-N). It is instructive to consider, a priori, what one would expect to uncover in the model surface energy budgets in this region based on the annual water budget validation in section 3.1. This region is a transition zone between quadrants SE and SW in Figure 3. Yet over the bulk of this region, Figures 1a and 1b show that warm-season green vegetation fraction is not sparse, ranging

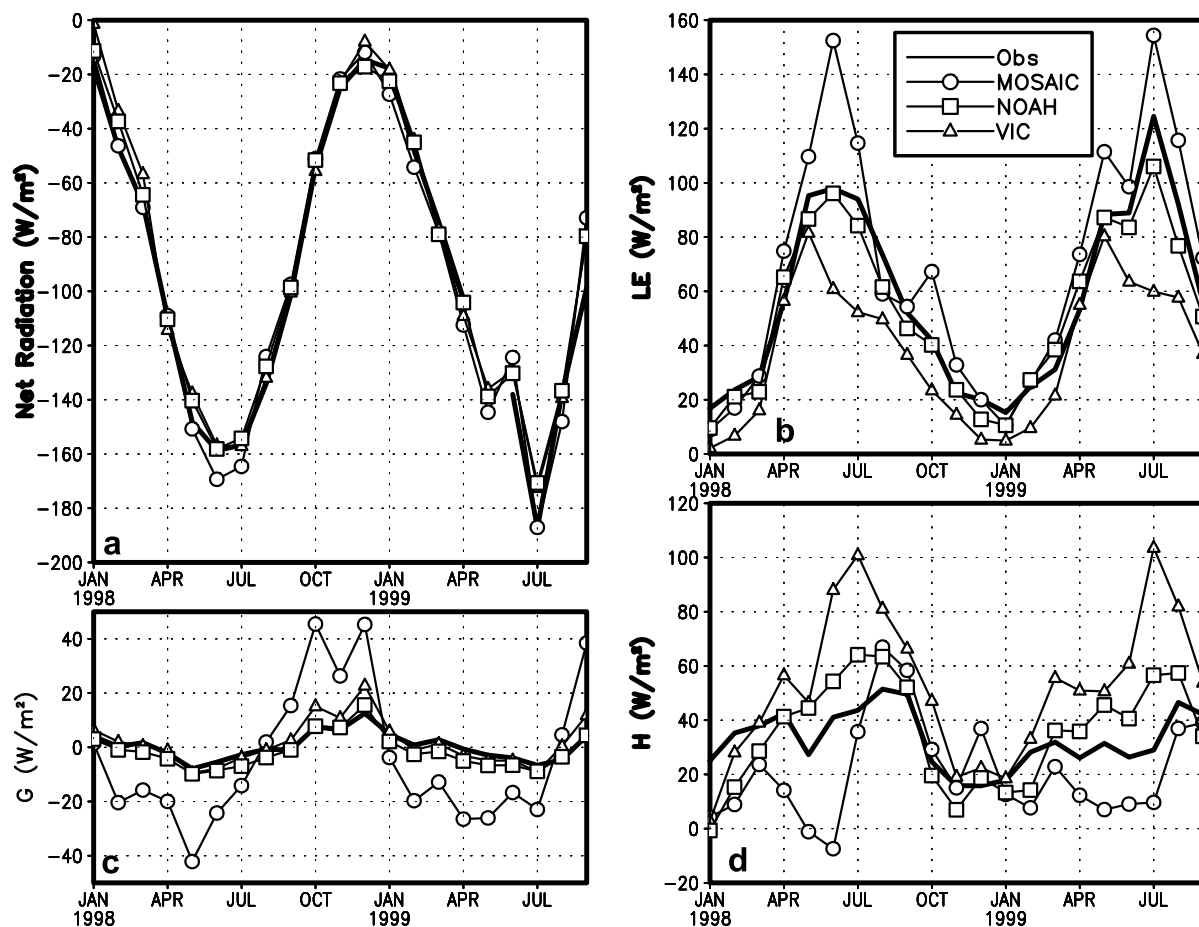


Figure 13. Time series of monthly mean surface energy fluxes (W m^{-2}) of (a) net radiation R , (b) latent heat LE , (c) ground heat G , and (d) sensible heat H averaged over the ARM/CART sites during January 1998 to September 1999 from observations (bold line, no symbols) and control runs for Noah (squares), VIC (triangles), and Mosaic (circles). The y axis range varies among panels. Positive flux is heat sink to surface, except for G .

from 0.4–0.8, and the mean annual precipitation exceeds 800 mm during the study period. Hence inspection of Figure 1 leads one to expect the mean surface water and energy budget of the region to be more similar in nature to that of SE than SW. Therefore, given that warm-season evaporation dominates annual evaporation, the following expectation is warranted: latent heat fluxes of the three SVATS over the ARM network are likely to reflect the model evaporation biases established over the SE quadrant in Figure 3b, namely, that Mosaic, Noah, and VIC will manifest latent heat flux averages that are substantially higher than observed, close to observed but somewhat low, and substantially less than observed, respectively. The validations below in Figures 13 and 14 confirm this expectation.

[69] Figure 13 gives a 21-month time series of monthly mean observed versus modeled surface energy fluxes of the NLDAS control runs, including net radiation (R), latent heat flux (LE), sensible heat flux (H), and ground heat flux (G). Figure 14 shows corresponding monthly mean diurnal cycles for the two months of July and April 1999. Columns 1 and 2 of Figure 14 depict only control-run results, while columns 3 and 4 additionally show experiments described later. There is

rather good agreement between observed and simulated R in all the models. Figure 14 shows some small model phase errors in R in Noah and VIC, which are further diagnosed in RL-N.

[70] Of more interest are the several situations of substantial bias in simulated LE , H , and G , examined first in Figure 13. As correctly anticipated above, in spring and summer, Mosaic has a substantial high bias in LE and, correspondingly, a substantial low bias in H . VIC has a substantial low bias in LE and high bias in H throughout most of the year (except spring), while Noah shows much smaller bias in LE (slightly low in warm season) and H (modestly high in warm season). The LE results here agree with the sign and relative magnitude of the model evaporation biases inferred in section 3.1. The counterpart to evaporation bias in section 3.1 was runoff bias of opposite sign. Here the analog counterpart to LE bias is H bias of opposite sign. The warm season LE and H biases of Mosaic, VIC, and Noah are highlighted further in the July 1999 midday biases in the diurnal cycles of Figure 14, again showing Mosaic with a significant positive bias in daytime LE and substantial negative bias in daytime H , thus very low Bowen ratio ($BR =$

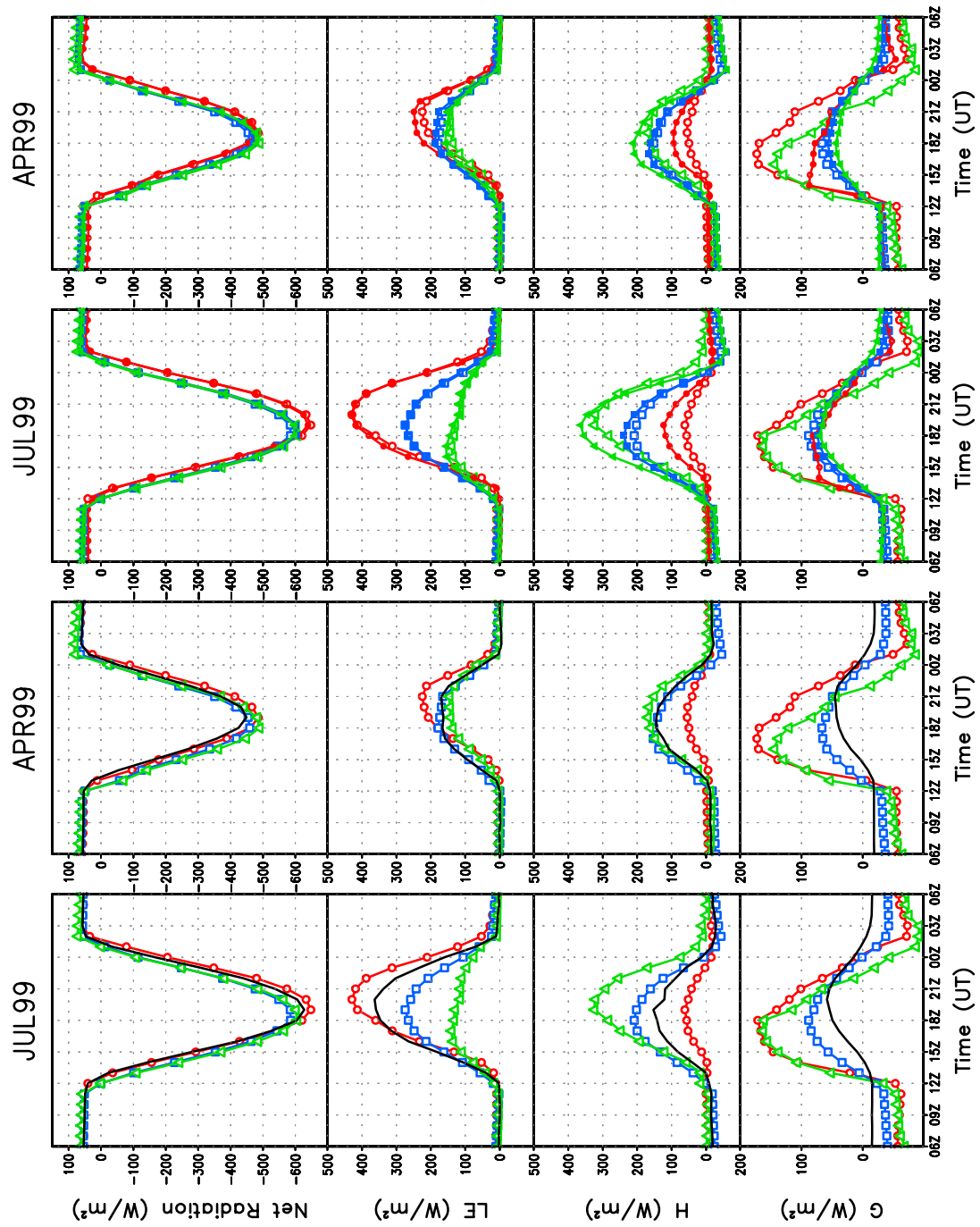


Figure 14.

H/LE , not shown), with VIC showing the opposite, and Noah showing bias similar to VIC in sign but notably smaller in magnitude. In April, Mosaic retains high LE bias and low H bias, while VIC and Noah show little bias in LE or H .

[71] Figures 13 and 14 exhibit serious errors in ground heat flux G in VIC (diurnal cycle), and especially in Mosaic (both diurnal and annual cycles). Noah shows comparatively little error in G , with virtually no bias in monthly mean (Figure 13) and rather modest high bias in diurnal daytime G from phase error. Mosaic has a large bias in monthly mean G during most months (Figure 13) and most hours (Figure 14), though the sign changes depending on the time in the diurnal/annual cycle. VIC has large hourly biases in G for most hours (Figure 14), but small monthly mean biases (Figure 13). VIC's daytime and nighttime biases in G are rather symmetric and opposite in sign, so they nearly cancel on a daily or monthly mean basis. Both Mosaic and VIC show large daytime diurnal high bias and phase error in G in both July and April of Figure 14, with daytime maximums 2–3 times larger than observed. Finally, the simultaneous and very high daytime biases in G and LE in Mosaic during April–July conspire to yield dramatically low sensible heat flux (H) during these months. Mosaic monthly mean H values during May–June are less than the annual winter minimums in observed H .

[72] With NLDAS infrastructure now in place (section 2), increasing attention is being given to model experiments. The large biases in G in Mosaic and VIC prompted such experiments, described in RL-N and presented here briefly. The VIC and Noah models use a surface energy balance approach for their surface radiative temperature, calculated in VIC for a thin but nonvanishing skin layer that has nontrivial heat storage, while Noah assumes an infinitesimally thin skin layer with negligible (zero) heat storage. As a test, VIC was re-executed by imposing zero heat storage in its surface energy balance treatment. The formulation G in Mosaic [Koster and Suarez, 1996] is based on the standard force-restore or “slab” treatment. It assumes that the “aggregate” surface/canopy medium of vegetation and near-surface soil has nonnegligible heat-storage capacity, specified by a heat capacity parameter (C_H) that strongly impacts G . In NLDAS, the C_H value in Mosaic's control run ($175,000 \text{ J m}^{-2} \text{ K}^{-1}$) was one calibrated in an earlier, independent temperature data assimilation system [Radakovich *et al.*, 2001], and not the lower traditional C_H value ($70,000 \text{ J m}^{-2} \text{ K}^{-1}$) specified by Koster and Suarez [1996] and used in several Mosaic PILPS experiments. To gage the impact of C_H , a Mosaic test was executed in NLDAS using the traditional lower value. Columns 3–4 of Figure 14 give the results of these Mosaic and VIC tests. (Aside: the Noah test in Figure 14 is presented later in section 4.2 and does not involve a change to surface heat capacity.) In both Mosaic and VIC, model ground heat flux was dramatically improved in the tests, both in July and April, becoming competitive with that in Noah (though

Mosaic manifests an unusual anomaly in G during the early morning). This improvement in simulated G holds throughout the year in both models, as shown in RL-N.

[73] However, the improvement in G in the Mosaic and VIC tests provided no improvement in the large LE biases in Mosaic or VIC in Figure 14. Rather, the increase in daytime available energy ($R-G$) gained by reducing daytime high bias in G acted only to increase the sensible heat flux H , in both models, which helped the low H bias in Mosaic and worsened the high H bias in VIC. The nonresponsiveness of LE and the high response in H strongly suggests that the canopy resistance is substantially higher than the aerodynamic resistance in both models in this vegetated region in the warm season. Sensitivity tests of canopy resistance will be a focus in all three SVAT models in NLDAS follow-on studies. As a start, Mosaic tests (not shown) have been run in which the fixed thicknesses of Mosaic's soil layers (with fixed 40-cm root zone) in Mosaic's control run was replaced with the traditional Mosaic approach of letting soil layer thickness and root depth vary tile by tile according to the tile's vegetation type. The impact on the warm season latent heat flux of Mosaic was significant, but the improvement with respect to ARM flux observations was mixed, being either negative or positive, depending on which warm season month was examined. Thus further evaluation is underway.

4.2. In Situ Validation of Land Surface Skin Temperature Over the SGP

[74] A chief goal of NLDAS is assimilation of satellite data to improve soil moisture, and in turn, surface fluxes. One keen interest is the assimilation of satellite-derived LST. Positive impact from LST assimilation will be greatly enhanced if errors in modeled LST stem primarily from errors in the background model's Bowen ratio that arise from errors in model soil moisture states. Prospects for success are much lower if LST errors arise from Bowen ratio errors caused not by soil moisture, but by errors in the model's treatment of (1) vegetation cover and nonsoil moisture attributes of its canopy conductance, (2) the surface air layer and its aerodynamic conductance, (3) albedo and net solar insolation, or (4) ground heat flux, G , and its impact on the available energy ($R - G$). Section 4.1 uncovered substantial errors in G . This section uncovers significant impact on model LST from intermodel differences in aerodynamic conductance.

[75] At the top of the hour in NLDAS, all three LSMs output an instantaneous, grid-cell mean, radiometric surface temperature, referred to here as the land surface skin temperature, LST. To obtain LST, each model applies the Stephan-Boltzmann Law given by $L = \epsilon\sigma(LST)^4$, in which ϵ is the surface emissivity (=1 in all three LSMs), σ is the Stephan-Boltzmann constant, LST is the skin temperature (K), and L is the upwelling longwave radiation (W m^{-2}). In Noah, which is nontiled, the surface energy budget is solved once for each grid cell to obtain LST and then L . In Mosaic

Figure 14. Monthly mean diurnal cycle of surface energy fluxes (W m^{-2}) of net radiation R (row 1), latent heat LE (row 2), sensible heat H (row 3) and ground heat G (row 4) for July (columns 1 and 3) and April (columns 2 and 4) of 1999. Columns 1 and 2: observed (bold black line, no symbols) and control runs for Noah (open squares, blue), VIC (open triangles, green), and Mosaic (open circles, red). Columns 3 and 4: same control runs as columns 1 and 2, plus one test each for Noah (solid squares, blue), VIC (solid triangles, green), and Mosaic (solid circles, red). See text for test features. The y axis range varies between rows. Positive flux is surface heat sink.

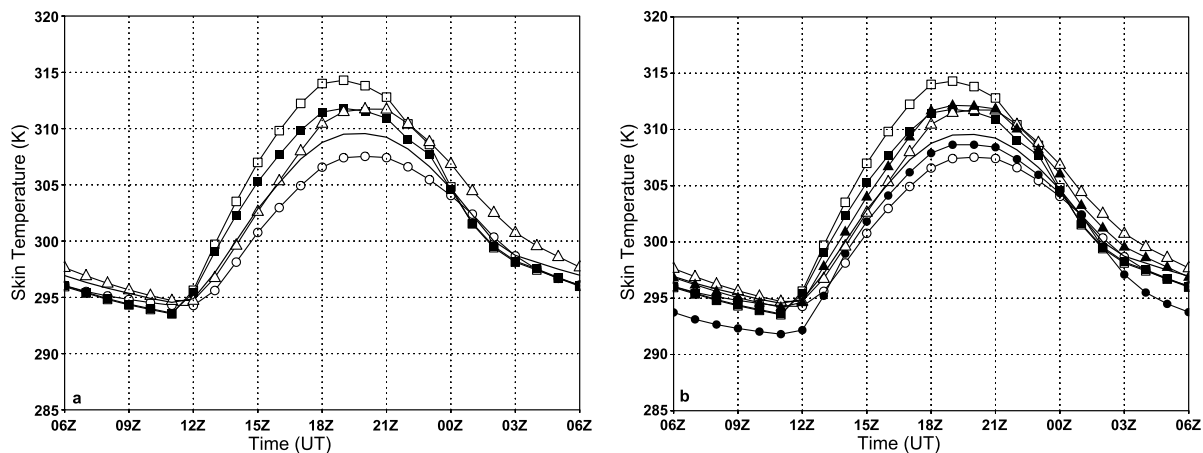


Figure 15. Monthly mean diurnal cycle of LST (K) averaged over all ARM/CART SIRS sites for July 1999 from observations (solid line, no symbols), control runs (open symbols) and test runs (solid symbols) for Noah (squares), VIC (triangles), and Mosaic (circles). (a) All three control runs plus Noah test of aerodynamic conductance. (b) As in Figure 15a, plus Mosaic and VIC tests of lower surface heat capacity.

and VIC, LST and L are obtained for each tile from a tile-specific energy budget, then the tile-weighted mean L over the grid cell is obtained, from which the grid-cell mean LST is derived from $L = \epsilon \sigma (LST)^4$. Finally, at each SIRS station, the observed L is time averaged to the top of the hour, and LST is obtained also from the latter relation using $\epsilon = 1$.

[76] Figure 15a shows the multistation average, monthly mean diurnal cycle of SIRS-observed and LSM co-located LST of the control runs, averaged over the SIRS sites, for July 1999. (The Noah test in Figure 15a is described later.) Mosaic has a midday cool bias in both months, as expected, given its high bias in LE and G and low bias in H in Figure 14 (columns 1 and 2). VIC and Noah have midday warm biases in July, also as expected, given their low LE and high H bias in July. While the sign of the models' midday LST bias in Figure 15a is as expected, the comparative magnitude of the bias between the models is perplexing at first, given the LE and H fluxes in Figure 14. Specifically, the VIC midday (19–20 UTC) warm bias in July (about +2 K) is about half as large as Noah (about +4 K), despite VIC's Bowen ratio ($BR = 2.91$) at this time being much higher than Noah's ($BR = 0.70$) and the observed ($BR = 0.38$). Thus VIC does not yield the largest midday warm bias, despite having by far the largest high bias in Bowen ratio.

[77] The daytime high bias in G in VIC's control run does not answer the paradox. Figure 15b shows the LST of the Mosaic and VIC tests of lower heat capacity (and the later Noah test) alongside control runs. Figure 15b reveals that the lower heat capacity (and its lower midday ground heat flux) in the VIC test does raise VIC's midday LST as physically expected, but only slightly, leaving it still well below the warmer LST of Noah. Figure 15b shows that the lower heat capacity test in Mosaic does yield a nontrivial increase in midday LST, reducing Mosaic's July midday cool bias by roughly half (though introducing a nighttime cool bias, because of the reduced nighttime release of stored ground heat diminished by the reduced heat capacity).

[78] The chief explanation of the paradox of VIC versus Noah midday summer LST lies in significant intermodel

differences in aerodynamic conductance. In the three models, the sensible heat flux H (W m^{-2}) is computed from the typical bulk transfer formulation given by

$$H = -\rho c_p C_h |V| (T_a - LST) \quad (2)$$

where ρ is the air density (kg m^{-3}), c_p the heat capacity for air ($1004.5 \text{ J kg}^{-1} \text{ K}^{-1}$), $|V|$ the wind speed (m s^{-1}), T_a the air temperature (K), and C_h the surface turbulent exchange coefficient for heat. The product $C_h |V|$ is the aerodynamic conductance (m s^{-1}), and its reciprocal is the aerodynamic resistance. C_h manifests a strong diurnal cycle with larger values during daytime heating. In (2), positive H means a heat source to the atmosphere and heat sink to the land surface for daytime LST exceeding T_a . The models get common surface forcing values of ρ , $|V|$, and T_a . Only C_h and LST in (2) are computed uniquely in each model. Therefore Noah can have higher midday values of LST than VIC simultaneously with lower midday values of H than VIC if and only if Noah has lower values of C_h . Figure 16 depicts the July 1998 monthly mean diurnal cycle of C_h for each model, averaged across the 14 EBBR stations. The line with solid squares in Figure 16 is a Noah test, discussed later, using a modified roughness length for heat. Indeed, the Noah control run has substantially smaller daytime values of aerodynamic conductance, and hence C_h , than Mosaic, and far smaller values than VIC. Follow-on research will seek to derive C_h explicitly from the EBBR observations.

[79] The smaller C_h values for Noah inferred from Figure 16 motivated a sensitivity run. The treatment of C_h in Noah was the subject of the NCEP study by *F. Chen et al.* [1997], which examined the impact on C_h of the chosen formulation for the roughness length for heat, z_{0h} . From a suite of tests, Chen et al. recommended the z_{0h} formulation of *Zilitinkevich* [1995], which is based on the dynamic roughness Reynolds number and includes an adjustable parameter, denoted here C_z , in the range 0–1. The Noah control run in NLDAS uses $C_z = 0.2$. Decreasing C_z increases z_{0h} , which increases C_h (thus increasing aerodynamic conductance) and the land/atmosphere coupling, thereby decreasing daytime

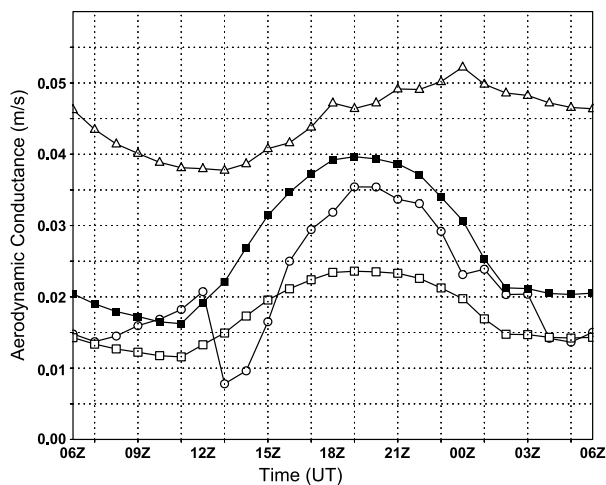


Figure 16. Monthly mean diurnal cycle of aerodynamic conductance (ms^{-1}) averaged over all ARM/CART SIRS sites for July 1998 for the control runs of Noah (open squares), VIC (open triangles), and Mosaic (open circles) and the Noah test (solid squares) using a modified formulation for roughness length for heat.

LST. This was investigated by a Noah sensitivity test, depicted in Figure 16, which used $C_z = 0.05$ in place of $C_z = 0.2$.

[80] The July monthly mean, midday aerodynamic conductance values in the Noah test in Figure 16 exceed the control values by nearly 70%. The surface fluxes and LST of this Noah test are depicted in Figure 14 (right columns) and Figure 15, respectively. Figure 15 shows a pleasing 2–3 K decrease in Noah's July peak daytime LST, cutting the midday warm bias by about half. Inspection of Noah test versus control surface fluxes in Figure 14 reveals small midday changes in R (increase) and G (decrease) of about $10\text{--}20 \text{ W m}^{-2}$ each, as expected from the lower midday LST. This slightly improved the already small biases in Noah R and G and yielded a roughly $20\text{--}40 \text{ W m}^{-2}$ increase in midday “available energy” ($R - G$) for sensible and latent heat fluxes. Once again however, very similar to the earlier response in Mosaic and VIC to increases in ($R - G$), the LE change in the Noah test in both July and April was negligible, especially so in July, with the $R - G$ increase in July once again being realized almost entirely by an increase in H .

[81] The lack of change in LE in July strongly indicates that the evaporative resistances to canopy transpiration (canopy resistance) and bare soil evaporation are much larger than, and thus dominant over, the aerodynamic resistance in influencing LE in this situation. Again, Figure 1a shows that the vegetation cover over the bulk of the ARM-SGP region in July is of order 40–80%. The analysis of Vogel *et al.* [1995] shows that even over an irrigated midlatitude wheat field in June, the LE change from a 20% change in aerodynamic resistance yielded only a 2% change in LE; that is, canopy resistances over non-sparse vegetation in midlatitude summer are typically much larger than, and dominant over, aerodynamic resistances, even when the soil is wet and contributing to a rise in canopy conductance. These results strongly suggest that the

remaining July bias of +2 K in the Noah test in Figure 15 is caused by an overly high canopy resistance. Moreover, the aerodynamic conductance of VIC in Figure 16 may well be too high, thus acting to preclude a much higher midday warm bias in LST that would more properly reflect VIC's high bias in Bowen ratio (H/LE) in July in Figure 14. Sensitivity tests of aerodynamic resistance and canopy resistance will be one focus in follow-on NLDAS studies.

4.3. Satellite-Based Validation of Land Surface Skin Temperature

[82] The GOES LST fields are produced by the GCIP partnership of NESDIS and UMD in GOES land surface products. In this section, after validating GOES LST against ARM LST as a benchmark, we use GOES LST to assess NLDAS LST over north central CONUS. The study is limited to nonmountainous regions, to avoid shadowing effects on the GOES LST retrievals. The retrievals are obtained from GOES-East (GOES-8) and provide fields of hourly LST at 0.5° spatial resolution in cloud-free conditions during daytime. The LST retrieval provides a single aggregate LST for each 0.5° target scene. We bilinearly interpolate the LST fields to the $1/8^\circ$ NLDAS grid. The GOES LST is retrieved only at 0.5° targets deemed 100% cloud-free. Cloud detection is based on that of earlier GOES insolation-retrieval studies such as Tarpley [1979], as refined in later studies such as PT-N. Despite the 100% cloud-free criteria, clouds may still be present in the scene owing to (1) optically thin cirrus, (2) subresolution or “subpixel” cloud (fair weather cumulus), and (3) difficulty of cloud detection over snow cover.

[83] GOES LST is retrieved by the so-called “split-window” technique of Wu *et al.* [1999], in which LST is obtained from a linear regression of the GOES brightness temperatures in the $11 \mu\text{m}$ and $12 \mu\text{m}$ bands. The regression coefficients were derived assuming a surface emissivity of $\epsilon = 1$. This assumption is valid over land surfaces of nonsparse vegetation or snowpack, but less valid over rather bare soils (wherein $\epsilon = 0.91\text{--}0.97$). Uncertainty from emissivity issues is avoided in this study by staying over nonsparse vegetation and by our universal application of $\epsilon = 1$ in (1) the NLDAS models, (2) the in situ ARM/SIRS sites, and (3) the GOES retrievals.

[84] We assess GOES LST here against the in situ LST observations of the 22 SIRS sites. We limit the assessment to nonwinter, as our future assimilation of GOES LST will generally be confined to the warm season of stronger coupling between LST and soil moisture. Figure 17 presents the monthly and multistation mean of the daytime hourly diurnal cycle of GOES LST and ARM LST for April and July in 1998 and 1999. The data samples for Figure 17 (and Figure 18) represent only locations and times when the GOES cloud screening detected zero cloud. In Figure 17, the GOES LST demonstrates a remarkable ability to match the station-observed mean diurnal cycle, though it shows a small cool bias (likely from undetected clouds) of order 0–1.5 K before 18 UTC (local noon) and 1–2.5 K thereafter. The smaller cool bias in the morning is likely from less prevalent cloud cover than (e.g., subpixel cumulus). In future data assimilation, one may mitigate this cool bias by assimilating the 3-hour rise in GOES LST before noon (about 15–18 UTC here), rather than LST itself. Tarpley [1994] applied the morning rise of GOES LST to infer

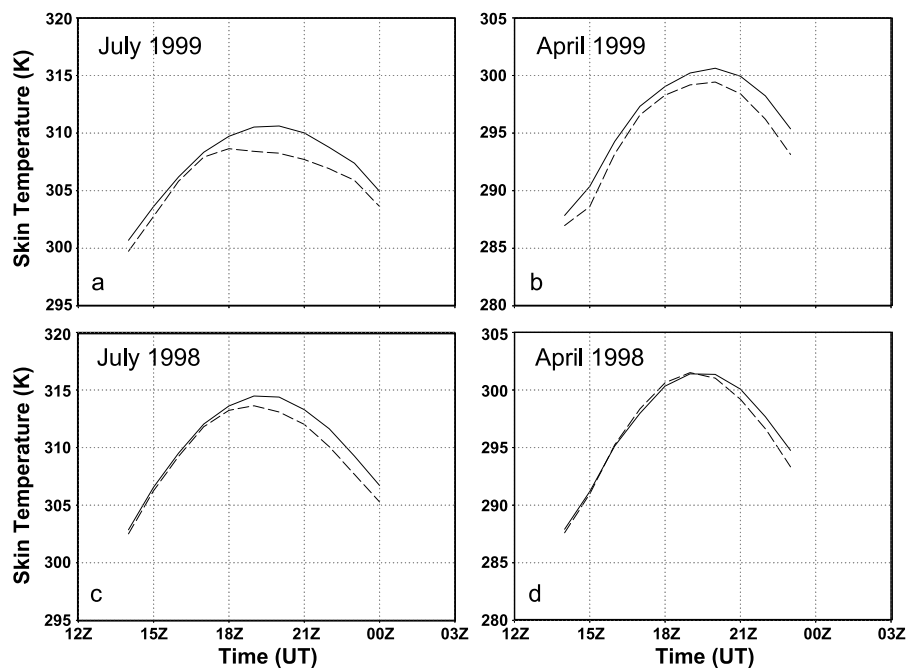


Figure 17. Monthly mean diurnal cycle of LST (K) averaged over all ARM/CART SIRS sites for (left) July and (right) April during (a and b) 1999 and (c and d) 1998 from SIRS observations (solid lines) and GOES-East retrieval (dashed lines).

monthly mean surface-moisture availability. Like the ARM LST, the 18 UTC GOES LST in Figure 17 is warmer (3–4 K) and its preceding 3-hour morning rise is larger (by about 1K) in July 1998 than July 1999. This interannual LST variability reflects the drought episode over the ARM region in July 1998, thus conditions were warmer and drier than in 1999 (and likely less cloudy, hence the smaller GOES LST afternoon cool bias in July 1998 versus 1999).

[85] Our goal is to use GOES LST retrievals to assess NLDAS LST over large regions that lack in situ observations of LST or surface fluxes. As a benchmark for that goal, Figure 18 illustrates, at the SIRS sites during July and April 1999 for 18 UTC, a pleasing similarity between GOES-based and ARM-based site-by-site match-ups with model LST. Moreover, all three models show good skill in either the GOES or ARM validation setting by yielding rather tight clusters close to the diagonal (and hence high correlations, shown later). In each month, the separate GOES and ARM match-ups use the same sample of instances where the GOES deemed the site to be cloud free. In Figure 18, the sample size of 198 in April (out of a possible $660 = 30 \text{ days} \times 22 \text{ stations}$) is notably smaller than that of 334 (out of 682) in July, as the GOES cloud screening detects cloud more often in the spring. One would expect this from the natural trend of decreasing cloud cover from spring to summer and the greater likelihood in July of shallow, subresolvable cumulus. Indeed in July, the GOES LST in Figure 18 manifests a small (but nonnegligible) leftward-pointing “cold tail” of outlier values that are not present in either the ARM observations or the models and thus likely represent GOES cloud detection failures. Similarly, the ARM observations

in April 1999 show several warm outlier values (near 315 K), not present in either the GOES or model LST, likely representing bad ARM station observations.

[86] Most importantly, as desired, the GOES versus model match-ups yield the same sense of model midday LST bias as we derived from ARM data alone in the prior section. Table 5 compares the GOES-based versus ARM-based model bias, error standard deviation and correlation obtained from the Figure 18 match-up and listed top-down from warmest to coldest model bias. The table shows good agreement between the sign and magnitude of the GOES-based and ARM-based model bias. The GOES-based model bias is order 1 K warmer than the ARM-based model bias, owing to the aforementioned GOES LST cool bias of order 1 K versus ARM LST. The GOES-based model LST bias essentially reproduced the ARM-based model bias, both in (1) the absolute sense of correct sign and reasonably good magnitude and (2) the relative sense between models and between spring and summer season. Specifically, in agreement with the ARM-based LSM signatures of control-run midday LST bias presented for the entire annual cycle in RL-N, the GOES-based model bias results in Figure 18 show that in summer (1) Noah has the largest warm bias, which becomes much smaller in spring, (2) VIC has a smaller and modest warm bias, which becomes virtually zero in spring, and (3) Mosaic has a modest cool bias, which becomes larger in spring. Similarly, the GOES-based and ARM-based standard deviations in Table 5 from Figure 18 are in reasonable agreement, both in the range of 3–4 K. Finally, the GOES-based correlations with model LST in Table 5 are very encouraging, ranging between 0.66 and 0.78, with

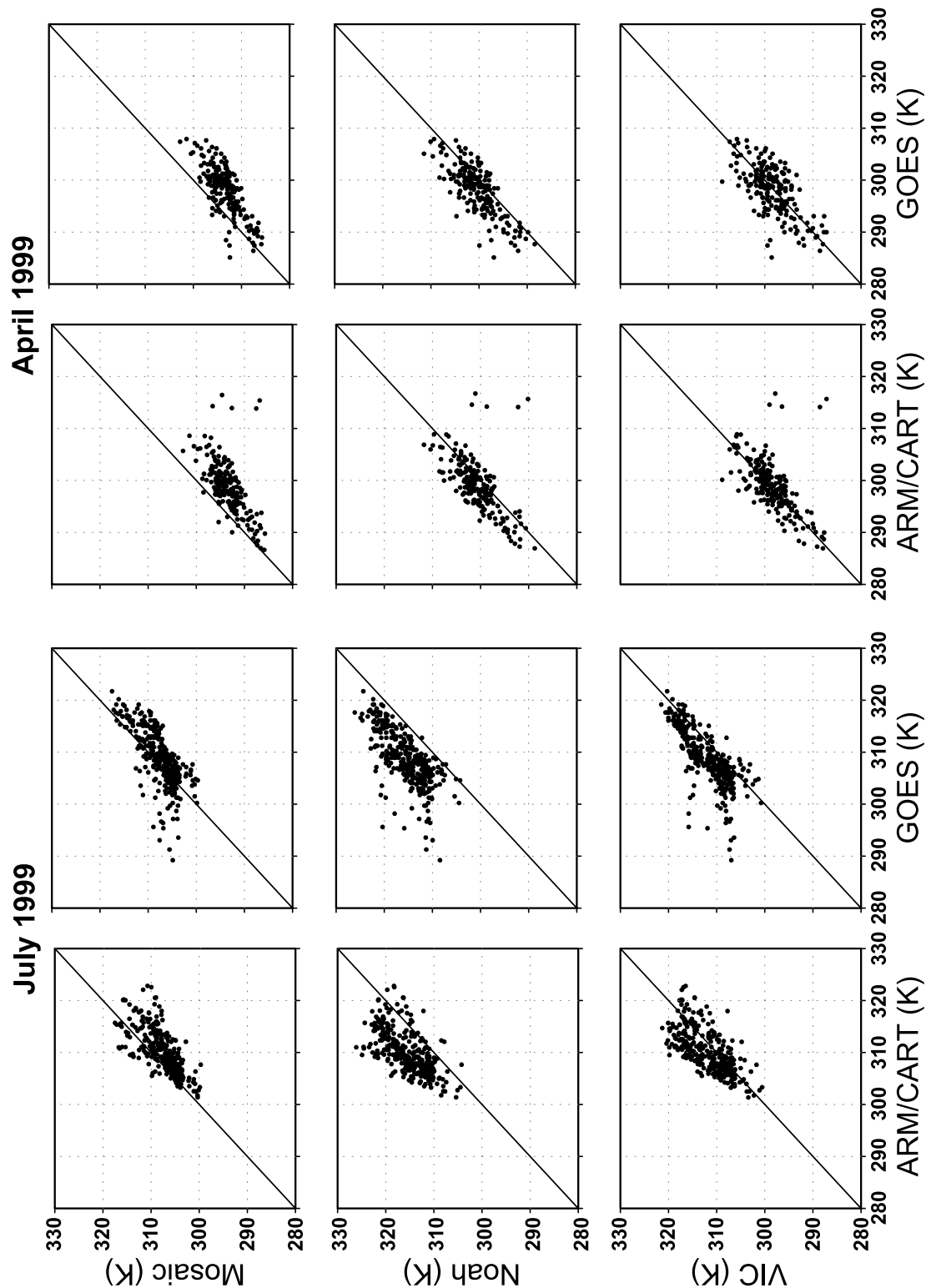


Figure 18.

Table 5. Bias, Error Standard Deviation, and Correlation of Model LST Versus Both ARM LST and Unscreened GOES LST Over the SGP From Figure 18^a

Model	July 1999 (GOES/ARM)			April 1999 (GOES/ARM)		
	Bias, K	STDE, K	Correlation	Bias, K	STDE, K	Correlation
Noah	+6.5/+5.4	3.8/3.5	0.70/0.64	+2.4/+1.7	2.9/4.1	0.78/0.61
VIC	+2.8/+1.7	3.4/3.3	0.76/0.68	+0.3/−0.5	3.9/4.3	0.66/0.57
Mosaic	−1.2/−2.3	3.7/3.0	0.70/0.72	−4.7/−5.4	3.0/4.0	0.77/0.62

^aSTDE, error standard deviation.

five of six values of 0.70 or more, all without any screening of the GOES cold LST outliers. The correlations of model LST with GOES LST are consistently higher than the ARM-based correlations, likely from better match of the spatial scale of the GOES footprint and the NLDAS grid-cell size, versus the point scale of ARM values.

[87] Encouraged by Figures 17 and 18, we evaluated model LST against GOES LST in Figure 19 across a large region of the northern Midwest, bounded by latitudes 39°N and 53°N and longitudes 82°W and 98°W (straddling 90°W longitude, representing strict local noon at 18 UTC). This region is chosen for its (1) spatial separation from SGP, (2) nonsparse green vegetation in summer (Figure 1a), and (3) vivid model differences in evaporation in Figure 2. GOES versus model LST over this region at 18 UTC for July and April 1999 are presented in Figure 19. The “screened” results therein are described later. Sample counts in Figure 19 are 70,000–100,000 (2–3 orders larger than Figure 18, owing to the larger region). In Figure 19, we binned the data into 1 K intervals (for display only, kept full precision in statistics) and used colors to denote 4 orders of data counts: 0–10 (red), 10–99 (yellow), 100–999 (light green), and 1000–9999 (dark green). The green shades depict the vast majority of the sample and they manifest well-behaved, elongated clusters lying near and parallel to the diagonal, as desired.

[88] Yet the red and yellow “tails” of cold GOES LST in the unscreened panels of Figure 19 depict a nontrivial number of points that likely represent cloud detection failures, similar to the cold tails in the July GOES panels of Figure 18. Hence we used model simulated LST to screen the cold GOES LST tails in columns 2 and 4 in Figure 19. Specifically, we rejected a GOES LST if model-minus-GOES LST exceeded the unscreened, sample-wide model-minus-GOES LST bias by more than two times the model-minus-GOES LST standard deviation of the unscreened sample. Figure 19 shows that this screening preserves the high-density core region of the original data cluster, while eliminating the cold tails. The data counts of the three panels in any column of Figure 19 are identical in the unscreened case, but differ slightly (less than 1.2%) in the screened case, as the screening for each panel uses the given model’s LST simulation. Last, in Figure 19, we point out the “lower lobe” of cold model LST in the April results of Mosaic and VIC. The lower edge of the lobe ends at freezing, suggesting that Mosaic and VIC have

sustained remnants of melting snowpack too late into the spring in this region.

[89] Table 6 shows the GOES-based model LST bias, error standard deviation and correlation obtained from Figure 19. The screened model-bias results are warmer, as expected, but by a modest 0.5 to 1.0 K. More notably, the standard deviation of the model errors are significantly less by around 1.0–1.5 K and the already high correlations increase by around 0.1 to 0.76 or higher for almost all entries. It is revealing to compare the unscreened results from the north central CONUS in Table 6 with the SGP results in Table 5 (all unscreened). Table 6 preserves the relative nature and order noted in the biases in Table 5; namely, Noah is warmest and Mosaic is coolest in July, with Noah notably less warm and Mosaic notably more cool in April, while VIC falls in-between in both months. Yet the unscreened biases in Table 6 are consistently 1–3 K warmer than those in Table 5, reflecting that either the models are warmer in this region relative to GOES LST than over the ARM SGP region, or the GOES LST has a larger cool bias (of order 1–3 K) than the GOES cool bias over the ARM region. We suspect the latter owing to more cloud contamination in this more humid region.

[90] We look forward to use of GOES LST in future assimilation studies and LSM assessments. The validation and utility of GOES LST in mountainous regions awaits future study. Efforts continue at UMD and NESDIS to improve cloud detection and spatial resolution in GOES LST.

5. Summary and Conclusions

[91] A multi-institution partnership under the GCIP program has developed and evaluated the backbone for a North American Land Data Assimilation System (NLDAS). This paper is the overview of nine NLDAS papers (Table 2), which appear together in the online HTML version of the GCIP3 special section of *Journal of Geophysical Research*. These partners assembled a wide set of GCIP-sponsored products and other data sources into robust forcing data sets and multiscale validation databases. Validation applied surface stations measuring energy fluxes, surface meteorology, soil moisture and temperature, and mountain snowpack, plus daily streamflow observations and satellite-derived land surface temperature (LST) and snow cover.

[92] NLDAS features nonmodel sources of precipitation and solar insolation and the four LSMs of Noah, VIC,

Figure 18. Comparison of model (*y* axis) versus observed (*x* axis) LST (K) at 18 UTC over all ARM/CART SIRS sites for July (columns 1 and 2) and April (columns 3 and 4) 1999 for (top) Mosaic, (middle) Noah, and (bottom) VIC versus SIRS observations (columns 1 and 3) and GOES-East observations (columns 2 and 4). Match-up point included only if GOES LST is available (cloud free), yielding sample sizes of 334 in columns 1 and 2 and 198 in columns 3 and 4.

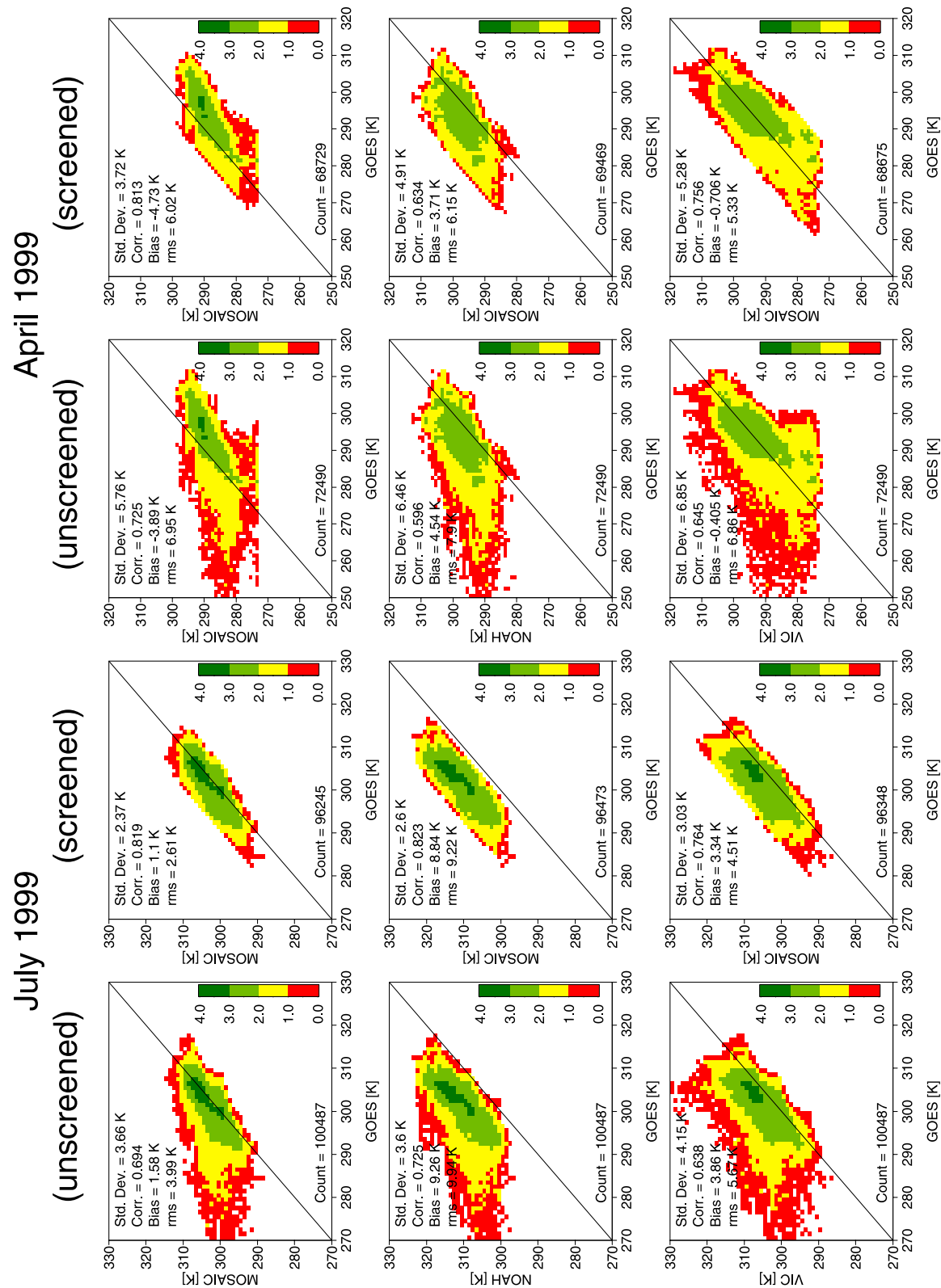


Figure 19.

Table 6. Bias, Error Standard Deviation, and Correlation of Model LST Versus Screened and Unscreened GOES LST for North Central CONUS From Figure 19^a

Model	July 1999 (GOES: Unscreened/Screened)			April 1999 (GOES: Unscreened/Screened)		
	Bias, K	STDE, K	Correlation	Bias, K	STDE, K	Correlation
Noah	+9.3/+8.8	3.6/2.6	0.73/0.82	+4.5/+3.7	6.5/4.9	0.60/0.63
VIC	+3.9/+3.3	4.2/3.0	0.64/0.76	−0.4/−0.7	6.9/5.3	0.65/0.76
Mosaic	+1.6/+1.1	3.7/2.4	0.69/0.82	−3.9/−4.7	5.8/3.7	0.73/0.81

^aSTDE, error standard deviation.

Mosaic, and Sacramento (SAC) executing in parallel on a 1/8° CONUS domain to provide land-state background fields for data assimilation experiments. The infrastructure of NLDAS includes streamflow routing and provides both real-time and retrospective execution to support both operations and research. The hourly NLDAS forcing, now spanning seven years from October 1996 to present (at time of writing), represents an important NLDAS by-product.

[93] The NLDAS thrust here was the forerunner to recent companion initiatives in 50 + year retrospective executions of VIC on the NLDAS grid by *Maurer et al.* [2002] and Noah on the NLDAS grid by H. van den Dool of NCEP (private communication, 2003). We encourage researchers to compare our NLDAS water and energy budgets here with (1) these 50 + year retrospectives, (2) operational global and regional coupled 4DDA, and (3) global and regional reanalysis.

[94] A central distinction between the above suites is the source and bias in the surface forcing. In NLDAS here, the forcing is anchored by gage-based daily precipitation analyses (with hourly disaggregation using radar-derived precipitation) and hourly surface insolation derived from GOES satellites. All remaining forcing is from NCEP's mesoscale 4DDA system, known as EDAS. NLDAS surface forcing compared well against Mesonet observations over the SGP. In tests that replaced NLDAS forcing with local-station forcing, the test versus control differences in states and fluxes were pleasingly small. Yet we continue thrusts to further improve the forcing. The GOES-based solar insolation shows some high bias at low sun angles and over snowpack, though less bias than the fallback insolation from EDAS. At mountain SNOTEL sites in western CONUS, NLDAS precipitation has a nearly 50% low bias. Thus NLDAS partners have implemented a PRISM-based technique [*Daly et al.*, 1994] into the real-time forcing as of 1 February 2002 (with plans to reproduce the retrospective forcing using PRISM).

[95] Observed precipitation and streamflow applied to the annual water budget provided observation-based estimates of evaporation and runoff over large regions. This revealed substantial biases and intermodel differences in evaporation. The ARM-observed surface energy budget of the SGP confirmed the same evaporation bias anticipated from the annual water budget analysis. The three SVAT-type models, though they treat vegetation cover explicitly,

nevertheless yield strikingly different warm season evaporation over vegetation. This canopy conductance disparity among the models is a foremost issue. Moreover, evapotranspiration bias can run counter to intuition. Though Mosaic has the shallowest root zone in NLDAS of the three SVATs, it has the highest warm season evaporation rates and hence highest warm season storage change in soil moisture, as it allows vigorous upward diffusion of water from the subroot zone.

[96] Aerodynamic conductance ($C_h|V$) was a second area of large disparity. Overly large or small midday C_h values were found to substantially distort the expected correlation between daytime LST bias and Bowen ratio bias. Such distortion has crucial implications for the prospects of successful assimilation of satellite LST. In summer of the SGP, though Noah had a small high bias in Bowen ratio while VIC had a large high bias, Noah had the largest midday LST warm bias and VIC the smallest. The cause was the substantially lower C_h values in Noah versus VIC.

[97] Soil moisture storage emerged as a third area of large disparity, similar to previous PILPS studies. The forward radiative transfer models that are crucial to modern-era assimilation of satellite data are sensitive to absolute moisture states. Thus, while one can simulate evaporation and runoff well from good simulation of temporal change in soil moisture, land assimilation of satellite data brings a more stringent need for good absolute states of soil moisture.

[98] NLDAS simulations of snowpack water equivalent (SWE) at mountain SNOTEL sites showed a substantial low bias in all four LSMs, with an attendant low bias in runoff, owing to the cited high insolation bias over snow and low precipitation bias in mountains in NLDAS. Yet there was still notable disparity across the models in snow cover fraction, snow albedo and timing of spring snowmelt. Noah exhibited particularly low snow albedo, which conspired with the high insolation bias in the forcing to yield a very early bias in Noah seasonal snowmelt. The VIC and SAC models yielded the smallest biases in simulated SWE and regional snow cover, with VIC yielding the best snowmelt timing. The elevation tiling unique to VIC in NLDAS provides an advantage in snow state modeling. SAC snowpack simulations also performed well, as SAC's simple temperature index-based snow model bypasses surface energy balance and snow albedo, and thus avoids the

Figure 19. Comparison of model (y axis) versus GOES-East (x axis) LST (K) at 18 UTC over the northern Midwest during July (columns 1 and 2) and April (columns 3 and 4) 1999 for (top) Mosaic, (middle) Noah, and (bottom) VIC versus unscreened (columns 1 and 3) and screened (columns 2 and 4) GOES LST. See text for color scale definition and latitude/longitude range of region.

positive feedback loops that can plague snowpack simulations in surface energy balance models [Slater *et al.*, 2001; Bowling *et al.*, 2003].

[99] We emphasize that the Mosaic, VIC, and SAC LSM configurations in NLDAS differed in important aspects from their traditional configurations. For example, Mosaic executed with fixed soil-layer thicknesses and root depth, rather than vegetation-dependent spatial variability. VIC executed with one-hour rather than its typical three-hour time steps and with hourly rather than uniform daily disaggregation of precipitation. Moreover, the SAC runs, by design, were the first executions over a continental scale of the newly developed, semi-distributed version of SAC, with a priori noncalibrated parameters. Thus NLDAS execution of SAC provided an essential benchmark for future CONUS-wide SAC execution in semi-distributed mode.

[100] Indeed, all the model intercomparisons thus far in NLDAS must be viewed as incomplete and providing only a benchmark. The models were not calibrated to NLDAS configuration (e.g., spatial resolution, model time step, given fields of surface characteristics, temporal character of precipitation disaggregation). As in PILPS, our purpose is not to rank the models, but rather to build an enduring test bed via the NLDAS infrastructure for development of objective calibration approaches on very large continental scales that far exceed and complement the regional scales applied recently in PILPS [Wood *et al.*, 1998; Bowling *et al.*, 2003; Boone *et al.*, 2004].

[101] Last, this paper concluded with assessments and application of GOES-based LST at $1/2^\circ$ hourly resolution. The diurnal cycle of GOES LST validated well against SGP flux-stations. Validation of model LST by means of GOES LST over the SGP gave validation scores similar to those from ground-based ARM observations. By using GOES LST to validate model LST over the northern Midwest, we found model LST biases consistent with those over the SGP. Hence GOES LST offers a powerful large-scale LSM validation tool.

[102] We are now assembling the tools to perform actual land data assimilation experiments. For this purpose, we will be adding one or two forward radiative transfer models into our common NLDAS infrastructure. Additionally, we are pursuing development of adjoint models and ensemble Kalman filter approaches. Finally, the NLDAS initiative here represented a pathfinder for a companion extension to a Global Land Data Assimilation System (GLDAS) by the NASA and NCEP partners of NLDAS. The GLDAS is described by Rodell *et al.* [2003] and is presently being ported from NASA to NCEP. With NLDAS and GLDAS together, NCEP and NASA and partners are striving to provide land state initial conditions for (1) land-memory predictability studies and (2) operational weather and climate model forecasts on daily to seasonal timescales.

Notation

4DDA	four-dimensional data assimilation
ARM/CART	Atmospheric Radiation Measurement/Cloud and Radiation Testbed (DOE)
ARS FAO	Agricultural Research Service Food and Agriculture Organization

AVHRR	Advanced Very High Resolution Radiometer on NOAA polar satellites
CAPE	convective available potential energy
CONUS	continental United States
CPC	Climate Prediction Center
CL-N	NLDAS paper by Cosgrove <i>et al.</i> [2003a]
CM-N	NLDAS paper by Cosgrove <i>et al.</i> [2003b]
DMIP	Distributed Model Intercomparison Project
DOE	Department of Energy
EDAS	NCEP Eta-model-based 4-D Data Assimilation System
EMC	Environmental Modeling Center of NCEP
EBBR	energy balance Bowen ratio flux stations in ARM/CART network
EF	extended facility flux stations in ARM/CART network
GAPP	GEWEX America Prediction Project
GCIP	GEWEX Continental-Scale International Project
GEWEX	Global Energy and Water Cycle Experiment
GOES	Geosynchronous Operational Environmental Satellite (USA)
GRIB	Gridded Binary data file format (WMO standard)
GSFC	Goddard Space Flight Center (NASA)
GSWP	Global Soil Wetness Project
GTOPO30	Global (30 arc seconds) digital elevation database
GVF	green vegetation-cover fraction
IMS	Interactive MultiSensor Snow (NESDIS)
LAI	leaf area index
LDAS	land data assimilation system
LM-N	NLDAS paper by Lohmann <i>et al.</i> [2004]
LR-N	NLDAS paper by Luo <i>et al.</i> [2003]
LSM	land surface model
LST	land surface skin temperature
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCDC	National Climatic Data Center
NCEP	National Centers for Environmental Prediction
NDVI	normalized difference vegetation index
NESDIS	National Environmental Satellite, Data, and Information Service
NLDAS	North American LDAS
NOAA	National Oceanic and Atmospheric Administration
NRCS	National Resources Conservation Service (USDA)
NWIS	National Water Information System
NWP	numerical weather prediction
NWS	National Weather Service (NOAA)
OGP	Office of Global Programs (NOAA)
OHD	Office of Hydrologic Development (NWS, formerly Office of Hydrology)
ORA	Office of Research and Applications of NESDIS
PAR	photosynthetically active radiation
PILPS	Project for Intercomparison of Land-Surface Parameterization Schemes

- PRISM Parameter-Elevation Regressions on Independent Slopes Model
- PS-N NLDAS paper by *Pan et al.* [2003]
- PT-N NLDAS paper by *Pinker et al.* [2003]
- RFC River Forecast Center (NWS)
- RL-N NLDAS paper by *Robock et al.* [2003]
- SAC Sacramento model (Sacramento Soil Water Accounting Model)
- SCE snow cover extent
- SD-N NLDAS paper by *Schaake et al.* [2004]
- SGP Southern Great Plains (field program)
- SIRS Solar and Infrared Radiation Station
- SNOTEL Snowpack Telemetry network of the NRSC
- SNOW-17 Snow accumulation and ablation model (NWS/OHD)
- SP-N NLDAS paper by *Sheffield et al.* [2003]
- SST sea surface temperature
- STATSGO State Soil Geographic Database
- SURFRAD Surface Radiation Budget Network (NOAA/OAR-ARL, OGP)
- SVAT Surface-Vegetation-Atmosphere Transfer (model)
- SWE snowpack water equivalent
- TOGA Tropical Ocean Global Atmosphere
- UMD University of Maryland
- USGS U.S. Geological Survey
- VIC Variable Infiltration Capacity LSM
- WMO World Meteorological Organization
- WSR-88D Weather Service Radar-Doppler
- [103] **Acknowledgments.** The work by NCEP/EMC, NWS/OHD, and NESDIS/ORA was supported by the NOAA OGP grant for the NOAA Core Project for GCIP/GAPP (K. Mitchell, J. Schaake, J. Tarpley, Co-PIs). The work by NASA/GSFC was supported by NASA's Terrestrial Hydrology Program (P. Houser, PI). The work by Rutgers University was supported by NOAA OGP GAPP grant GC99-443b (A. Robock, PI), the Cook College Center for Environmental Prediction, and the New Jersey Agricultural Experiment Station. The work by Princeton was supported by NOAA OGP GAPP grant NA86GPO258 (E. Wood, PI). The work by NCEP/CPC was supported by NOAA/NASA GAPP Project 8R1DA114 (R. Higgins, PI). The work by University of Maryland was supported by grants NA56GPO233, NA86GPO202, and NA06GPO404 from NOAA OGP and by NOAA grant NA57WC0340 to University of Maryland's Cooperative Institute for Climate Studies (R. Pinker, PI). Some figures were drawn with GrADS, created by Brian Doty. We thank DOE for the ARM/CART meteorological and heat flux data that were provided to the project at no cost and the NOAA Office of Global Programs and NASA Land Surface Hydrology Program for their purchase of the Oklahoma Mesonet meteorological and soil moisture and temperature data for their funded investigators. Dag Lohmann worked on the GCIP Core Project at NCEP under the University Corporation for Atmospheric Research Visiting Scientist Program.
- ## References
- Abdulla, F. A., D. P. Lettenmaier, E. F. Wood, and J. A. Smith (1996), Application of a macroscale hydrologic model to estimate the water balance of the Arkansas-Red river basin, *J. Geophys. Res.*, **101**, 7449–7459.
- Anderson, E. A. (1973), National Weather Service River Forecast System: Snow Accumulation and Ablation Model, *NOAA Tech. Memo. NWS Hydro-17*, Natl. Weather Serv., Silver Spring, Md.
- Berbery, E. H., Y. Luo, K. E. Mitchell, and A. K. Betts (2003), Eta model estimated land surface processes and the hydrological cycle of the Mississippi basin, *J. Geophys. Res.*, **108**(D22), 8852, doi:10.1029/2002JD003192.
- Betts, A., F. Chen, K. Mitchell, and Z. Janjic (1997), Assessment of the land surface and boundary layer models in two operational versions of the NCEP Eta model using FIFE data, *Mon. Weather Rev.*, **125**, 2896–2916.
- Boone, A., et al. (2004), The Rhone-Aggregation Land Surface Scheme Intercomparison Project: An overview of results, *J. Clim.*, **17**, 187–298.
- Bowling, L. C., et al. (2003), Simulation of high latitude hydrological processes in the Torne-Kalix basin: PILPS Phase 2(e): 1. Experiment description and summary intercomparisons, *Global Planet. Change*, **38**, 1–30.
- Burnash, R. J. C., R. L. Ferral, R. A. McGuire (1973), A generalized streamflow simulation system: Conceptual models for digital computers, technical report, Joint Fed.-State River Forecast Cent., U.S. Natl. Weather Serv. and Calif. Dep. of Water Resour., Sacramento, Calif.
- Chen, F., K. Mitchell, J. Schaake, Y. Xue, H.-L. Pan, V. Koren, Q. Y. Duan, M. Ek, and A. Betts (1996), Modeling of land-surface evaporation by four schemes and comparison with FIFE observations, *J. Geophys. Res.*, **101**, 7251–7268.
- Chen, F., Z. Janjic, and K. Mitchell (1997), Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model, *Boundary Layer Meteorol.*, **85**, 391–421.
- Chen, T. H., et al. (1997), Cabauw experimental results from the Project for Intercomparison of Land-surface Parameterization Schemes (PILPS), *J. Clim.*, **10**, 1194–1215.
- Cherkauer, K. A., and D. P. Lettenmaier (1999), Hydrologic effects of frozen soils in the upper Mississippi River basin, *J. Geophys. Res.*, **104**, 19,599–19,610.
- Cosgrove, B. A., et al. (2003a), Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project, *J. Geophys. Res.*, **108**(D22), 8842, doi:10.1029/2002JD003118.
- Cosgrove, B. A., et al. (2003b), Land surface model spin-up behavior in the North American Land Data Assimilation System (NLDAS), *J. Geophys. Res.*, **108**(D22), 8845, doi:10.1029/2002JD003316.
- Daly, C., R. P. Neilson, and D. L. Phillips (1994), A statistical-topographic model for mapping climatological precipitation over mountainous terrain, *J. Appl. Meteorol.*, **33**, 140–158.
- Dimmeyer, P., A. Dolman, and N. Sato (1999), The pilot phase of the Global Soil Wetness Project, *Bull. Am. Meteorol. Soc.*, **80**, 851–878.
- Douville, H., P. Viterbo, J.-F. Mahfouf, and A. C. M. Beljaars (2000), Evaluation of the optimum interpolation and nudging techniques for soil moisture analysis using FIFE data, *Mon. Weather Rev.*, **128**, 1733–1756.
- Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley (2003), Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, *J. Geophys. Res.*, **108**(D22), 8851, doi:10.1029/2002JD003296.
- Gibson, J. K., P. Kallberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano (1997), ERA description, *ERA-15 Proj. Rep. Ser. 1*, 72 pp., Eur. Cent. for Medium-Range Weather Forecasts, Reading, UK.
- Gutman, G., and A. Ignatov (1998), The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models, *Int. J. Remote Sens.*, **19**, 1533–1543.
- Hansen, M. C., R. S. DeFries, J. R. G. Townshend, and R. Sohlberg (2000), Global land cover classification at 1 km spatial resolution using a classification tree approach, *Int. J. Remote Sens.*, **21**, 1331–1364.
- Higgins, R. W., W. Shi, E. Yarosh, and R. Joyce (2000), Improved United States precipitation quality control system and analysis, *NCEP/Clim. Predict. Cent. Atlas 7*, Clim. Predict. Cent., Camp Springs, Md.
- Hollinger, S. E., and S. A. Isard (1994), A soil moisture climatology of Illinois, *J. Clim.*, **7**, 822–833.
- Jarvis, P. G. (1976), The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field, *Philos. Trans. R. Soc. London, Ser. B*, **273**, 593–610.
- Ji, M., A. Kumar, and A. Leetmaa (1994), A multiseason climate forecast system at the National Meteorological Center, *Bull. Am. Meteorol. Soc.*, **75**, 569–577.
- Kalnay, E., et al. (1996), NCEP/NCAR 40-year Reanalysis Project, *Bull. Am. Meteorol. Soc.*, **77**, 437–472.
- Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. Hnilo, M. Fiorino, and G. Potter (2002), NCEP-DOE AMIP-II Reanalysis (R-2), *Bull. Am. Meteorol. Soc.*, **83**, 1631–1643.
- Koren, V., J. Schaake, K. Mitchell, Q. Duan, F. Chen, and J. Baker (1999), A parameterization of snowpack and frozen ground intended for NCEP weather and climate models, *J. Geophys. Res.*, **104**, 19,569–19,585.
- Koren, V. I., M. Smith, D. Wang, and Z. Zhang (2000), Use of soil property data in the derivation of conceptual rainfall-runoff model parameters, paper presented at 15th Conference on Hydrology, Am. Meteorol. Soc., Long Beach, Calif., 10–14 Jan.
- Koster, R., and P. Milly (1997), The interplay between transpiration and run-off formulations in land-surface schemes used with atmospheric models, *J. Clim.*, **10**, 1578–1591.
- Koster, R., and M. Suarez (1994), The components of a SVAT scheme and their effects on a GCM's hydrological cycle, *Adv. Water Resour.*, **17**, 61–78.
- Koster, R., and M. Suarez (1996), Energy and water balance calculations in the Mosaic LSM, *NASA Tech. Memo.*, **104606**, vol. 9, 60 pp.

- Koster, R., M. Suarez, and M. Heiser (2000), Variance and predictability of precipitation at seasonal-to-interannual timescales, *J. Hydrometeorol.*, **1**, 26–46.
- Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of land surface water and energy fluxes for GCMs, *J. Geophys. Res.*, **99**, 14,415–14,428.
- Liang, X., D. P. Lettenmaier, and E. F. Wood (1996a), One-dimensional statistical dynamic representation of subgrid spatial variability of precipitation in the two-layer variable infiltration capacity model, *J. Geophys. Res.*, **101**, 21,403–21,422.
- Liang, X., E. F. Wood, and D. P. Lettenmaier (1996b), Surface soil moisture parameterization of the VIC-2L model: Evaluation and modifications, *Global Planet. Change*, **13**, 195–206.
- Lohmann, D., et al. (1998), The project for intercomparison of land-surface parameterization schemes (PILPS) phase 2(c) Red-Arkansas river basin experiment: 3. Spatial and temporal analysis of water fluxes, *Global Planet. Change*, **19**, 161–179.
- Lohmann, D., et al. (2004), Streamflow and water balance intercomparison of four landsurface models in the North American Land Data Assimilation System (NLDAS), *J. Geophys. Res.*, **109**, D07S91, doi:10.1029/2003JD003517.
- Luo, L., et al. (2003), Validation of the North American Land Data Assimilation System (NLDAS) retrospective forcing over the southern Great Plains, *J. Geophys. Res.*, **108**(D22), 8843, doi:10.1029/2002JD003246.
- Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen (2002), A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States, *J. Clim.*, **15**, 3237–3251.
- Miller, D. A., and R. A. White (1998), A conterminous United States multi-layer soil characteristics data set for regional climate and hydrology modeling, *Earth Inter.*, **2**, Paper No. 2.
- Nijssen, B., D. P. Lettenmaier, X. Liang, S. W. Wetzel, and E. F. Wood (1997), Streamflow simulation for continental-scale river basins, *Water Resour. Res.*, **33**, 711–724.
- Nijssen, B., R. Schnur, and D. P. Lettenmaier (2001), Global retrospective estimation of soil moisture using the variable infiltration capacity land surface model, 1980–93, *J. Clim.*, **14**, 1790–1808.
- Oki, T., T. Nishimura, and P. Dirmeyer (1999), Assessment of annual runoff from land surface models using total runoff integrating pathways (TRIP), *J. Meteorol. Soc. Jpn.*, **77**, 69–89.
- Pan, H.-L., and L. Mahrt (1987), Interaction between soil hydrology and boundary layer development, *Boundary Layer Meteorol.*, **38**, 185–202.
- Pan, M., et al. (2003), Snow process modeling in the North American Land Data Assimilation System (NLDAS): 2. Evaluation of model-simulated snow water equivalent, *J. Geophys. Res.*, **108**(D22), 8850, doi:10.1029/2003JD003994.
- Pinker, R. T., et al. (2003), Surface radiation budgets in support of the GEWEX Continental-Scale International Project (GCIP) and the GEWEX Americas Prediction Project (GAPP), including the North American Land Data Assimilation System (NLDAS) project, *J. Geophys. Res.*, **108**(D22), 8844, doi:10.1029/2002JD003301.
- Radakovitch, J. D., P. R. Houser, A. da Silva, and M. G. Bosilovich (2001), Results from global land-surface data assimilation methods, paper presented at 5th Symposium on Integrated Observing Systems, Am. Meteorol. Soc., Albuquerque, N. M., 14–19 Jan.
- Ramsay, B. H. (1998), The interactive multisensor snow and ice mapping system, *Hydrol. Processes*, **12**, 1537–1546.
- Reynolds, C. A., T. J. Jackson, and W. J. Rawls (2000), Estimating soil water-holding capacities by linking the Food and Agriculture Organization soil map of the world with global pedon databases and continuous pedotransfer functions, *Water Resour. Res.*, **36**, 3653–3662.
- Roads, J. O., et al. (2003), GCIP water and energy budget synthesis (WEBS), *J. Geophys. Res.*, **108**(D16), 8609, doi:10.1029/2002JD002583.
- Robock, A., et al. (2003), Evaluation of the North American Land Data Assimilation System over the southern Great Plains during the warm season, *J. Geophys. Res.*, **108**(D22), 8846, doi:10.1029/2002JD003245.
- Rodell, M., et al. (2003), global land data assimilation System, *Bull. Am. Meteorol. Soc.*, in press.
- Rogers, E., D. Deaven, and G. Dimego (1995), The Regional Analysis System for the operational “early” Eta model: Original 80-km configuration and recent changes, *Weather Forecast.*, **10**, 810–825.
- Schaake, J. C., et al. (2004), An intercomparison of soil moisture fields in the North American Land Data Assimilation System (NLDAS), *J. Geophys. Res.*, **109**, D01S90, doi:10.1029/2002JD003309.
- Schlosser, C. A., A. G. Slater, A. Robock, A. J. Pitman, J. K. Y. Vinnikov, A. Henderson-Sellers, N. A. Speranskaya, K. Mitchell, and PILPS 2(d) Contributors (2000), Simulations of a boreal grassland hydrology at Valdai, Russia: PILPS Phase 2(d), *Mon. Weather Rev.*, **128**, 301–321.
- Sellers, P., Y. Mintz, Y. Sud, and A. Dalcher (1986), A Simple Biosphere model (SiB) for use within general circulation models, *J. Atmos. Sci.*, **43**, 505–531.
- Sheffield, J., et al. (2003), Snow process modeling in the North American Land Data Assimilation System (NLDAS): 1. Evaluation of model-simulated snow cover extent, *J. Geophys. Res.*, **108**(D22), 8849, doi:10.1029/2002JD003274.
- Slater, A. G., et al. (2001), The representation of snow in land surface schemes: Results from PILPS 2(d), *J. Hydrometeorol.*, **2**, 7–25.
- Tarpley, J. (1979), Estimating incident solar radiation at the surface from geostationary satellite data, *J. Appl. Meteorol.*, **18**, 1172–1181.
- Tarpley, J. D. (1994), Monthly evapotranspiration from satellite and conventional meteorological observations, *J. Clim.*, **7**, 704–713.
- Verdin, K. L., and S. K. Greenlee (1996), Development of continental scale digital elevation models and extraction of hydrographic features, paper presented at Third International Conference/Workshop on Integrating GIS and Environmental Modeling, Natl. Cent. for Geogr. Inf. and Anal. Santa Fe, N. M., 21–26 Jan.
- Vogel, C. A., D. D. Baldocchi, A. K. Luhr, and R. K. Shankar (1995), A comparison of a hierarchy of models for determining energy balance components over vegetation canopies, *J. Appl. Meteorol.*, **34**(10), 2182–2196.
- Wigmosta, M. S., L. W. Vail, and D. P. Lettenmaier (1994), A distributed hydrology-vegetation model for complex terrain, *Water Resour. Res.*, **30**, 1665–1679.
- Wood, E. F., D. P. Lettenmaier, X. Liang, B. Nijssen, and S. W. Wetzel (1997), Hydrological modeling of continental-scale basins, *Annu. Rev. Earth Planet. Sci.*, **25**, 279–300.
- Wood, E., et al. (1998), The Project for Intercomparison of Land-Surface Parameterization Schemes (PILPS) Phase 2(c) Red-Arkansas River basin experiment: 1. Experimental description and summary intercomparisons, *Global Planet. Change*, **19**, 115–136.
- Wu, X., W. P. Menzel, and G. S. Wade (1999), Estimation of sea surface temperatures using GOES-8/9 radiance measurements, *Bull. Am. Meteorol. Soc.*, **80**, 1127–1138.
- Zilitinkevich, S. S. (1995), Non-local turbulent transport: Pollution dispersion aspects of coherent structure of convective flows, in *Air Pollution III*, vol. 1, *Air Pollution Theory and Simulation*, edited by H. Power, N. Moussiopoulos, and C. A. Brebbia, pp. 52–60, Comput. Mech., Billerica, Mass.

A. A. Bailey, B. H. Ramsay, and J. D. Tarpley, Office of Research and Applications, NESDIS, E/RA1 WWBG Room 712, 5200 Auth Road, Camp Springs, MD 20746, USA. (baileya@comcast.net; bruce.ramsay@noaa.gov; dan.tarpley@noaa.gov)

B. A. Cosgrove, J. K. Entin, and P. R. Houser, Hydrological Sciences Branch and Data Assimilation Office, NASA Goddard Space Flight Center, Mail Code 974.1, Greenbelt, MD 20771, USA. (brian.cosgrove@gsfc.nasa.gov; jared.entin@gsfc.nasa.gov; paul.r.houser@nasa.gov)

Q. Duan, V. Koren, and J. C. Schaake, Office of Hydrologic Development, NOAA/NWS, 1325 East-West Highway, SSMC2, Room 8356, Silver Spring, MD 20910, USA. (qingyun.duan@noaa.gov; victor.koren@noaa.gov; john.schaake@noaa.gov)

R. W. Higgins and W. Shi, Climate Prediction Center, National Centers for Environmental Prediction, NOAA/NWS, 5200 Auth Road, Room 605, Camp Springs, MD 20746-4304, USA. (wayne.higgins@noaa.gov; wshi@ncep.noaa.gov)

D. P. Lettenmaier, Department of Civil and Environmental Engineering, University of Washington, Roberts Hall, FX-10, Box 352700, Seattle, WA 98195-2700, USA. (dennisl@u.washington.edu)

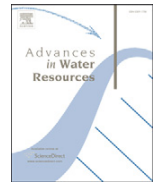
D. Lohmann, J. Meng, and K. E. Mitchell, Environmental Modeling Center, National Centers for Environmental Prediction, NOAA/NWS, 5200 Auth Road, Camp Springs, MD 20746-4304, USA. (dag.lohmann@noaa.gov; jesse.meng@noaa.gov; kenneth.mitchell@noaa.gov)

L. Luo, M. Pan, J. Sheffield, and E. F. Wood, Department of Civil and Environmental Engineering, Princeton University, Room E208, E-Quad, Olden Street, Princeton, NJ 08544, USA. (lluo@princeton.edu; mpan@princeton.edu; justin@princeton.edu; efwood@princeton.edu)

C. H. Marshall, Department of Atmospheric Sciences, Colorado State University, Fort Collins, CO 80523, USA. (curtis@atmos.colostate.edu)

R. T. Pinker, Department of Meteorology, University of Maryland, College Park, 2213 Computer and Space Sciences Building, College Park, MD 20742-2425, USA. (pinker@atmos.umd.edu)

A. Robock, Department of Environmental Sciences, Rutgers University, 14 College Farm Road, New Brunswick, NJ 08901-8551, USA. (robrock@envsci.rutgers.edu)



Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter

Xianhong Xie^{a,*}, Dongxiao Zhang^{a,b,c}

^a Department of Energy and Resources Engineering, College of Engineering, Peking University, Beijing 100871, China

^b Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, California, USA

^c Mork Family Department of Chemical Engineering and Materials Sciences, University of Southern California, Los Angeles, California, USA

ARTICLE INFO

Article history:

Received 17 June 2009

Received in revised form 23 March 2010

Accepted 23 March 2010

Available online 30 March 2010

Keywords:

Data assimilation

Ensemble Kalman filter

Distributed hydrological modeling

Catchment scale

Uncertainty

ABSTRACT

Catchment scale hydrological models are critical decision support tools for water resources management and environment remediation. However, the reliability of hydrological models is inevitably affected by limited measurements and imperfect models. Data assimilation techniques combine complementary information from measurements and models to enhance the model reliability and reduce predictive uncertainties. As a sequential data assimilation technique, the ensemble Kalman filter (EnKF) has been extensively studied in the earth sciences for assimilating in-situ measurements and remote sensing data. Although the EnKF has been demonstrated in land surface data assimilations, there are no systematic studies to investigate its performance in distributed modeling with high dimensional states and parameters. In this paper, we present an assessment on the EnKF with state augmentation for combined state-parameter estimation on the basis of a physical-based hydrological model, Soil and Water Assessment Tool (SWAT). Through synthetic simulation experiments, the capability of the EnKF is demonstrated by assimilating the runoff and other measurements, and its sensitivities are analyzed with respect to the error specification, the initial realization and the ensemble size. It is found that the EnKF provides an efficient approach for obtaining a set of acceptable model parameters and satisfactory runoff, soil water content and evapotranspiration estimations. The EnKF performance could be improved after augmenting with other complementary data, such as soil water content and evapotranspiration from remote sensing retrieval. Sensitivity studies demonstrate the importance of consistent error specification and the potential with small ensemble size in the data assimilation system.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Understanding the response of a river catchment to atmospheric forcing is critically important to climate studies, agriculture (irrigation planning and vegetation and crop growth), natural hazards prevention and mitigation (e.g. floods and droughts), and other water resources managements (e.g. water transfer and storages). These studies are highly dependent on advanced simulation models and large amounts of environmental data that are increasingly being made available. Hydrological models are built based on a set of principles coupled with a number of assumptions and imperfectly defined parameters, and measurements are usually scarce in space and discontinuous in time. These result in great uncertainties about the measurements, model structures and parameters. In applications, numerical hydrological modeling generally requires estimation of

model parameters through calibration with observed data to mitigate the output uncertainties. Many calibration methods have been developed, including the automatic calibration with multiple objectives or criteria [20,24,32]. But most of the calibration methods are to find a set of reasonable parameter values and to attribute all errors to parameter uncertainties [23]. Moreover, general calibration methods are difficult to completely utilize multiple sources of data from remote sensing and automated ground-based sensors. These limitations can be remedied with the recently developed data assimilation techniques.

The hydrological data assimilation method shares the basic tenet of merging models and observations and accounts for uncertainties from different sources of information [21]. The idea has been inspired by and adapted from atmospheric and oceanic data assimilation systems that concern estimation of initial conditions. In contrast, land surface dynamics, including hydrological processes at the catchment scale, are fundamentally damped in nature, and the process assimilation is all about estimating errors in uncertain meteorological forcing conditions and model parameterizations [29]. So the sequential data assimilation methods, such as the Kalman filter (KF) and the

* Corresponding author. Present address: Room 301, Founder Building, No. 298, Chengfu Road, Haidian District, Beijing, China. Tel.: +86 10 62757432; fax: +86 10 82529010.

E-mail addresses: xiexh@pku.edu.cn (X. Xie), donzhang@usc.edu (D. Zhang).

ensemble Kalman filter (EnKF), are more suitable for hydrological applications and it can continually update the hydrological states and model parameters when new data become available. Especially, the EnKF is becoming popular in many areas of earth sciences because it is easy to use, is flexible, and makes relatively few restrictive assumptions [9,10]. It can also account for nonlinearities and partially non-Gaussianity.

The EnKF has been successfully applied in the land surface data assimilation [17,39,40]. The modern adaptive filtering techniques were investigated to address the error estimation problem that is a crucial component in the EnKF [6,29]. However, most of these studies focused on the regional or global scale of interest in hydrometeorology, reflecting on the land surface–atmosphere interactions, while the horizontal movements of water, including the overland flow, stream flow routing and water interactions between sub-catchments, are often omitted in the land surface models. On the other hand, the fundamental operative unit for water resources management is the catchment or river basin. There is therefore a need to extend the applicability of data assimilation in hydrology from the regional land surface processes to catchment scale hydrological issues [37]. At this more local scale, the main issues are to estimate hydrologic parameters (e.g. the permeability) and to characterize hydrologic responses such as runoff, soil water movement, evapotranspiration and groundwater movement.

There have been a few of studies with encouraging results concerning the local scale hydrological state estimation with data assimilation. Kitanidis and Bras [14,15] originally reformulated a nonlinear (lumped) conceptual catchment model into a form amenable to linear estimation for the KF application, and they found that the assimilating feedback significantly improved the real-time forecasting of river discharges even when the model and input error statistics are not perfectly known. Schuurmans et al. [33] applied the KF with a constant Kalman gain to assimilate the remotely sensed latent heat flux for improving the water balance computation, and demonstrated that data assimilation has much potential for analyzing and improving distributed hydrological model prediction. Aubert et al. [1] investigated the extended KF in an operational forecasting context by introducing soil moisture data into stream flow modeling. They also indicated remote sensing data coupled with sequential assimilation to be well adapted to streamflow forecasting. Komma et al. [16] examined the benefits of updating soil moisture with the EnKF in forecasting large floods. Das et al. [8] used an ensemble square root filter (EnSRF) scheme to assimilate the aircraft-based soil moisture observations in a distributed hydrological model.

These studies greatly increase the potential to successfully apply ensemble filtering methods in catchments for hydrological modeling that allows coupled physical processes including precipitation, overland flow, infiltration, evapotranspiration, groundwater and streamflow. The representation of these processes resorts to empirical and physical nonlinear equations instead of well-defined governing equations like the general groundwater models or the land surface models. In distributed hydrological models, unavoidably, there are a large number of states and parameters [28,35], which pose difficulties for assimilating different types of observations even with the EnKF. Hence data assimilation approaches are often used to estimate the dynamic states while the parameters are excluded from the assimilation update group [23]. For examples, Pauwels et al. [26] perturbed parameters associated with TOPMODEL but those parameters for every member were not updated when assimilating runoff to estimate soil moisture, and Clark et al. [5] used streamflow observations to update dynamic hydrological states based on prior calibrated parameters. This could obtain reasonable historical values, but may not be applicable to the long-term hydrological predictions or the environment with time-varying model parameters [23].

The filter techniques implicitly provide a combined estimation for the dynamic state and the parameters by augmenting a joint state

vector [2,8]. The standard KF is limited to a linear dynamic system [23], while the EnKF bypasses this limitation and it has been successfully used in hydrogeology for the parameter estimation with synthetic experiments [4,36] and with field scale flow and transport experimental data sets [18]. In order to reduce the degree of freedom in the joint state vector, a dual state-parameter estimation strategy was developed based on the EnKF and has been demonstrated in conceptual rainfall-runoff models [23], while this dual estimation approach neglects the effect of cross-state and parameter dependencies [3], and it may be limited due to its overwhelming computational burden in the presence of large numbers of computational units.

Even though many studies focus on joint estimation using the EnKF with the state augmentation, little attention has been paid to the distributed hydrological modeling in which the state and parameters are rather high dimensional. The applicability of the state augmentation is also not clear when assimilating multi-site observations with biased error estimation. Therefore, the potential of the EnKF needs further exploration to address the issues about the combined estimation for distributed catchment models.

In this paper, we assess the performance of the EnKF with the state augmentation for a physical-based distributed hydrological model, SWAT, and focus on the parameter estimation of the curve number (CN_2) in the SWAT, as well as the estimation of prognostic variables such as the runoff, soil water content and evapotranspiration. The parameters are assembled into an augmented state vector because there are strong relations between parameters and dynamic states. We do so on the basis of synthetic examples. Multiple types of measurements are introduced and their effects are discussed. In order to better demonstrate its performance, sensitivities from the error prescription, the initial realization and the ensemble size are investigated.

The remainder of the paper is organized as follows. After a brief review of the SWAT model and the EnKF method in Section 2, we describe the design of the data assimilation system and illustrate the synthetic experiments in Section 3. Diagnostics of the capability and sensitivity of EnKF are discussed in Sections 4 and 5, respectively. Conclusions and a summary of future work are given in Section 6.

2. Sequential data assimilation

2.1. Dynamic hydrological process

With the time evolving, a hydrological system is integrated with dynamic states and static parameters under driving forces. A dynamic hydrological model can be expressed as a nonlinear stochastic process [2],

$$X_{t+1} = f(X_t, U_{t+1}) + \omega_{t+1} \quad (1)$$

where t denotes the time step, X is an augmented state vector consisting of dynamic variables (e.g., water content in the soil profile) and static parameters (e.g., hydraulic conductivity), f is the nonlinear hydrological model forward operator, U is a set of externally specified time-dependent forcing variables (e.g., precipitation), and the noise term ω accounts for model errors that represent all uncertainties related to model structure, forcing variables and parameters.

Observations using different instruments or techniques can be obtained from the hydrological system. It could be directly or indirectly related to the hydrological states and be written as,

$$Y_t = H(X_t) + \varepsilon_t \quad (2)$$

where Y denotes the observation vector from measurement instruments, H is the linear or nonlinear observation operator specifying deterministic relationship between the observation data and the true state X , and ε is a noise term accounts for both measurement error

(instrument and processing errors in the measurement) and representation error (introduced by imperfect observation model H), like the random variable ω in Eq. (1). In general, these noise terms are assumed to be Gaussian.

In order to get optimal estimates for hydrological states and parameters of interest, we should combine the complementary information from hydrological models and observations, which is the basic idea of data assimilation. Sequential assimilation algorithms march forward in time, alternating between model propagation and data assimilation update whenever observations become available.

2.2. Framework of ensemble Kalman filter

The EnKF is based upon Monte Carlo method and the Kalman filter formulation to approximate the true probability distribution of the model state, conditioned on a series of observations of the model states [3]. Instead of computing the error covariance explicitly, the probability density of model states is represented by a large ensemble of model states, including parameters, and each member of the ensemble is integrated forward in time by the model independently. Similar to Eq. (1), the model forecast is executed in the EnKF for each ensemble member as follows:

$$X_{t+1}^i = f(X_t^i, U_t^i) + \omega_{t+1}^i, \omega_{t+1}^i \sim N(0, Q_{t+1}), i = 1, \dots, n \quad (3)$$

$$U_{t+1}^i = U_{t+1} + \xi_{t+1}^i, \xi_{t+1}^i \sim N(0, R_{t+1}) \quad (4)$$

where n is the ensemble size, namely the number of ensemble members, X_{t+1}^i is the component of the i th ensemble member forecast at time $t+1$, X_t^i is the i th updated ensemble member at time t , U_t^i is the i th perturbed forcing variables, ω_{t+1}^i and ξ_{t+1}^i are independent white noises for the forecast model and forcing terms, drawn from multi-normal distributions with zero mean and specified covariance Q_{t+1} and R_{t+1} . At time $t+1$, the observation ensemble member can be written as,

$$Y_{t+1}^i = HX_{t+1}^i + \varepsilon_{t+1}^i, \varepsilon_{t+1}^i \sim N(0, S_{t+1}) \quad (5)$$

where HX_{t+1} is a simplification of $H(X_{t+1})$, which is the original observation vector obtained from the true hydrological catchment state X_{t+1} , and ε_{t+1}^i is the noise term with zero mean and specified covariance S_{t+1} .

With the model forecasts and observations being available, the assimilation or updating process can be expressed as,

$$X_{t+1}^+ = X_{t+1}^- + K_{t+1}(Y_{t+1}^i - HX_{t+1}^-) \quad (6)$$

$$K_{t+1} = P_{t+1}^- H^T (HP_{t+1}^- H^T + S_{t+1})^{-1} \quad (7)$$

$$P_{t+1}^- \approx \frac{1}{N-1} \sum_{i=1}^N [(X_{t+1}^i - \langle X_{t+1}^- \rangle)(X_{t+1}^i - \langle X_{t+1}^- \rangle)^T]^{-1} \quad (8)$$

$$\langle X_{t+1}^- \rangle = \frac{1}{N} \sum_{i=1}^N X_{t+1}^i \quad (9)$$

where X_{t+1}^+ is the new optimal estimate vector of hydrological states after assimilation, K_{t+1} is the Kalman gain, which determines the weight between the modeling and observation states in assimilation processes, P_{t+1}^- , approximated with the ensemble, is the prior model error covariance. Similarly, the updated (posterior) error covariance could be estimated with the updated ensemble members. This covariance is not indispensable for the EnKF, whereas its estimation is necessary in the standard Kalman filter.

As shown in Eqs. (8) and (9), the EnKF algorithm uses the first and second order moments to represent the probability density, which means the Gaussian hypothesis is implicit for the model states, including parameters. Even though this hypothesis is often violated for the hydrological modeling, the EnKF appears to provide a good approximation for nonlinear and non-Gaussian land surface problems [40].

3. Data assimilation setup

3.1. SWAT model and hydrological state

The data assimilation algorithm is implemented on a popular distributed hydrological model, Soil and Water Assessment Tool (SWAT). Since it is physically based and computationally efficient, uses readily available inputs, and enables users to study long-term impacts, this model has been widely used to predict impacts of land management practices on water, sediment, and agricultural chemical yields in large, complex watersheds with varying soils, land use, and management conditions over long periods of time [11,19,25]. For modeling purposes, a catchment is partitioned into a number of sub-catchments or sub-basins according to the property of Digital Elevation Model (DEM) data, and then input information for each subbasin is grouped or organized into different hydrologic response units (HRUs) that are comprised of unique land cover, soil, and management combinations.

The land phase of the hydrologic cycle is simulated at the HRU scale based on a water balance equation with a daily step:

$$SW_t = SW_0 + \sum_{i=1}^t (R_i - Q_{surf,i} - ET_i - W_{seep,i} - Q_{gw,i}) \quad (10)$$

where SW_t is the final soil water content (mm H₂O), SW_0 is the initial soil water content (mm H₂O), t is the time (days), R_i is the amount of precipitation on day i (mm H₂O), $Q_{surf,i}$ is the amount of surface runoff on day i (mm H₂O), ET_i is the amount of evapotranspiration on day i (mm H₂O), $W_{seep,i}$ is the amount of percolation and bypass flow exiting the soil profile bottom on day i (mm H₂O), and $Q_{gw,i}$ is the amount of return flow on day i (mm H₂O). In this equation, the term of $Q_{surf,i}$ is the main component that determines streamflow in reaches and the soil moisture in soil profiles. It can be expressed with an empirical model, SCS runoff equation:

$$Q_{surf,i} = \frac{(R_i - 0.2S_i)^2}{R_i + 0.8S_i} \quad (11)$$

where S_i is the retention storage of soil profiles (mm H₂O). It is empirically defined as:

$$S_i = \frac{25400}{CN_i} - 254 \quad (12)$$

where CN_i is the temporal curve number for the day. In order to capture the temporal-spatial variations of the retention storage (S_i) due to changes in soil water content, soil properties, land use, management and slope, the temporal curve number is implicitly expressed as:

$$CN_i = f(CN_2, SW_i) \quad (13)$$

where CN_2 is the static curve number corresponding to the average moisture condition, and it depends on the soil's permeability, land use and antecedent soil water conditions [25]. Its value is often between 30 and 100. It is a dominant parameter for the surface runoff generation in an HRU, and consequently influences the streamflow process and the soil moisture states.

After the surface runoff generation, other processes including soil percolation, evapotranspiration, groundwater and seepage are carried out at the HRU level and subsequently water from these processes is aggregated to the subbasin level. The total runoff generated in a subbasin is routed through the channel network using the variable storage routing method or the Muskingum River routing method [19,25].

Although the SWAT model is capable of simulating other processes, such as the plant growth and the sediment movement, here we exclusively focus on the water movement in catchments. Moreover, we select 11 primary state variables to diagnose the performance of data assimilation on this model, even though there are hundreds of state variables to support the model run.

As shown in the Table 1, the first seven general states, including water stored or lagged in the soil profile and the drainage reaches, are model-dependent variables to characterize water storage conditions in HRUs or in subbasins. The subsequent three dynamic variables, Runoff (R), Soil water content (SW) and Evapotranspiration (ET) are viewed as observable or prognostic states since they could be measured with specific instruments or techniques. In addition, there are a large number of static variables (parameters) needed to be identified for a simulation application, and each of them may make a different contribution. However, the SCS runoff curve number, CN_2 , is the most sensitive parameter and its uncertainty could induce great effects on runoff modeling and other hydrological processes [13,24,32,34]. It is often needed to be estimated by calibration methods or data assimilation methods. In the SWAT model, it should be mentioned that the curve number is used to compute the maximum retention value and other two shape coefficients [25]. This computational process is implicitly expressed with Eq. (13). The three parameters will be continuously corrected with the CN_2 updated in the data assimilation process.

3.2. Experimental area description

In this study, an experimental catchment is adopted in the Lake Fork Watershed in Northeast Texas, USA. This catchment area covers 489.85 km² and the altitudes of the basin vary between 106 m and 195 m above sea level. The pasture (51.3%), range-grasses (28.4%) and deciduous forest (16.0%) are the three main land use types.

The data set including topographic information, land use information, soil data, daily precipitation, and climate records is taken from the SWAT visual software modular [19]. As a synthetic data assimilation experiment, nevertheless, the data about stream runoff

(streamflow), soil water content and actual evapotranspiration in subbasins are drawn from a reference simulation that will be described in the next subsection.

Based on a digital elevation model, the catchment is partitioned into 20 subbasins and subsequently 49 HRUs according to the land use and soil type information. The soil profile is divided into seven layers with different soil properties. As shown in Fig. 1, the serial number of subbasins is not in a natural sequence because some outlets are added to get approximately uniform subbasin areas. This is beneficial to the hydrological simulation and the data assimilation. Consequently, the joint state vector consists of 717 such variables as stated in Table 1.

3.3. Data assimilation procedure

Synthetic experiments are designed to assess the capability and sensitivity of the assimilation process to model parameterizations and physical representations, and this is also a general and effectual way to investigate data assimilation performance on predictive systems as shown by Chen and Zhang [4] and Kumar et al. [17]. Since the true parameter values are known in the experiments, we can easily assess the performance of the assimilation approach. In these experiments, one reference field is randomly picked up from a predefined Gaussian distribution with given statistics. And then the hydrological model is integrated to obtain the reference state (or “true” state), referred to as the “control run” or reference modeling. At the measurement locations, the observations (e.g. the runoff) to be used in the assimilation system are drawn from the reference state. Finally, the assimilation integrations are conducted by introducing the synthetic observations on the basis of a stochastic modeling platform.

In the hydrological modeling, a warming up process is often needed to initialize the model state and harmonize hydrologic responses with the meteorological forcing, and thus the warming up period should be set before the assimilation performs. However, the computation may be expensive to run an ensemble with a large number of modeling members through a long warming up period, which is often more than one year. Alternatively, we could run only one member of the ensemble to warm up (spin up) the model and then perturb the meteorological forcing terms and generate stochastic realizations of the parameters. In this study, three running periods are

Table 1
Selected dynamic states and parameters for the assimilation update.

Order	State variable	Description	Scale	Class
1	$Q_{sufstor}$	Amount of surface runoff stored or lagged	HRU	V
2	$Q_{latstor}$	Amount of lateral flow stored or lagged	HRU	V
3	Q_{shall}	Amount of shallow water stored or lagged	HRU	V
4	Q_{rchrg}	Amount of recharge entering the aquifer	HRU	V
5	Q_{pregw}	Amount of groundwater flow into the main channel	HRU	V
6	W_{sol}	Amount of water stored in the soil layer for each HRU	$HRU \times N_{lay}$	V
7	W_r	Amount of water stored in the reach	subbasin	V
8	R	Amount of water flow (Runoff) out of a reach	subbasin	V_o
9	SW	Amount of water in soil in a subbasin	subbasin	V_o
10	ET	Amount of actual evapotranspiration in subbasin	subbasin	V_o
11	CN_2	SCS runoff curve number for moisture condition II	HRU	P

Note: The class V denotes general state variables, V_o denotes the observable or prognostic variables, and P denotes principle parameters; N_{lay} is the number of layers of soil profiles in a hydrologic response unit (HRU).

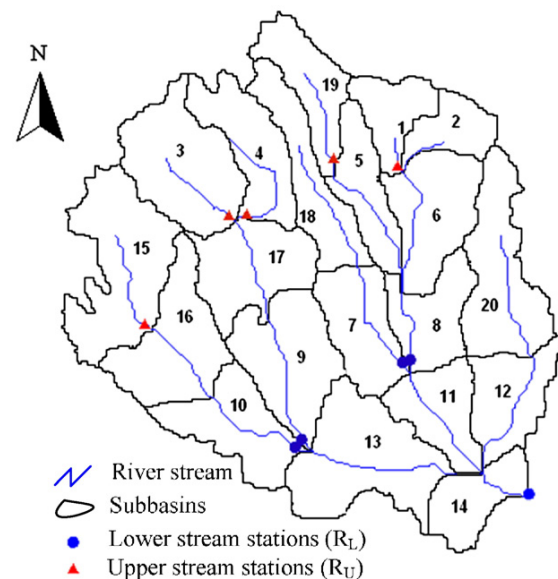


Fig. 1. Subbasin distribution of the experimental catchment from the Lake Fork Watershed in Northeast Texas, USA.

specified for the data assimilation procedures: (1) Warming up period: the model propagates forward with the same boundary conditions (meteorological driving force) as the “control run”, and the parameters (here is the curve number, CN_2) could be specified randomly; (2) Perturbation period: At the beginning of this period an ensemble of stochastic parameter fields are generated using a predefined Gaussian distribution, and the meteorological condition is perturbed using additive white noise throughout this period. Each member of the ensemble propagates forward independently and the hydrological states are simulated given the continuously perturbed the driving force; (3) Data assimilation period: After the ensemble has been generated and initialized, the hydrological states are updated by assimilating observations drawn from the “control run”. The modeling states, as well as the driving force, are perturbed with white noises throughout this period. Clearly, these procedures are not only applicable to the synthetic experiments in this study but also could be carried out in a real-world application. Furthermore, when there is a poor knowledge about the initial hydrological conditions of a catchment for short-period modeling, the first two procedures are able to create reasonable distributions of hydrological states to approximate the initial conditions and be beneficial to the subsequent data assimilation operations. For long period modeling, however, there is no need to perturb meteorological conditions during the perturbation period, since the impact from poor initial conditions on assimilation performances may not be significant owing to the damping nature of hydrological systems [29].

In this study, the length of the control run is three years, 1095 days in total. As to the data assimilation modeling, correspondingly, the first 445 days is set as warming up period, and then it is the perturbation period (from 446th to 455th day). Without exceptional specification, subsequently, the assimilation is carried out from the 456th to the 1095th time step, 640 days in total. It should be reminded that the data assimilation method used here focuses on the estimations of CN_2 and prognostic variables exclusively, and other parameters in the control run and assimilation run hold identical values.

In terms of the landuse and soil types in our study area, the values of the curve numbers (CN_2) for each HRU are around 75.0 [25], and uncorrelated stochastic realizations of CN_2 s are generated for the 49 HRUs by specifying Gaussian distributions, $N(75.0, 5.0^2)$. The assumption of the Gaussianity is made for the convenience and should be subject to further investigation. Moreover, the standard deviation is empirically set as 5.0 that is large enough to account for their uncertainties, because the range of CN_2 s are not beyond (60, 90) in this area according to the reference values [25]. Based on this statistics, the Latin hypercube method [12,27] is used to generate a set of realizations. Fig. 2 shows 200 realizations and the black squares are randomly selected as a reference parameter field to simulate the “true” states under the control run.

Since the first seven general variables in Table 1 are temporary and their model errors are difficult to be identified, we assume they are free of model errors in our experiments. In contrast, the prognostic variables, namely the runoff (R), evapotranspiration (ET) and soil water content (SW), are perturbed by additive white noises to represent the possible measurement and model errors. These noises are mutually uncorrelated in space and time with means equaling to zero (indicating unbiased estimations) and standard deviations scaling to the current values of variables. Likewise, the precipitations are perturbed, and the standard deviations also equal to the products of a scaling factor and the current precipitation values. Furthermore, two different scaling factors for precipitation perturbations could be used in the perturbation period (period 2) and the assimilation period (period 3). Specifically, in the period 2, giving a larger scaling factor (here it is 0.30) can generate broad spreads of precipitation and hydrological response to account for uncertainty in the initial conditions when the prior knowledge is poor. Without exceptional

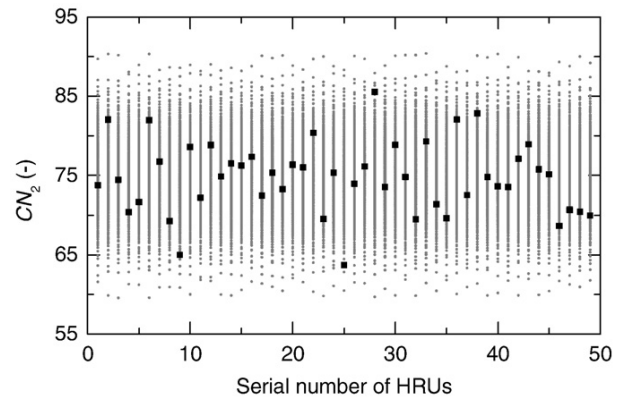


Fig. 2. Stochastic realizations of the CN_2 s with Gaussian distribution, $N(75.0, 5.0^2)$, 200 members.

specification, the ensemble size is 200 for all cases, and detailed specifications are shown in Table 2.

It should be mentioned that these errors may be prescribed in a number of ways. One way is to specify time-dependent errors [9,30] or even to use a modern adaptive filter algorithm [6,29]. But the other way described above is much more straightforward and operational and has been widely used in data assimilation works [5,17]. In addition, the hydrological system in our study is free of errors, since both of the reference modeling and the data assimilation modeling are implemented based on the same model (SWAT), and the observations are directly drawn from the reference simulation. In fact, the success of EnKF is quite dependent on the way of error prescription that should be consistent with the real uncertainties of the model and the measurement [9,10,26]. Therefore, the errors prescribed with scaling factors (Table 2) have overestimated the real uncertainties of the interested hydrological system, and this inconsistency is favorable to examine the robustness of EnKF.

4. Capability of EnKF

In practice, runoff is a comprehensive response to the hydrological cycle and it could be easily obtained from hydrological stations. We take it as a preferred observation data in the assimilation work as it is the most general type of measurement in catchments. Moreover, the remote sensing products, for example the soil moisture retrieval data, play more and more important roles in hydrological modeling with the recent technology development and they could be considered as complementary data sets in data assimilations.

4.1. Runoff measurement

We first take five runoff observations in the lower catchments as shown in Fig. 1, from the subbasins 7, 8, 9, 10 and 14. The runoff observations are drawn from reference modeling and assimilated at every time step. The initial realizations of curve number (CN_2) are generated with Gaussian distribution $N(70, 5.5^2)$, which is biased with respect to the true solution, representing an initial guess to (prior knowledge of) the unknown true properties. It is also feasible to specify relatively accurate estimation of CN_2 for each HRU according to prior knowledge (soil and land cover properties), due to the assumption of independence for all HRUs. But here we simply set identical distributions for all HRUs. In order to assess the capability of EnKF, we also prescribe relatively small scaling factors for the standard deviations of the errors: 0.001 for the observed runoff and 0.01 for the precipitation and the three model-based prognostic variables. These detailed specifications are exhibited as Case 1 in

Table 2

Statistical parameter specification for the EnKF assimilation cases and the control run.

Case	Measurement type	CN_2 distribution	S_o for observation			S_m for model			S_p for precipitation
			R	SW	ET	R	SW	ET	
True		$N(75, 5.0^2)$							
1	R_L	$N(70, 5.5^2)$	0.001			0.01	0.01	0.01	0.01
2	R_U	$N(70, 5.5^2)$	0.001			0.01	0.01	0.01	0.01
3	R_O	$N(70, 5.5^2)$	0.001			0.01	0.01	0.01	0.01
4	$R_O + SW$	$N(70, 5.5^2)$	0.001	0.001		0.01	0.01	0.01	0.01
5	$R_O + ET$	$N(70, 5.5^2)$	0.001		0.001	0.01	0.01	0.01	0.01
6	R_L	$N(70, 5.5^2)$	0.05			0.01	0.01	0.01	0.01
7	R_L	$N(70, 5.5^2)$	0.1			0.01	0.01	0.01	0.01
8	R_L	$N(70, 5.5^2)$	0.1			0.1	0.1	0.1	0.1
9	R_L	$N(70, 5.5^2)$	0.3			0.3	0.3	0.3	0.3
10	R_L	$N(70, 5.5^2)$	0.4			0.4	0.4	0.4	0.4
11	R_L	$N(70, 7.0^2)$	0.05			0.03	0.03	0.03	0.05
12	R_L	$N(65, 7.0^2)$	0.05			0.03	0.03	0.03	0.05
13	R_L	$N(65, 5.5^2)$	0.05			0.03	0.03	0.03	0.05

Note: S_o , S_m and S_p are the scaling factors for error standard deviations corresponding to variables; R_L , R_U and R_O denote the runoff observations from five lower, five upper streams and the catchment outlet, respectively; SW and ET denote observations of the soil water content and the evapotranspiration; “+” denotes a combination of two kinds of measurements. The SW and the ET observations are only assimilated at the 1st, 31st, 61st, 91st and 121st assimilation steps.

Table 2. Impacts of larger scaling factors will be discussed in the next section.

4.1.1. Parameter estimation

After assimilation under the conditions described above, scatter-plots for the ensemble means of the CN_2 s vs. the reference field are depicted in Fig. 3 for the initial, the 60th, the 180th and the 640th

assimilation steps. It is apparent that the initial estimates of CN_2 s are about 70 and do not have any resemblance of the reference field. Based on this set of initial realizations, the model replicates run to the next time step and are assimilated with the five runoff observations. The scatter points of the final result are around the 45-degree line, which indicates a good agreement between the ensemble means of CN_2 s and the true values. Even just at the 180th assimilation step, the

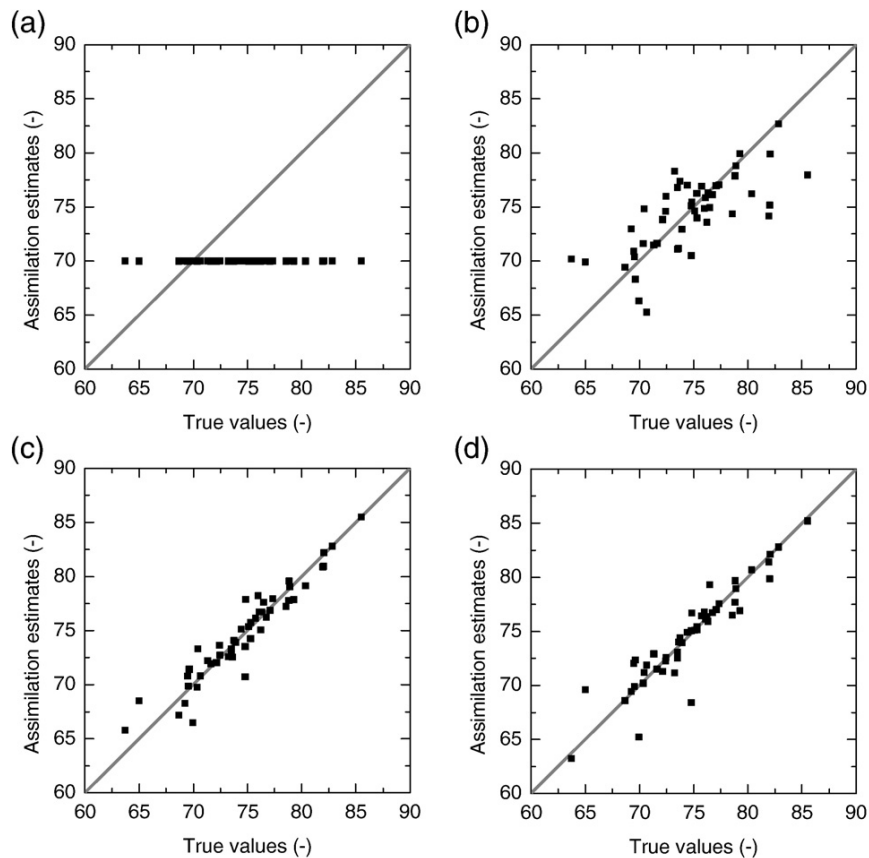


Fig. 3. Assimilation estimates of the CN_2 s vs true (reference) values for Case 1: (a) ensemble means of initial realizations; (b) ensemble means at the 60th assimilation step; (c) ensemble means at the 180th assimilation step, and (d) ensemble means at the 640th assimilation step (final results).

ensemble means capture the major features of the reference field, and there is a slight improvement for the CN_2 estimation in the rest of assimilation steps.

It is of interest to note that some estimates are still discrepant to the true values although the prognostic variables are matched very well. This implies the well-known nonuniqueness in the inverse problems. Specifically, the biased estimation values of CN_2 s could provide acceptable hydrological responses, such as the runoff estimation. As typical examples taken from the 7th HRU in subbasin 5 and the 28th HRU in subbasin 10 (Fig. 4), the ensemble means of the CN_2 s match the true values very well after 120 assimilation steps. In contrast, for the 31th HRU in the subbasin 11, the value of CN_2 is underestimated, and the ensemble spreads to a broad range.

4.1.2. Prognostic variable estimation

Getting accurate estimation for the prognostic variables is an objective of the hydrological simulation and the data assimilation. To compare the performance of the assimilation integrations, open loop simulations are conducted using the same initial realizations of the CN_2 s as well as the uncertain precipitation inputs but without assimilating the model with observations. Thus each member of the ensemble integrates forward independently without updating parameters and variables. Root mean square errors (RMSEs) with respect to the true simulation are computed based on the ensemble means of the three prognostic variables to assess the estimation accuracy.

Fig. 5 shows average time series of RMSEs of prognostic variables for the entire modeling domain. It is clear that the data assimilation in Case 1 reduces the RMSEs of the runoff and soil water content estimations. Especially in peak runoff occurring periods, the RMSEs are lower than that in open loop simulations. However, the assimilation does not give any significant improvement for the evapotranspiration estimates as the two RMSE series keep at the same levels approximately. Evapotranspiration is a complex process containing the soil water evaporation, plant transpiration and canopy

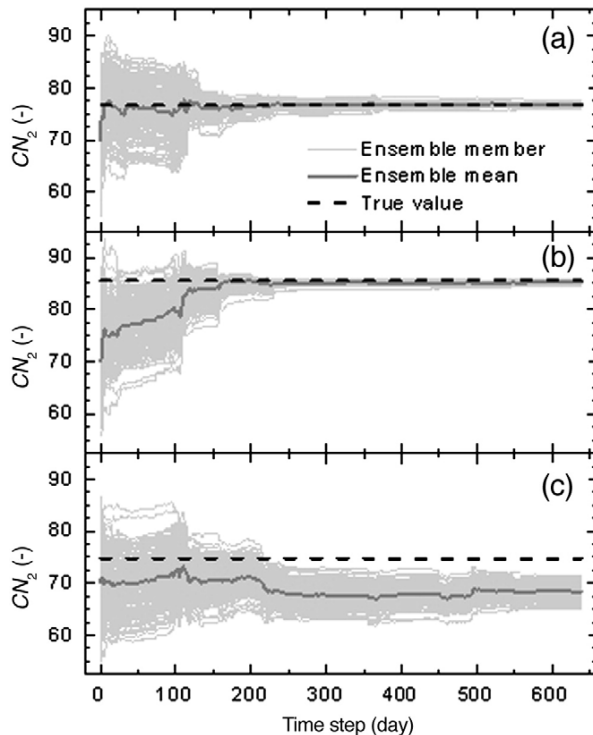


Fig. 4. Typical assimilation processes of the CN_2 s: (a) for HRU 7, Subbasin 5; (b) for HRU 28, Subbasin 10; and (c) for HRU 31, Subbasin 11.

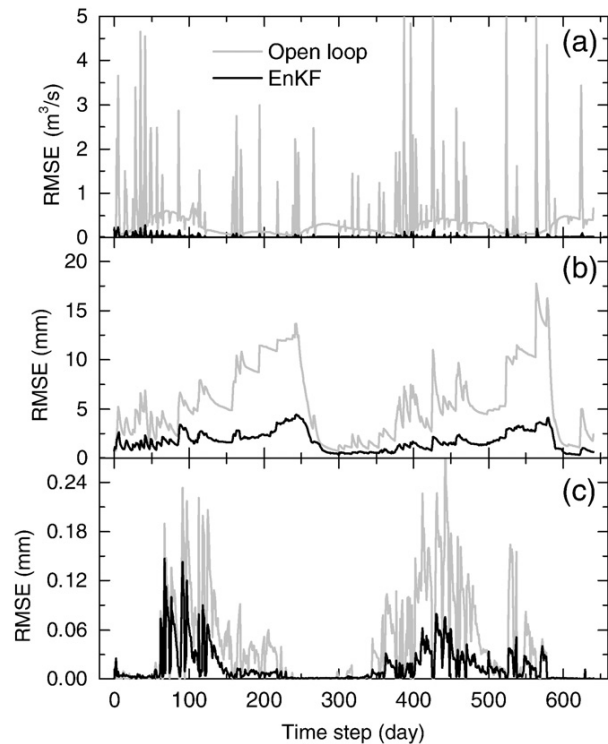


Fig. 5. RMSE series of the three prognostic variables based on the data assimilation (Case 1) and the open loop simulation: (a) runoff, (b) soil water content, and (c) evapotranspiration.

free water evaporation, and these processes are dominated by many parameters and variables. In this study, however, we only focus on the curve number (CN_2) estimation that dominates the surface runoff generation and soil water balance (Eqs. (10)–(13)) and indirectly influences the evapotranspiration processes. In order to diagnose the performance of evapotranspiration estimation, other parameters related to these processes may be needed under this assimilation condition.

Furthermore, taking subbasin 2 as a typical example which is far away from the observation locations, we exhibit the results monthly in order to distinguish the discrepancies of the EnKF and the open loop. It shows that the data assimilation estimates provide better matches with true processes than do the open loop simulations in which great discrepancies are present at the peak value points (Fig. 6). In fact, these results are general exhibitions that are shared by the other subbasins. Consequently, the data assimilation is capable of reducing the simulation uncertainties of these prognostic variables and improving their estimation accuracy.

4.2. Impact of runoff measurement sites

Runoff measurement stations may be assembled at the upper streams or the lower streams in a catchment. Runoff processes at different measurement sites are impacted by hydrological responses from their upper regions. Thus, it is necessary to examine the impact of runoff measurement sites on data assimilation. Two sets of data assimilation runs are conducted: (1) The runoff observations from the five upper streams, in subbasins 1, 3, 4, 15 and 19, are assimilated into the SWAT model at every simulation time step, and (2) the runoff observation only from the outlet at subbasin 14 is assimilated. As described in Case 2 and Case 3 in Table 2, other assimilation conditions

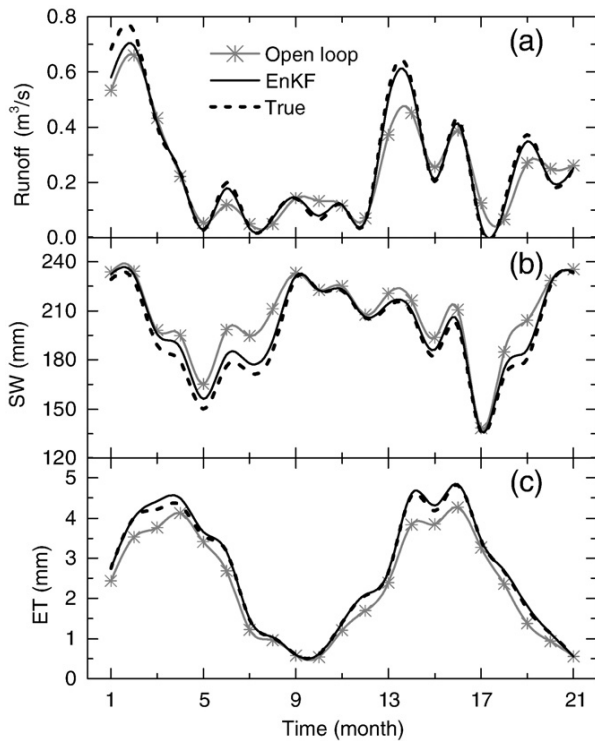


Fig. 6. Monthly true processes and estimations from the data assimilation and the open loop simulation in the subbasin 2: (a) runoff, (b) soil water content, and (c) evapotranspiration.

are not changed. The initial ensemble means for the two cases are still about 70 as shown in Fig. 3a.

Fig. 7 exhibits the final results of Case 2 and Case 3. In Case 2, most of the points are below the 45-degree line, which indicates that the values of CN_2 s are underestimated and that reasonable estimates for CN_2 s are not obtained by assimilating the five observations from the upper streams (Fig. 7a). However, the Case 3 shows much better results since the assimilated-true points are around the 45-degree line (Fig. 7b).

When comparing the RMSE series of Case 1, Case 2 and Case 3 as shown in Fig. 8 (a), we can see that the RMSE level obtained with the five lower observations is lowest in the assimilation period, while that obtained with the upper observations is at the highest level. Furthermore, another favorable error measure, mean absolute error (MAE), is employed as it describes a natural error magnitude and avoids the effect of long-tail probability distributions [38]. However, the MAEs show identical patterns to the RMSEs (Fig. 8), which means long-tail probability distribution does not exist evidently and the parameter estimates can be improved by assimilation observations. Particularly, more reasonable results could be obtained by assimilating the runoff observations from the lower stream stations. This trait attributes to the fact that the runoff from lower streams, especially the catchment outlet, contains the full information about the hydrological states of the upper regions.

4.3. Multiple measurements

There are many types of measurements that can be combined to estimate the hydrological states. In the measurement family, hydrologic remote sensing plays an important role since it can provide much more information about the land surface conditions at large scales, for example the radiance, surface soil moisture and evapotranspiration. Even though the EnKF method is capable of

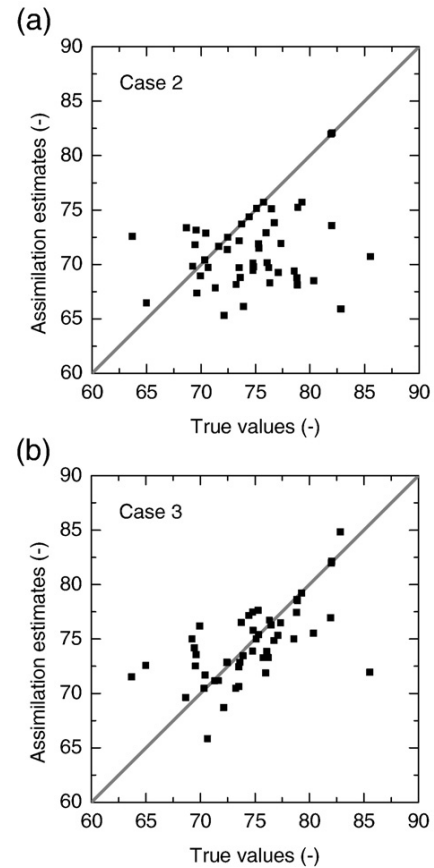


Fig. 7. Final results of the CN_2 estimates vs true (reference) values: (a) Case 2 and (b) Case 3.

assimilating the satellite radiances directly with nonlinear observation operators (as Eq. (2)), this treatment is too complex to get successful application for poor knowledge of radiative transfer processes [31]. Alternatively, it is possible to assimilate the remote sensing retrieval data that is considered as complementary information when the runoff observations are inadequate.

To test this idea, Case 4 and Case 5 are designed based on Case 3 as shown in the Table 2. In addition to the runoff observation from the outlet (subbasin 14), the soil water content and evapotranspiration of the each subbasin, drawn from the reference modeling, are assimilated at the 1st, 31st, 61st, 91st and 121st time steps. Hence, the measurement groups are multiple: five-time data of soil water content are added to the single runoff observation for Case 4, and evapotranspiration for Case 5 likewise. As the treatment on the runoff observations, the scaling factors of standard deviations of these new observations are prescribed with 0.001 to account for their uncertainties in our synthetic experiments, even though this value may be too optimistic for practical measurements.

It is clear from the Fig. 8 that assimilating multiple types of measurements reduces the RMSEs and the MAEs of CN_2 estimates and leads to systematic improvements over the case of only assimilating the outlet runoff (Case 3). The improvement of Case 5 (adding evapotranspiration observations) is not as significant as that of Case 4 since the relationship between the CN_2 and the evapotranspiration is weak as explained before. Fig. 9 gives final results of Case 4 and Case 5 for the CN_2 s estimates vs. the true values, indicating that their matches are acceptable despite the nonuniqueness and nonlinearity between parameters and hydrological responses. Moreover, Case 4 provides the best estimates about the three prognostic variables among the first five

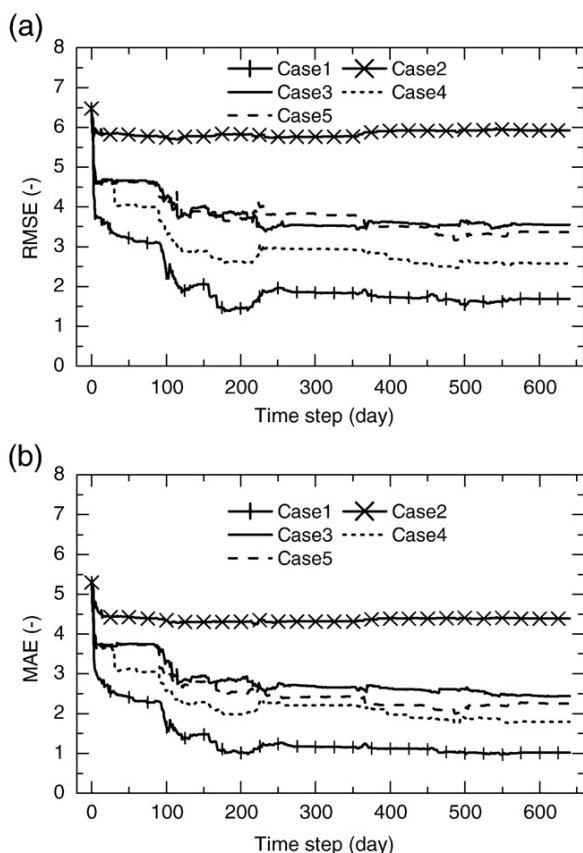


Fig. 8. Root mean square errors (RMSEs, (a)) and mean absolute errors (MAEs, (b)) of the CN_2 estimations for the first five cases.

cases after 100 assimilation steps even though its estimation accuracy for CN_2 is not comparable to Case 1 (Figs. 10 and 5). In fact, the soil water moisture condition plays an important role in controlling the surface runoff generation and the evapotranspiration processes, as expressed in Eq. (10) through Eq. (13). When its observations are assimilated, for example at the first assimilation step, the water storages in all subbasins approach the reference (true) values and these updated states result in acceptable estimates on the runoff and evapotranspiration at the second and subsequent steps. Therefore, the soil water content observation provides complementary data to the assimilation systems, even though it may not be available or utilized at every time step.

5. Sensitivity of EnKF

The EnKF is a Monte Carlo approximation based on a sequential Bayesian filtering process. When constructing a data assimilation framework, we should first consider various aspects, mainly including the model and observation errors, the prior knowledge about the parameter distribution as well as the ensemble size, because these factors have critical impacts on the accuracy and performance [4,17]. In this section, we diagnose sensitivities of the EnKF and focus on the performance with respect to the three factors. The detailed design for the cases is given in Table 2.

5.1. Error specification

The EnKF requires estimates of the model and observation errors to properly merge model predictions with observations [17], while the accurate specification of the errors is generally difficult because

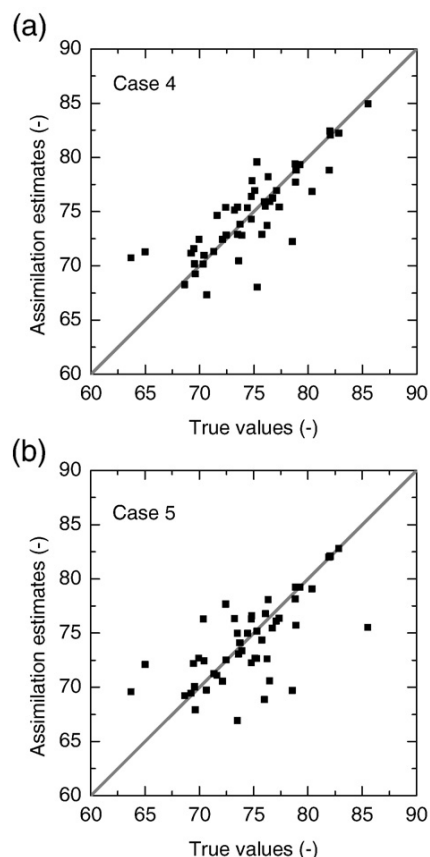


Fig. 9. Final results of the CN_2 estimates vs true (reference) values: (a) Case 4 and (b) Case 5.

the source and statistical structure of these errors are often unknown [7]. Adaptive filtering approaches addressing this problem are capable of estimating model and observation error covariance information during the online cycling of a data assimilation system [6]. To date, however, these approaches have received little attention and been rarely applied to hydrological modeling [29]. Here we hence prescribe the error noises explicitly by representing the standard deviations scaling to the current values of the variables, as described in Section 3.3. This treatment holds an implicit assumption that larger values of model and observation variables could introduce greater errors. In actual applications, it is preferable to overestimate rather than underestimate errors as the underestimation may result in filter divergences [5]. So here we examine the impacts of large errors by varying the scaling factors and focus on the parameter (CN_2) estimation.

The scaling factors (S_o and S_m) for the observation and the model as well as the precipitation (S_p) are changed intentionally to represent different error magnitudes. Five cases (from Case 6 to Case 10) are designed based on Case 1 in order to identify their contributions to the assimilation system. Note that all errors in these cases are overestimated compared to the true ones in the error-free synthetic hydrological system.

Fig. 11a illustrates that the assimilation estimates of the CN_2 s for Case 6 ($S_o=0.05$) agree with their true values very well, and the result is comparable to that of Case 1 with small errors. When the three scaling factors increase to 0.1 (Case 8), the result is suboptimal as the matching points are symmetrically distributed around the 1:1 line (Fig. 11b); and further the assimilation estimates will only capture synoptic properties of the distribution of true values as the

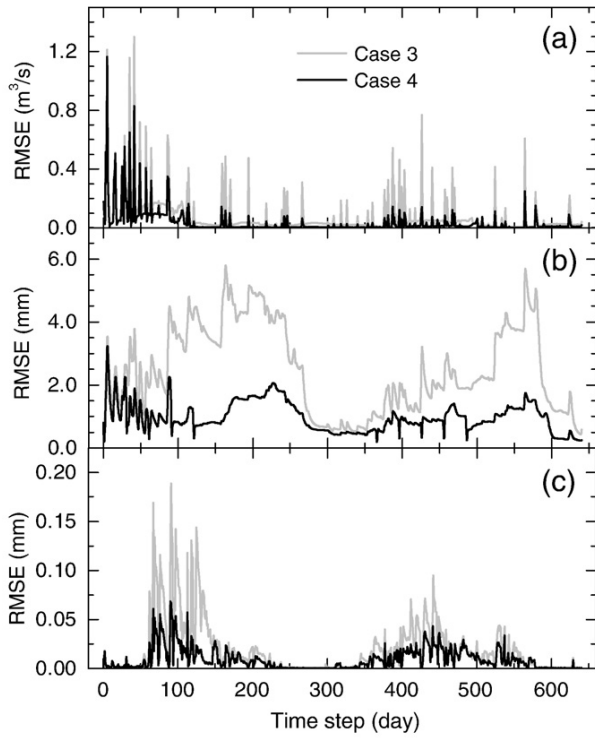


Fig. 10. RMSE series based on Case 3 and Case 4 for (a) runoff, (b) soil water content, and (c) evapotranspiration.

scaling factors reach 0.4 (Case 10, Fig. 11c). Therefore, it is preferable that the overestimated level of the errors should be kept below 10% (corresponding to the 0.1 for the scaling factors), and larger levels may degrade the performance of EnKF for a catchment hydrological system [7]. These characteristics are further interpreted by Fig. 12 that describes the estimated errors of the CN_2 for the five cases. As to Case 6, for example, the RMSE and the MAE series drop dramatically within the first 100 assimilation steps, and the final errors approximate to the results of Case 1. Moreover, with the increase of the scaling factor, the series of the RMSE and the MAE would go up gradually.

In addition to the error magnitude, the data assimilation performance also depends on a consistent combination of errors for

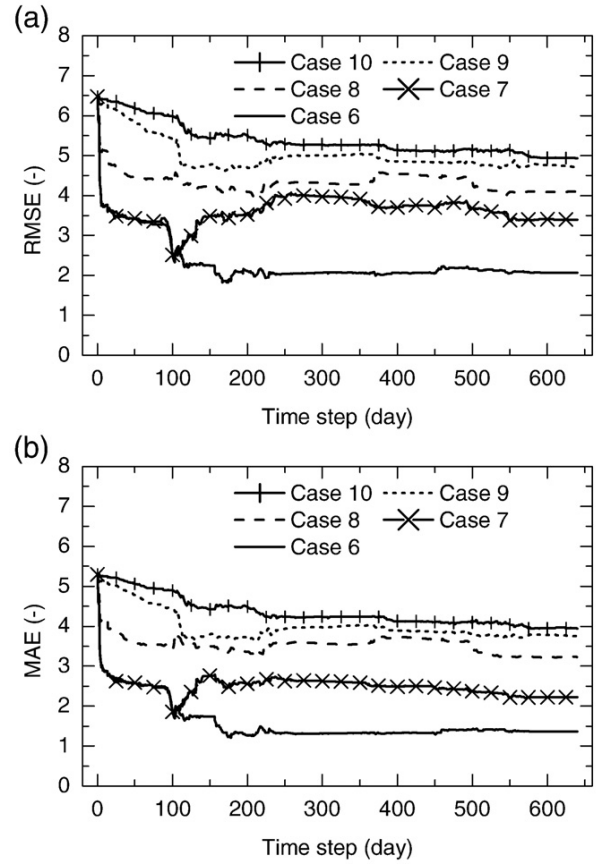


Fig. 12. RMSEs (a) and MAEs (b) of the CN_2 estimations with different error specifications (Case 6 to Case 10).

the forcing boundary (i.e. precipitation in this study), the observation and the model. For example, Case 7, whose scaling factor for runoff observation ($S_o=0.1$) is ten times the size of the model and the precipitation ($S_m=S_p=0.01$), exhibits quite discrepant estimates. Even though its RMSE series and MAE series follow the levels of Case 6 within the first 100 assimilation steps, they jump up subsequently and keep at high levels (Fig. 12). This property is not unique for Case 7, and

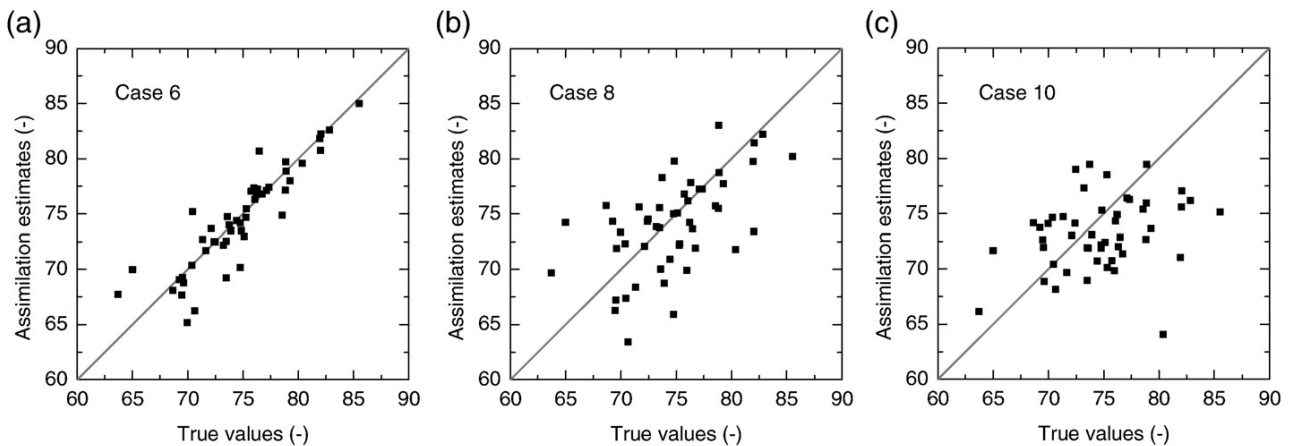


Fig. 11. Final results of the CN_2 estimates vs true (reference) values: (a) Case 6, (b) Case 8, and (c) Case 10.

Case 1 also presents oscillatory behaviors as shown in Fig. 8. In fact, more unacceptable results can appear in the cases with inconsistent combination of errors, especially the case with large error magnitudes (e.g., for $S_o = 0.05$, $S_m = 0.01$ and $S_p = 0.05$, not shown). This kind of combination may not render negative impact at the beginning steps, but at the late assimilation steps it spoils the data assimilation performance or even induces filter divergences. Remedy to this issue may be made by prescribing time-correlated errors or using adaptive filter techniques [6,29].

5.2. Initial realization of parameters

Usually according to the prior knowledge of the landuse and soil type in a catchment, the initial realizations can be generated with raw statistical properties including means and variances. If the land cover type is pasture and the soil type is in “B group” with fair hydrological condition in an HRU, for example, the mean of CN_2 could be set as 69 and the variance would be small [25]. In practice, this kind of rough prescription may be far away from the true statistics of parameters, particularly when the HRU contains an erratic type of landuse and complex soil formations. So the subjectivity and randomness are often unavoidable. The goal of this subsection is to examine the impact of the biased prior statistics on the estimation results. The mean and variance are taken into account in Case 11, Case 12 and Case 13 whose means and variances are far away from the true values (Table 2).

Fig. 13 illustrates final results of the three cases. It is clear that these assimilation estimates have captured principal properties of the CN_2 distribution for the HRUs. Case 11 with relative small initial biases provides the best result while there are no significant differences among the three. This has also been demonstrated by the evolutions of errors. As shown in Fig. 14, the RMSEs and MAEs drop dramatically at the first five assimilation steps even for Case 12 and Case 13 with large initial biases. After 250 assimilation steps, each of them keeps constant approximately and there are small differences mutually. The overestimated errors of the observation, the model and the forcing boundary (Table 2) should be partially responsible for the estimated results, which are not as good as that of Case 1. Therefore, the data assimilation performance is not greatly dependent on the initial guess of the parameters. In addition, compared to Case 13, Case 12 with larger initial biases exhibits less RMSE and MAE within the beginning 90 assimilation steps and even its final errors are not comparable to that of Case 13. So it is preferable to enlarge the magnitudes of the variances if the prior means are far away from the truths (i.e., with poor prior knowledge), especially for a short-period assimilation. Wide range distributions of the initial realizations are beneficial for

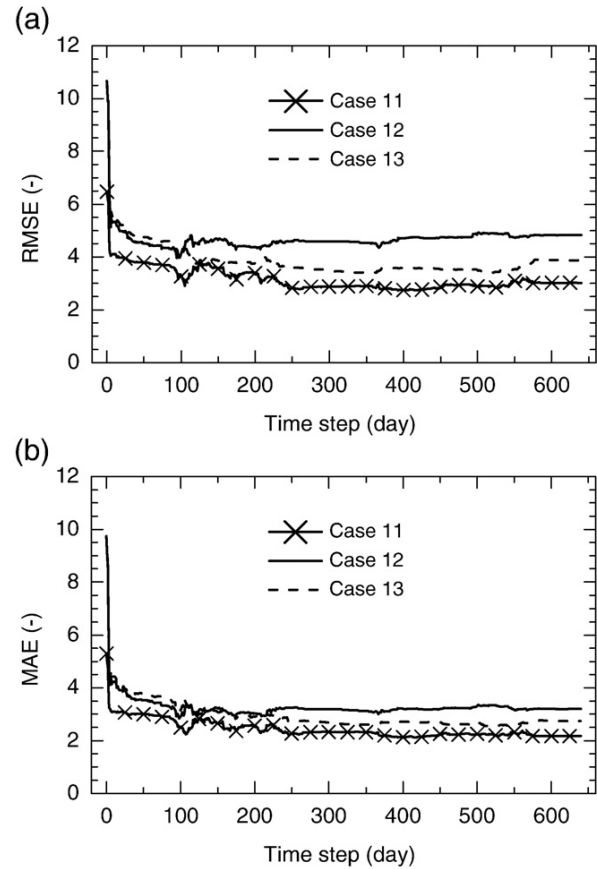


Fig. 14. RMSEs (a) and MAEs (b) of the CN_2 estimations under different initial guesses (Case 11 to Case 13).

broadly covering true values and making each realization approach truths by assimilating observations.

5.3. Ensemble size

In the EnKF, a cluster of realizations (samples) are used to approximate the probability distribution of the states. Generally, increasing the ensemble size enables the algorithm to propagate the error information more accurately but at the same time it increases

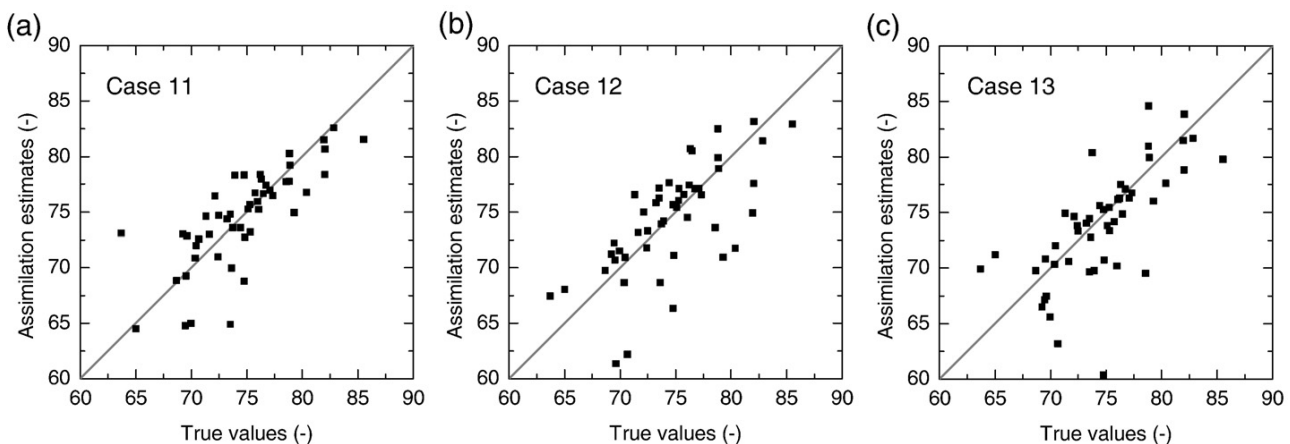


Fig. 13. Final results of the CN_2 estimates vs true (reference) values: (a) Case 11, (b) Case 12, and (c) Case 13.

the computational burden. Moreover, the assimilation estimation results also depend on specified random realizations even with the same ensemble size. Therefore, it is of great importance to explore the appropriate ensemble size to balance the estimation accuracy and the computational efficiency.

On the basis of Case 13, other three cases with ensemble size of 100, 400, 1000 are assembled with 300 assimilation time steps. For each of the four (Case 13 also included), we perform 30 sets of assimilation runs by varying the random seeds to generate different realizations. In addition to the RMSE, another measure of the goodness is the Ensemble Spread, which represents the estimated uncertainty. It should be close to the RMSE if the EnKF estimates the uncertainty of the state properly [4]. The confidence intervals for RMSE and Ensemble Spread at each assimilation step, with the 0.95 confidence intervals, are computed based on the 30 sets of assimilated results.

Table 3 displays the final results of the four cases. As expected, the spatial average RMSE decreases with the increase of ensemble size, while it exhibits slight improvement when the size is in excess of 400. For the case with a small ensemble size (e.g., 100, 200) the Ensemble Spread is smaller than the RMSE, which indicates that the realizations systematically underestimate the uncertainty of the state. This kind of underestimation can be eliminated by increasing the ensemble size. For example, in the case with size = 1000, the Ensemble Spread is close to the RMSE after 50 assimilation steps (Fig. 15). Nevertheless, this improvement comes with the price of a large computational time. With a computer of dual 3.50G HZ processor, the execution time increases from 78 s to 3977 s as the size goes from 100 to 1000, given the 300 time steps. The cost with size = 1000 is not quite expensive for our problems, but the computational burden will increase and become overwhelming if there are more than a hundred of computational units (sub-basins) in a different catchment. Certainly, the reasonable size depends on the nature of the problem. For example, the atmospheric estimation could be performed well with 64 ensemble members [22] and the land surface assimilation may require the size only about 10 [17], while the hydraulic conductivity estimation in geologic formations requires about 200 members [4].

With equal ensemble sizes, furthermore, different sets of parameter realizations will provide different estimation accuracy, especially for the small-sized cases. While the ensemble-spread confidence interval is negligible, there is a certain RMSE confidence interval for every case as shown in Table 3 and Fig. 15, and the case with size = 100 exhibits the broadest interval. This characteristic results from the basic principle of the EnKF which uses finite random samples to represent a notional probability distribution. So it is preferable to try different random realizations and compare their results when we conduct data assimilation with the EnKF.

6. Summary and conclusions

This paper has investigated the performance of the ensemble Kalman filter (EnKF) for the catchment scale hydrological data assimilation. Distributed hydrological models outperform lumped hydrological models in capturing the spatial patterns of the coupled hydrological processes in catchments [28,35], but their high dimensions of states and parameters will pose difficulties for obtaining optimal

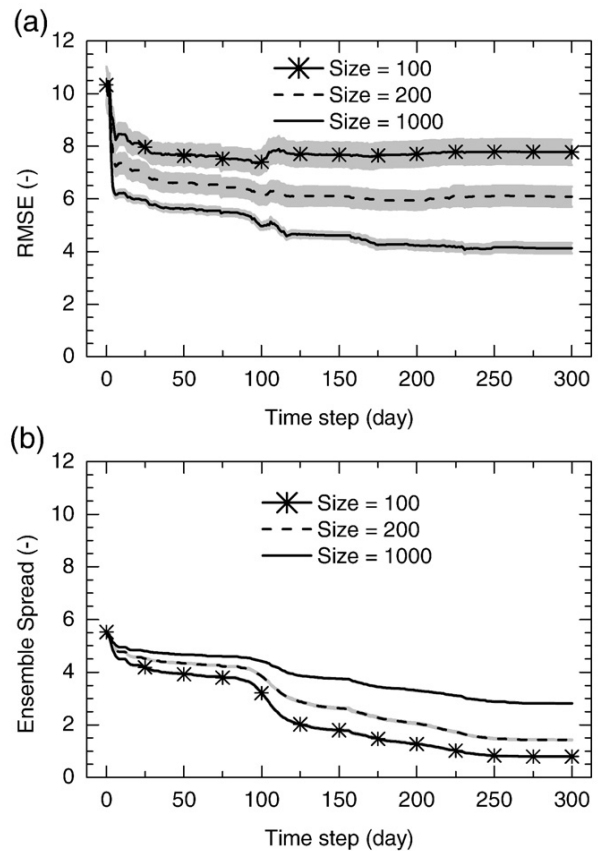


Fig. 15. RMSEs (a) and Ensemble Spreads (b) for CN_2 estimations with different ensemble sizes (the light gray areas are the 95% confidence intervals).

estimations. Our objective here has been to better understand the capability of combined state-parameter estimation with EnKF and to obtain a more complete understanding of its sensitivities. Accordingly, a data assimilation system has been constructed based on a physically based hydrological model (SWAT), and synthetic assimilation experiments have been performed to estimate a dominant parameter, CN_2 , as well as the three prognostic states, namely the runoff, soil water content, and evapotranspiration.

It has been found that by assimilating observations, such as the runoff, the EnKF can effectively update the hydrological states and progressively improve the parameter estimation. The hydrologic states (e.g., the soil water content) can be well reproduced even with biased prior knowledge of the parameter. The hydrologic parameters are more difficult to get perfect estimates owing to the nonuniqueness problem in distributed models.

Moreover, the runoff observations from different sites play different roles in the assimilation. The observations from lower streams, particularly from the outlet station, make the greatest contribution to the estimations. So the outlet runoff is a preferred observation for data assimilation if only a limited number of stations are available. In addition, taking the soil water content as complementary observations can obviously improve the inverse estimate of parameters, while the evapotranspiration observations are less effective owing to its weak correlation to the interested parameter (CN_2).

Besides the measurement sites and types, the EnKF is also sensitive to other factors about the assimilation system setup. The success of the data assimilation is quite dependent on the error specification, which should be appropriate to the model capability and the measurement flexibility. Not only is the level of error overestimation

Table 3
Estimated confidence intervals of the CN_2 s and computational requirements for different ensemble sizes.

Ensemble size	RMSE		Ensemble spread		Execution time (s)
	Mean	Confidence Interval	Mean	Confidence Interval	
100	7.775	(7.299, 8.250)	0.790	(0.774, 0.806)	78
200	6.080	(5.698, 6.461)	1.428	(1.411, 1.444)	215
400	4.614	(4.387, 4.842)	2.040	(1.938, 2.142)	699
1000	4.118	(3.917, 4.319)	2.814	(2.797, 2.830)	3977

limited (less than 10% being recommended), but also the combination of errors (about the model, the observation and the forcing boundary) should be kept consistent to avoid spoiling the assimilation performance. Furthermore, the biased initial estimates of parameters do not pose significant impacts because they can be improved sequentially by assimilating observations, but relatively large variances are preferable when short-term data assimilation is conducted. In addition, before achieving statistical convergence the assimilation results also depend on the particular set of initial realizations because the EnKF is basically a Monte Carlo method. It may be a good idea to compare results from different sets of realizations for the case with a small ensemble size. A reasonable ensemble size should be taken into account to balance estimation accuracy and computational feasibility.

Our analysis is limited by several factors that could be addressed in the future work. In this study, first, the state-parameter estimation was based on the EnKF by concatenating the uncertain parameters into a single joint state vector (state augmentation). This approach highly relies on the state-parameter dependencies. It is not able to provide good estimates if the dependencies are weak (e.g., Case 5 in this study, see Fig. 8). To address this challenge, the novel variants of the EnKF, including the smoothed EnKF [3] and other dual estimate approaches [23], may be useful remedies. Furthermore, we focused only on one dominant parameter estimate in the synthetic data assimilation system by drawing observations from reference simulation and assumed the parameter to obey a Gaussian distribution. In reality, a number of additional parameters and states may control the hydrological processes, and the Gaussian assumption may be violated. These problems will need to be examined in real-world hydrological environment instead of synthetic experiments. In addition, the model and observation errors are assumed to vary with the magnitude of the variables according to the prescribed scaling factors. This simple treatment may be augmented with the adaptive filtering approach that estimates the error information during the online cycling of a data assimilation system [6,29]. Finally, a more general analysis would examine a broad range of observations including the in-situ measurements, multi-scale remote sensing soil moisture, evapotranspiration and land surface temperature. Our ongoing work is addressing some of these issues.

Acknowledgements

We are grateful to the supports by Natural Science Foundation of China (No. 50688901), the Chinese National Basic Research Program (No. 2006CB705800), and the China Postdoctoral Science Foundation (No. 20080440271). We would like to thank Dr. Witold F. Krajewski and the anonymous reviewers for their constructive comments.

References

- [1] Aubert D, Loumagne C, Oudin L. Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall–runoff model. *J Hydrol* 2003;280:145–61.
- [2] Bras RL, Rodriguez-Iturbe I. *Random functions and hydrology*. Reading, MA, USA: Addison Wesley; 1985. 559 pp.
- [3] Chen M, Liu S, Tieszen LL, Hollinger DY. An improved state-parameter analysis of ecosystem models using data assimilation. *Ecol Model* 2008;219:317–26.
- [4] Chen Y, Zhang D. Data assimilation for transient flow in geologic formations via ensemble Kalman filter. *Adv Water Res* 2006;29:1107–22.
- [5] Clark MP, Rupp DE, Woods RA, Zheng X, Ibbitt RP, Slater AG, Schmidt J, Uddstrom MJ. Hydrological data assimilation with the ensemble Kalman filter: use of streamflow observations to update states in a distributed hydrological model. *Adv Water Res* 2008;31:1309–24.
- [6] Crow WT, Reichle RH. Comparison of adaptive filtering techniques for land surface data assimilation. *Water Resour Res* 2008;44. doi:10.1029/2008WR006883 W08423.
- [7] Crow WT, Van Loon E. Impact of incorrect model error assumptions on the sequential assimilation of remotely sensed surface soil moisture. *J Hydrometeorol* 2006;7:421–32.
- [8] Das NN, Mohanty BP, Cosh MH, Jackson TJ. Modeling and assimilation of root zone soil moisture using remote sensing observations in Walnut Gulch Watershed during SMEX04. *Remote Sens Environ* 2008;112:415–29.
- [9] Evensen G. *Data assimilation: the ensemble Kalman filter*. Springer; 2007. p. 279.
- [10] Evensen G. The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics* 2003;53:343–67.
- [11] Gassman PW, Reyes MR, Green CH, Arnold JG. The soil and water assessment tool: historical development, applications, and future research directions. *Trans ASABE* 2007;50(4):1211–50.
- [12] Helton JC, Davis FJ. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab Eng Syst Saf* 2003;81:23–69.
- [13] Immerzeel WW, Droogers P. Calibration of a distributed hydrological model based on satellite evapotranspiration. *J Hydrol* 2008;349:411–24.
- [14] Kitanidis PK, Bras RL. Real-time forecasting with a conceptual hydrologic model 1. Analysis of uncertainty. *Water Resour Res* 1980;16(6):1025–33.
- [15] Kitanidis PK, Bras RL. Real-time forecasting with a conceptual hydrologic model 2. Applications and results. *Water Resour Res* 1980;16(6):1034–44.
- [16] Komma J, Blöschl G, Reszler C. Soil moisture updating by ensemble Kalman filtering in real-time flood forecasting. *J Hydrol* 2008;357:228–42.
- [17] Kumar SV, Reichle RH, Peters-Lidard CD, Koster RD, Zhan X, Crow WT, Eylander JB, Houser PR. A land surface data assimilation framework using the land information system: description and applications. *Adv Water Res* 2008;31:1419–32.
- [18] Liu GS, Chen Y, Zhang DX. Investigation of flow and transport processes at the MADE site using ensemble Kalman filter. *Adv Water Res* 2008;31:975–86.
- [19] Luzio MD, Srinivasan R, Arnold JG, Neitsch SL. *Arview Interface for SWAT2000: user's guide*. Blackland Research Center, Texas Agricultural Experiment Station; 2002.
- [20] Madsen H. Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives. *Adv Water Res* 2003;26:205–16.
- [21] McLaughlin D. An integrated approach to hydrologic data assimilation: interpolation, smoothing, and filtering. *Adv Water Res* 2002;25:1275–86.
- [22] Mitchell HL, Houdekamer PL. Ensemble size, balance, and model-error representation in an ensemble Kalman filter. *Mon Weather Rev* 2002;130:2791–808.
- [23] Moradkhani H, Sorooshian S, Gupta HV, Houser PR. Dual state-parameter estimation of hydrological models using ensemble Kalman filter. *Adv Water Res* 2005;28:135–47.
- [24] Muleta MK, Nicklow JW. Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. *J Hydrol* 2005;306:127–45.
- [25] Neitsch SL, Arnold JG, Kiniry JR, Williams JR. *Soil and water assessment tool, theoretical documentation: Version 2000*. Temple: Blackland Research Center, Texas Agricultural Experiment Station; 2001.
- [26] Pauwels RN, De Lannoy JM. Improvement of modeled soil wetness conditions and turbulent fluxes through the assimilation of observed discharge. *J Hydrometeorol* 2006;7:458–77.
- [27] Pebesma EJ, Heuvelink GM. Latin hypercube sampling of Gaussian random fields. *Technometrics* 1999;41(4):303–12.
- [28] Reed S, Koren V, Smith M, Zhang Z, Moreda F, Seo DJ. DMIP participants. Overall distributed model intercomparison project results. *J Hydrol* 2004;298(1–4):27–60.
- [29] Reichle RH, Crow WT, Keppenne CL. An adaptive ensemble Kalman filter for soil moisture data assimilation. *Water Resour Res* 2008;44. doi:10.1029/2007WR006357 W03423.
- [30] Reichle RH, McLaughlin D, Entekhabi D. Hydrologic data assimilation with the ensemble Kalman filter. *Mon Weather Rev* 2002;130:103–14.
- [31] Reichle RH. Data assimilation methods in the Earth sciences. *Adv Water Res* 2008;31(11):1411–8.
- [32] Rouhani H, Willems P, Wyseure G, Feyen J. Parameter estimation in semi-distributed hydrological catchment modelling using a multi-criteria objective function. *Hydrol Process* 2007;21:2998–3008.
- [33] Schuurmans JM, Troch PA, Veldhuizen AA, Bastiaanssen WGM, Bierkens MFP. Assimilation of remotely sensed latent heat flux in a distributed hydrological model. *Adv Water Res* 2003;26:151–9.
- [34] Shen ZY, Hong Q, Yu H, Liu RM. Parameter uncertainty analysis of the non-point source pollution in the Daning River watershed of the Three Gorges Reservoir Region, China. *Sci Total Environ* 2008. doi:10.1016/j.scitotenv.2008.06.009.
- [35] Smith MB, Seo DJ, Koren VI, Reed SM, Zhang Z, Duan Q, Moreda F, Cong S. The distributed model intercomparison project (DMIP): motivation and experiment design. *J Hydrol* 2004;298:4–26.
- [36] Sun AY, Morris A, Mohanty S. Comparison of deterministic ensemble Kalman filters for assimilating hydrogeological data. *Adv Water Res* 2009;32:280–92.
- [37] Troch PA, Paniconi C, McLaughlin D. Catchment-scale hydrological modeling and data assimilation (Preface). *Adv Water Res* 2003;26:131–5.
- [38] Willmott CJ, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim Res* 2005;30:79–82.
- [39] Zhou YH, McLaughlin D, Entekhabi D, Crystal GH. An ensemble multiscale filter for large nonlinear data assimilation problems. *Mon Weather Rev* 2008;136:678–98.
- [40] Zhou YH, McLaughlin D, Entekhabi D. Assessing the performance of the ensemble Kalman filter for land surface data assimilation. *Mon Weather Rev* 2006;134:2128–42.

SPECIAL SESSION ON UNESCO-IHP

This part is a contribution of two or three experts from the UNESCO-IHP. The role of UNESCO-IHP is emphasized in addressing the climatic and hydrologic problems, and other world-wide affairs.

The role of United Nations Educational, Scientific and Cultural Organization–International Hydrological Programme in sustainable water resources management in East Asian countries

Ramasamy JAYAKUMAR, DUAN Xiaoli, Eunah KIM, *LIU Ke

Natural Science Sector, UNESCO Office Beijing, Beijing 100600, China

Abstract: For over 30 years, IHP (International Hydrological Programme) has been actively operating as a UNESCO's (United Nations Educational, Scientific and Cultural Organization) international scientific cooperative programme in water research, water resources management, education and capacity-building, and the only broadly-based science programme of the UN (United Nations) system in this region. By a number of initiatives and networks, the IHP has progressively carried out activities on the quantity and quality of global/regional water resources, transboundary water resources management, mitigation of water related hazards, and water education. While addressing comprehensive areas over water challenges, greater emphasis has been placed on the role of water resources management for sustainable development and with respect to the expected changes in climate and environmental conditions. WWAP (World Water Assessment Programme) and its major product WWDR (World Water Development Report) in East Asia are under the framework of IHP which supports field oriented activities on monitoring freshwater, developing case studies, enhancing national assessment capacity, and facilitating decision making processes. In light of transboundary waters in IHP, RSC (Regional Steering Committee) plays a focal role for facilitating regional cooperation in the Southeast and East Asia and Pacific States. Furthermore, ISI (International Sediment Initiative) and IFI (International Flood Initiative) have significant roles, respectively, for the management of erosion and sedimentation in line with river system or reservoir management, and for the flood management focusing on capacity building of each country in East Asia. There are other major areas of concern under UNESCO's IHP programme in East Asia, specifically in aspects including, mitigating water conflicts on transboundary aquifers through ISARM (International Shared Aquifer Resources Management), water management of arid areas through Water and Development Information for Arid Lands- A Global Network (Asian G-WADI), and sustainable management of groundwater by UNESCO Water Chair, as well as water education through the programme of Sustainable Water Integrated Management-Educational Component.

Keywords: UNESCO-IHP; East Asia; water resources

Received: 2008-04-14 **Accepted:** 2008-12-30

Author: Ramasamy Jayakumar, native of India, Ph.D, specialized in application of remote sensing, GIS, geo-statistical methods in hydrogeological modelling and irrigation water management. E-mail: r.jayakumar@unesco.org

***Corresponding author:** Liu Ke, E-mail: k.liu@unesco.org

1 Introduction

It was recognized in the World Summit on Sustainable Development, held in Johnhansberg 2002, that sciences have an indispensable role in providing solid underpinning for sound decision and policy-making process on sustainable development. The findings and recommendations stemmed from the summit, together with those from the World Sciences Forum held before, have guided United Nations Educational, Scientific and Cultural Organization (UNESCO) in further shaping its science programmes, defining missions of the programme as “*the Promoter and Broker of Sciences*”. The mission was further structured into two major programmes of:

- Science, environment and sustainable development, aimed to improve human security through a better management of the environment
- Capacity-building in science and technology for development, that seeks to enhance human and institutional capacities in science and technology to allow the widest possible participation in the knowledge society, and also to adapt science policy to societal needs

Both of the programmes are endeavored to pursue the UN Millennium Development Goals (MDGs), in particular those in relation to the eradication of poverty, gender, environmental sustainability and the development of a global partnership for peace and development.

2 The International Hydrological Programme (IHP)

IHP is UNESCO's international scientific cooperative programme in water research, water resources management, education and capacity-building, and the only broadly-based science programme of the UN system in this region. For 30 years until now (1975–2005), IHP has been committed to developing the science of hydrology to meet requests derived from social development. Its primary aim is to draw together scientists worldwide in order to establish the scientific and technological bases for the rational management of water resources with respect to water quantity and quality. UNESCO contributed to the preparation of the first United Nations Conference on Water held at Mar del Plate, Argentina, in March 1977. This conference attended by representatives of 116 governments and of many international organisations, had a paramount importance in the international recognition of water as a key factor in socio-economic development.

In January 1992, UNESCO was one of the twenty bodies and agencies of UN system which organized the International Conference on Water and the Environment (ICWE) in Dublin, Ireland. The Dublin Statement enunciates four basic principles:

1. Freshwater is a finite and vulnerable resource, essential to sustain life, development and the environment
2. Water development and management should be based on a participatory approach, involving users, planners and policy-makers at all levels
3. Women play a central part in the provision, management and safeguarding of water
4. Water has an economic value in all its competing uses and should be recognized as an economic good

The ICWE was designed as an input to the UN Conference on Environment and Devel-

opment (The Earth Summit) held in Rio de Janeiro, Brazil in June 1992. The Agenda 21 adopted by the conference mentions, *inter alia*, that “The holistic management of freshwater and the integration of sectoral water plans and programmes within the framework of national economic and social policies are of paramount importance of action in the 1990s and beyond”. The above-mentioned principles are fully in line with the basic philosophy of the IHP (Sorin Dumitrescu, 2006).

In the general framework of IHP, greater emphasis has been placed on the role of water resources management for sustainable development and with respect to the expected changes in climate and environmental conditions. Progress has also been achieved in methodologies for hydrological studies, training and education in the water sciences. One of IHP’s continuing objectives is to integrate developing countries into research and training efforts, thereby reinforcing regional aspects while maintaining global coordination.

The current phase IHP-VII (2008–2013) entitled ‘Water Dependencies: Systems under Stress and Societal Responses’ will focus on the strengthening of the existing scientific knowledge to provide new directions for science and research to develop scientific tools and responses to help mitigate and reverse these trends.

3 Overview of fresh water situation in East Asia

3.1 Global water resources scenario

In 2006, the 2nd version of the World Water Development Report (WWDR) was issued by the World Water Assessment Programme (WWAP). The report is the flagship report of WWAP, a programme under the UN Water—a cross-agency entity coordinating 24 UN agencies dealing with water issues. The report not only includes latest data and information reflecting the world water situation, but also grasps major themes of the general trend for researches and management on water issues.

As is addressed in the 2nd version of the World Water Development Report, one of the major themes identified is the changing context of the world water situation. The changes may come from both natural and social origins. In the natural part, more and more scientific data are verifying the fact that the natural environment is subject to the influences of a changing climate. This may have given rise to natural changes and even disasters such as the Indian Ocean Tsunami, and the frequent hurricane attacking the Caribbean and Pacific coasts. Moreover, changes are also brought forth from the social aspect, including population growth, immigration, and transboundary management of water resources. Some of these matters may even incur water related tensions and conflicts in areas suffering from serious water scarcity. In addition, the recent fast-paced economic development, especially in developing countries like China, has generated large amount of pollution that is threatening the safety of drinking water resources. Moreover, an efficient and strong governance mechanism needs also to be urgently set up to sustainably protect and manage precious water resources.

3.2 Water resources of East Asia

Table 1 shows that water availability varies greatly in this region. To illustrate, Mongolia enjoys the highest water resources availability rate. Although Mongolia is comparatively

rich in water resources stemming from precipitation in the high mountains, in its desert southern, western and eastern provinces, the water resources are much scarcer and are generally of poorer quality with increasing salts and diminishing water levels in groundwater tables, streams and lakes (FAO-AQUASTAT, 1999). Another country share the problem of uneven water distribution is China (FAO-AQUASTAT, 1999). Besides low water availability per capita, the water resources in China are unevenly distributed in terms of time and space. Affected by monsoons, precipitation occurs mostly in the summer months providing 60-80% of the total annual precipitation, which is advantageous for agriculture. This pattern of precipitation tends to result in droughts in spring and floods and waterlogging in the summer. Furthermore, quality of fresh water resources in this region varies greatly. Japan and R.O. Korea yield much higher water quality than the rest.

Water availability varies in different seasons. Most of the East Asian region is subject to the monsoon climate, the warm and wet winds in spring and summer times bring most of the annual precipitation. However, in the winter seasons, the climate tends to be cold and dry. This situation often makes this region vulnerable to natural disasters such as hurricane and storm in summer and drought in winter. For instance, the Chinese capital Beijing has been experiencing continuous drought from 1999 to 2006 in winters. The per capita water resources per year are less than 300 cubic meters, which is only 1/30 of the world's average. To cope with the situation, groundwater has become one of the major alternative resources for cities like Beijing. Unfortunately, the groundwater consumption often follows an excessive manner, leaving a series of problems such as land subsidence and sea water intrusion. The problems can be particularly worsened when climate change brought higher temperatures, reducing amount of precipitation but increasing that of evaporation.

Table 1 Water availability of East Asian countries

Country	Population	Precipitation Rate ¹	TARWR volume 2005	TARWR ² per capita 2000	TARWR per capita 2005	Surface water ratio to TARWR	Ground-water ratio to TARWR	Overlap ³ ratio	Incoming water ratio to TARWR	Outgoing ⁴ water ratio to TARWR	Total use ratio to TARWR
	10 ³	mm/yr	km ³ /yr	m ³ /yr	m ³ /yr	%	%	%	%	%	%
China	1,320,892	600	2,830	2,259	2,140	96	29	26	1	25	–
China, Taiwan Prov.	22,894	2,400	67	–	2,930	94	6	0	0	–	–
DPR Korea	22,776	1,400	77	3,464	3,390	86	17	16	13	6	12
Japan	127,800	1,700	430	3,383	3,360	98	6	4	0	0	21
Mongolia	2,630	200	35	13,739	13,230	94	18	11	0	76	1
RO Korea	47,951	1,100	70	1,491	1,450	89	19	15	7	–	27

Source: FAO-AQUASTAT, 2005, the table was extracted from World Water Development Report 2 (¹Average Precipitation (1961-90 from IPCC (Intergovernmental Panel on Climate Change) (mm/year). As in the FAO-AQUASTAT Database, for some countries large discrepancies exist between national and IPCC data on rainfall average. In these cases, IPCC data were modified to ensure consistency with water resources data. ²TARWR stands for Total Actual Renewable Water Resources. ³Overlap is the water that is shared by both the surface water and groundwater systems. ⁴Outflow - Sep. 2004 for surface water and Aug. 2005 for groundwater)

3.3 Regional cooperation on transboundary waters

Transboundary waters include any surface or groundwater that mark, cross or located on the

boundaries between two or more States. In other words, wherever transboundary waters flow directly into the sea, these transboundary waters end at a straight line across their respective mouths between points on the low-water line of the banks (Helsinki Convention 1992). By the year 2006, 263 transboundary waters were identified in the world and the number has increased by emergence of the newly independent states after the breakup of the former Soviet Union. Europe alone has 100 transboundary groundwater aquifers and more are expected to be identified in the future (United Nations Development Programme, 2006). About 60% of the global population depends on these transboundary waters which often preserve natural ecosystems. However, due to the increasing pressure of economic development and competition for scarce resources, many international water basins have to suffer serious environmental, social and political problems. The United Nations Conference on Environment and Development (UNCED) in Rio de Janeiro in 1992 adopted Agenda 21, recognizing the multi-sectoral nature of water resources development as well as the diverse interests in their utilization (United Nations University, 1990). While there have been much efforts among European, South-American and Middle-East countries for cooperation on transboundary water between neighbouring countries, East Asia has relatively low number of transboundary waters and thus less cooperative work regarding sharing waters.

China shares 12 main rivers with six neighbouring countries, including Mongolia, Pakistan, India, Kazakhstan, the Kyrgyz Republic, Myanmar, Lao People's Democratic Republic and Viet Nam. The mean annual volume of water entering the country is 17.2 km³, of which 4.2 percent in the Heilongjiang basin, 52.9 percent in inland rivers, 0.7 percent in rivers in the southwest, and 42.2 percent in the Pearl River basin. Table 2 presents data on rivers entering into and flowing out of China.

In light of international cooperation regarding water sharing problems UNESCO's PCCP (Potential Conflict to Cooperation Potential) facilitates multi-level and interdisciplinary dialogues to foster peace, cooperation and development. It uses methodologies of joint research and training activities to find ways to enhance the concerned parties' knowledge of the shared water resources as well as to achieve progress on the cooperation and development.

4 IHP in East Asia

4.1 Regional Steering Committee (RSC)

Regional Steering Committee (RSC) is a regional network of IHP in Southeast Asia. Its members consist of representatives from 14 IHP National Committees from this region. This has made it easier for regional cooperation to carry out IHP initiatives. Since 1993, annual RSC meetings have been convened in different countries of the region to report, evaluate and review a wide range of activities conducted in the framework of IHP as well as to design new ones.

With strong support from the Japanese Government through providing a Funds-in-Trust contribution on a regular basis, in co-operation with UNESCO Jakarta and the participating member states, the RSC has coordinated a wide spectrum of initiatives over the ten years covering research studies, technical projects, workshops, training courses and annual symposia, bringing together many specialists in the field of water activities. Among initiatives

generated by RSC, the most notable ones include the AP-FRIEND (Asian Pacific Flow Regimes from International and Experimental Network Data) project, the Asian Pacific Water Archive and the Catalogue of Rivers.

4.2 International Sediment Initiative (ISI)

The management of erosion and sedimentation has been an important part for catchments, river system, and reservoir management. It has been recognized that, without appropriate treatment methods, more than 50% of the world's reservoir capacity will be lost in the next decades due to erosion and sedimentation. Nevertheless, relevant investigations on sedimentation and erosion need to be further promoted in a world scale. In this general backdrop, the International Sedimentation Initiative (ISI) was launched, as a major activity, in the sixth phase of UNESCO-IHP (2002-2006).

The initiative is expected to improve awareness on the sedimentation and erosion issues. It is also aimed to provide advices and enhance the making and implementation of policies that would contribute to the sustainable management of erosion and sedimentation. To realize the aims, a series of activities have been carried out. For instance, under the general ISI framework, a global repository was constructed with inputs of data and information, collected by major international research institutes, derived from documents on erosion and sedimentation.

One of such reputable institutes in East Asia is the International Research and Training Center for Erosion and Sedimentation (IRTCES) located in Beijing. Through its competent investigations, IRTCES managed to provide the latest data and information on the situation of the erosion and sedimentation in major river basins such as the Yangtze River and the Yellow River. In addition, it also organizes trainings and conferences to improve awareness and capacity to sustainably manage and monitor the erosion and sedimentation in China.

4.3 International Flood Initiative (IFI)

As one of the greatest water-related natural disasters, flooding can cause devastating damage that affect millions of people's livelihood, claiming thousands of lives each year. However, on the other hand, the floods are naturally occurred phenomena providing elements not only to the sustainability of particular ecosystems, but also to the development of many human activities.

Unfortunately, water-related disaster is under constant rise as a result of urbanization, climate change and global warming. Since 1992, the yearly number of water-related disasters has risen from slightly over 50 to more than 150. The disasters have claimed about 25,000 lives and affected over 500 million annually, costing the world economy more than \$60 billion. Under the circumstances, during the World Conference on Disaster Reduction (WCDR), International Flood Initiative (IFI) was launched in 2005. IFI aims to reduce human and socio-economic losses from flooding and use of flood plains while increasing social, economic and ecological benefits.

In order to achieve its goals, IFI has engaged many carefully designed activities related to integrated flood management covering research studies, training, information networking, empowering communities with good governance and technical assistance. Through these activities, IFI is expected to develop the capacity of each country to better understand and

handle flood involved hazards, vulnerabilities and benefits.

IFI is based in the International Centre for Water Hazard and Risk Management (ICHARM) hosted by the Public Works Research Institute in Tsukuba, Japan. ICHARM was endorsed as the global facility and Secretariat responsible for the IFI at UNESCO's 33rd session of its General Conference.

4.4 G-WADI Asia

Approximately 80 countries, constituting 40% of the world's population, were suffering from serious water shortages by the mid-1990s (Asian G-WADI-UNESCO, 2007). In less than 25 years, two-thirds of the world's population will be living in water-stressed countries. Globally, arid and semi-arid areas face the greatest pressures to deliver and manage fresh-water resources. Furthermore, these areas are under growing pressure of water management derived from issues including, population growth, agricultural expansion, salinity increases, and agricultural/urban pollution. These problems have exerted further difficulties to realize goals of water resource availability, equity in water management, and strategies to support peace and security.

As a result of growing attention and challenges on water management in arid and semi-arid areas, the 15th session of the Intergovernmental Council of the International Hydrological Programme (IHP) decided to establish a Global Network on Water Resources Management in Arid and Semi-arid Zones, in December 2002. Following this decision, the Global Network on "Water and Development Information for Arid Lands (G-WADI)" was established in the initiating meeting held in Paris on 14 and 15 April 2003. G-WADI's strategic objective is to strengthen the global capacity to manage water resources in arid and semi-arid areas. Its primary aim is to create an effective global community to promote international and regional cooperation in arid and semi-arid areas.

In light of the goal of G-WADI, in March 2005, in Roorkee, India Asian G-WADI was established to create a network promoting international and regional cooperation of water management in arid and semi-arid areas with specific emphasis on Asia. This network covers eleven countries including Afghanistan, China, India, Iran, Kazakhstan, Kyrgyzstan, Mongolia, Pakistan, Tajikistan, Turkmenistan, and Uzbekistan. Asian G-WADI serves as a platform for information and resource sharing. Focusing on this aim, Asian G-WADI has formulated guidelines for proposing one or more basins as G-WADI pilot basins in a country, which can contribute to the knowledge base for the region. Asian G-WADI is working effectively in network building of data, groundwater modelling and research. Asian G-WADI has also identified specific research needs for arid and semi-arid areas of Asia that would be highlighted as areas of special emphasis and would provide assistance to conduct studies and research.

4.5 Groundwater for emergency situation

In recent years, the availability of and access to freshwater have been highlighted as among the most critical resource challenges in the world. With the blooming economic development and the rapid urban population increase in China, water shortage has become a widespread issue. More than 400 cities (about two-thirds of the cities in China) are being puzzled by seasonal or perennial water shortage; among them more than 100 cities are facing serious

water crisis.

As part of the global initiative, UNESCO Office Beijing initiated a pilot case study on "The Emergency Water Supply in Beijing for Coping with the Consecutive Drought." The Project is being implemented by the China Institute for Geo-Environmental Monitoring (CIGEM), China Geological Survey and Geo-Environmental Monitoring Station (GEMS), Beijing (UNESCO Office Beijing, 2007).

Beijing is regularly influenced by the continental climate with semi-humid monsoon, and faces serious shortage of water resources that is far below the UN standard water resource availability per capita and also below the lowest standard of water resource availability per capita to ensure the modern and easy social life and production presented by UNESCO.

In the beginning of 2007, UNESCO paid much attention to the serious water shortage in Beijing, trying to find out solutions for the shortage of water supply in the consecutive years of drought in Beijing, and to share the experiences of emergency water supply for other regions and countries with similar situation. To complete the mission successfully, CIGEM and Beijing Geological Environment Monitoring Station (GEMS) were selected to collaboratively conduct data collection, and general research incorporated five field investigations. A completed report on the Use of Groundwater for Emergency Cases in Beijing has been prepared.

4.6 Rainwater harvesting of Shenzhen urban area

As one of the biggest economic centres of China, Shenzhen is facing great challenges of extreme water shortage although enjoying fast economic growth. It ranks as one of the top seven cities in China facing most severe water shortage crisis. Shenzhen's geographic character, in line with its large amount of water consumption, and other human activities have jointly worsened its water shortage problem. Water resources in Shenzhen city heavily rely on rivers and groundwater resources. However, rainfall, as another source of important water supply, has not been sufficiently used yet. For instance, only one third of its total rain fall volume has been properly collected and stored for usage. Additionally, most of the rain harvesting venue concentrates on various reservoirs. Despite the fact of water shortage and problems on rain harvesting Shenzhen is facing, it is still one of the cities advocating rain harvesting also with successful experiences. The case of Shenzhen and downside of rain harvest activities can serve as example for other cities with similar situation. Therefore, conducting a case study of rain harvesting in Shenzhen city would greatly contribute to rain harvesting endeavours elsewhere in China.

In view of that, UNESCO Office Beijing sponsored a pilot case study on "Urban Rain Water Harvesting—Case Study in the Shenzhen City" carried out by College of Water Sciences, Beijing Normal University. This case study was conducted in 2007 and a report on Rain Harvest in Shenzhen City has been completed and submitted to UNESCO Office Beijing.

The report elaborates experiences of the Qiaoxiang district, a pilot district of rain harvesting in Shenzhen. Three major rain harvesting methods, including roof rainwater collection, road rainwater collection and green space rainwater collection were identified and investigated in detail. Data analysis of the report showed that annual mean rainfall volume in Shenzhen is above 1837 mm, but the value of water resources availability per capita is less than 200 m³. Additionally, Shenzhen City Reservoir Runoff accounts in the proportion of the

total runoff in the city of Shenzhen were 30.5%, 31.8% and 33%. In other words, only one-third of the Shenzhen annual rainwater resources are properly stored for usage. Therefore, Shenzhen has great potential in rain harvesting. Another contribution of this research paper is that it explored advantages, disadvantages and respective harvesting methods in three different areas including mountain areas, urban areas and drainage areas, where rain harvesting is being carried out. Lessons learned from this case study could well provide reference and benchmark to be shared for rain harvest endeavors in China.

4.7 Transboundary aquifer

Approximately 40% of the world's population lives in river basins and aquifer systems that cross the political boundaries of two or more countries. The aquifers, apart from its environmental function as vital natural resources, are also endowed with crucial social function, constituting solution for preventing conflicts over water resources. Transboundary Aquifers, as part of the global groundwater resource system, are important for building a society where civilizations coexist harmoniously and accommodate each other.

With the goal of promoting sound use of transboundary aquifers, ISARM (Internationally Shared Aquifer Resource Management) has led to several regional initiatives, including ISARM-America, ISARM-Europe, ISARM-Balkans and ISARM-Asia. As one of the outputs of ISARM-Asia programme, the report on "Transboundary Aquifers in Asia with Special Emphasis to China" represents advanced research on transboundary aquifers in the Asia-Pacific region. Through this report, UNESCO, as a major coordinator of ISARM, promoted awareness on transboundary aquifers as vital natural resources, especially among the policymakers, and also to enhance collaboration of countries with shared aquifers on management of such resources in a sustainable manner.

Based on previous experiences in other regions, the report of "Transboundary Aquifers in Asia with Special Emphasis to China", comparatively, is in need of further promotion. The authors supervise and develop regional preliminary data on selected shared aquifers in Asia and conduct a pilot case study on the Heilongjiang-Amur River flowing through China and Russian Federation under ISARM-Asia framework. This report covers topics including groundwater resources and transboundary aquifers in Asia, groundwater resources and transboundary aquifers in China, pilot study of shared aquifer between provinces within China, and a case study of shared aquifers between China and the Russian Federation - the aquifers underlying the Heilongjiang-Amur River basin.

In this report, UNESCO-IHP programme identified twelve significant transboundary aquifers in Asia, demarcated as Table 2, that are with significant socio-economic and ecological implications for the sharing nations. Also, the aquifers are often factors to maintain regional peace and cooperation. Based on data collected on groundwater resources contained in such aquifers, the groundwater systems in Central, East, and South Asia are analyzed. Particularly, aquifers on the national boundary of China are evaluated in details. Moreover, the research on the middle basin of Heilongjiang-Amur River that borders China and Russia has been taken as the first case study on transboundary aquifers in East Asia. The information presented through such case study is now available for sharing for other Asian countries that possess transboundary aquifers. The report also provided primary data and information for the joint management of transboundary water resources. For instance, joint man-

agement on the Heilongjiang-Amur River basin is investigated in the case study. This could contribute to the ISRAM-Asia, as one of the first initiatives, from the management perspective. The report could also serve to promote as well the ISRAM-Asia Network headed by China.

Table 2 Transboundary aquifers in Asia (Han Zaisheng *et al.*, 2006)

No	Name of transboundary aquifer system	Countries sharing this aquifer system	Type of aquifer system	Extension (km ²)
1	Ertix River Plain	Russia, Kazakhstan	1	120000
2	West Altai	Russia, Kazakhstan	1,2	40000
3	Ili River plain	China, Kazakhstan	1	53000
4	Yenisei upstream	Russia, Mongolia	1,2	60000
5	Heilongjiang River plain	China, Russia	1	100000
6	Central Asia	Kazakhstan, Kyrgyzstan, Uzbekistan, Tajikistan, Turkmenistan, Afghanistan	1,2	660000
7	India River plain	India, Pakistan	1	560000
8	Southern of Himalayas	Nepal, India	1	65000
9	Ganges River plain	Bangladesh, India	1	300000
10	South Burma	Burma, Thailand	2	53000
11	Mekong River plain	Thailand, Laos, Cambodia, Vietnam	1	220000
12	New Guinea Island	Indonesia, Papua New Guinea	2	870000

Type of aquifer system: 1 porous, 2 fissured/fractured, 3 karst

4.8 Water education

Designated by the UN general assembly as the leading agency for the Decade of Education for Sustainable Development (DESD), UNESCO as the UN technical Agency takes upon its leading role in water education. It has become imperative to teach and guide young children for a sustainable future through water education. Effective water education can greatly contribute to improving awareness and imparting applicable knowledge and techniques for sustainable water management to our next generation. In collaboration with the Chinese Ministry of Water Resources and the Italian Ministry for the Environment and the Territory, it launched Sustainable Water Integrated Management – Water Education Programme (SWIM-EDU). This programme aims to educate Chinese students from primary to the high school on water resources management in their daily lives. SWIM-EDU aims to improve awareness among general public on the preciousness of water resources, and ultimately advocate concrete actions to protect, cherish and conserve water.

SWIM-EDU has successfully carried out a poster competition among primary school students with theme of water protection in cooperation with CAST (China Association for Science and Technology). Winners of the competition were awarded certificates jointly issued by UNESCO and IMELS (Italian Ministry of Environment, Land and Sea). Additionally, the winning posters have been made into calendars. Through the promoted engagement of students in this activity, awareness and applied knowledge on daily water resources management was improved by an effective manner. More importantly, SIWM-EDU is also co-

operating with CAST and Ministry of Education (MOE), P.R. China in compelling a series of text books of water education, and a series of teacher training would also be initiated accordingly. Furthermore, a website of SWIM-EDU was established as one of the main outputs of this project.

Through providing a platform for hydrologists, education experts, national government and international organizations to work together, UNESCO is working effectively with both schools and policy makers in advancing water education in China.

4.9 Impact of climate change in water resources

Climate change has been a serious and urgent issue for the last few decades, calling upon related national and international sectors' attention to mitigate or adapt to the impact of global climate change. Its impacts are detrimental especially in developing countries where infrastructures are not ready to the changes, and consequently challenging to the achievement of the MDGs, specifically in regard to the eradication of poverty, combating diseases as well as ensuring environmental sustainability. According to China's First National Assessment of Global Climate Change released in June 2007, climate change will affect on China's production of wheat, corn and rice to drop by up to 37% over the next 50 years. In addition, increased intensity and frequency of flooding and droughts as well as deterioration of public health were expected (Lin Erda *et al.*, 2007).

UN Theme Group on Energy and Environment, in collaboration with the government counterparts in China has formulated a joint programme namely China Climate Change Partnership Framework to be carried out for three years starting from 2008. It is applicable to the category 'Environment and Climate Change' among the agenda of MDG fund focusing on an effort to achieve the MDGs. The principal outcomes of the joint programme are: 1) mainstreaming climate change mitigation and adaptation into national and sub-national policies, planning, and investment frameworks; 2) establishment of innovative partnerships and dissemination of technologies to mitigate climate change and increase local access to sustainable energy; and 3) accelerated action by China in assessing vulnerability to climate change and developing adaptation plans and mechanisms. More specified outputs under those main outcomes are assigned to the nine participating UN agencies in correspondence towards their specialized expertise so that the programme takes accumulative benefits of the agencies' strengths and lessons learned from past experiences.

Water resources are one of the vulnerable areas that have shown symptoms of worsening water scarcity, deteriorating water quality and increasing risks of water-related disasters. UNESCO takes the leading role on water issues enclosed under framework of the joint programme with cooperation with other relevant UN agencies and government partners. Water sector in the joint programme will focus on vulnerability assessment, development of adaptation measure, increasing monitoring capacity on climate change impact, as well as the definition and enact of remedial actions. Based on the level of vulnerability to climate change in China, Yellow River Basin was identified as the focus area for this project, where is needed to develop a water management scheme covering aspects of vulnerability assessment, monitoring, and the development of adaptation and remedial measures. UNESCO will implement the project at Yellow River Basin to carry out analysis on present status, need assessment, and a thematic report on the challenges would also be prepared which will be

included in the 3rd World Water Development Report (WWDR). In addition, a training workshop would also be organized under this project.

4.10 World Water Assessment Programme

The World Water Assessment Programme (WWAP) is one of the core programmes in UN-Water. Housed in UNESCO, it supports activities on monitoring freshwater, developing case studies, enhancing national assessment capacity, and facilitating decision making processes. The major product of the programme is a periodic publication of the WWDR comprising comprehensive reviews on the state of the world's freshwater resources.

WWDR includes a thematic component, a methodological component involving analysis and the production of indicators of water-related stress, and a case study component. As depicted in the Diagram 1, the key objectives of the WWDR are building national capacity on self-assessment of water resources and facilitating the application of the accumulated knowledges given, mutually exchanging progressive feedback for the next periodic report. In this regard, case studies from all around the world contribute to knowledge sets including indigenous experiences and cross-sectoral methodologies in water resources management from various natural and socio-economic environments.

4.11 Case study of Tokyo River Basin, Japan (WWDR-1, 2003; WWDR-2, 2006)

The first WWDR included the pilot case study 'Greater Tokyo Region' providing interesting geographical features of combining small river basins and a mega city scale which covered 5 river basins, namely Tone River, Ara River, Tama River, Sagami River, and Tsurumi River. Being a rich and industrialized country, its region has been known as best practices in water resource management and environmental protection by managing risks with well-established infrastructures. However, it was also reported to have water challenges such as deterioration of water quality, flood and drought.

Comprehensive indicators were presented as a major outcome of the region's successful water management to monitor water qualities of the five river basins with regard to the challenging areas, which had also adapted public opinion through internet based survey. Those indicators based on the clear criteria (relevance, cost, comprehensibility, clarity, continuity and social benefit) are important for future assessment though some of them are vague to measure exact values.

In WWDR-2, Japan provided an overview of water issues that challenge the country's water resource management and extended its concern about water related disasters and sustainability of water resources. The overview not only presents basic information on water supplies to the public and rates of population with access to sanitation, it also briefed on the environmental laws in fields of agriculture, industry, and energy regulating sustainable use of water resources, increasing public access to sanitation facilities, combating environmental degradation.

4.12 Case study of Tuul River Basin, Mongolia (WWDR-2, 2006)

Mongolia is a landlocked country covering about 1.5 million km². Most of the nation's territory is high plateau with an average altitude of 1,580 m above sea level. Due to such geographical features, Mongolia suffers from frequent water stress. Only 6% of the annual pre-

precipitation is available as water resources in form of surface water with other left seeps into the soil. Not only quantity, the quality of water resources in Mongolia is also a big challenge. Based on statistics enclosed in the report, it shows that 40% of the Mongolian population lacks access to safe water and only 25% has sustainable access to sanitation facilities. The major cause of such deficiency was recognized as the national poverty. As a result, water-related diseases are common, and the poor are often left out of due health services.

In the second WWDR, a brief summary on a pilot study of the Tuul River Basin, Mongolia was presented. It is a significant area of the country where the capital city Ulaanbaatar is located. This region, despite its small coverage within national territory (only 3.19%), is however accommodating more than half of the whole national population. Due to a large number of the poor (25% of the urban poor) living in Ulaanbaatar, the proportion of those with adequate access to safe water resources is low. Furthermore, rapid urbanization and growing mining industry significantly contaminated water resources of the Basin, also disrupted local ecosystems.

Even though the Mongolian government well recognizes importance of water resources and has made numerous efforts to protect ecosystem, for instance carrying out the National Water Programme, it has been difficult to actually implement such efforts due to the deficiency of resources, and hardship to coordinate various ministries and local and central governments. Consequently, water related policies developed at the national level often do not reach the local level. The report thus presents challenges towards water resources in Mongolia, including low level of per capita water consumption with standard far below sanitary requirement, deforestation, poor hydropower potential, lack of early warning systems, lack of public awareness, and poor management of water-related programme policies.

4.13 Case study of river basins in PR China and RO Korea

To contribute to the 3rd WWDR that will be published in 2009 during the fifth World Water Forum in Istanbul, Turkey, PR China and RO Korea have been carrying out case studies on Yellow River Basin and Han River Basin respectively. It is the first time for both countries to jointly show results of case study in the report. As the major theme of the report would focus on "Climate Change and Water," the case studies will focus on the trend of hydrological cycle and impact on land use in response to global climate change as well as the state of freshwater resources. Based on capacity assessment of the two countries in mitigating and adapting towards climate challenges, proper measurements would be projected corresponding to respective natural and socio-economic contexts of the two nations. In China, the case studies will be comprehensively further developed under the framework of the Spanish MDG Funds project for three years and presented again as a full paper in the 4th WWDR.

5 Conclusions

As one of the main four pillars supporting UNESCO's freshwater initiative, IHP has been actively operating for over 30 years. IHP is the only science and education programme devoted to freshwater issues in the UN system. Its purpose is to enhance scientific, technical, institutional and human capacities in Member States to achieve internationally agreed water-related goals, particularly those enclosed under the MDGs.

In the Asia-Pacific region, IHP activities are mainly implemented through UNESCO's Field Offices and the IHP National Committees in respective countries. This IHP regional network addresses a wide range of critical water challenges covering from availability and sanitation of water resources, transboundary water resources management, mitigation of water related hazard, to water education and impacts of climate change upon water resources.

Not a single country is able to solve the above-mentioned challenges and problems alone, therefore, joint regional researches are vital. In close cooperation with hydrologists, local governments and water related institutes, IHP serves as a platform for the sharing and exchange of information, knowledge, and technology on regional fresh water protection. It has been continuously providing support for decision makers in East Asia in developing adaptation strategies to cope with the impact of global changes, in socio-economic and ecological terms, on the precious freshwater resources.

References

- UNESCO. Asian G-WADI Information Brochure. <http://www.asian-gwadi.org/doc/Asian%20G-WADI%20brochure%20-final.PDF>, document accessed on January, 2008.
- FAO-AQUASTAT. Country Profile-China. <http://www.fao.org/nr/water/aquastat/countries/china/index.stm>, document accessed on January, 2008.
- FAO-AQUASTAT. Country Profile-Mongolia. <http://www.fao.org/nr/water/aquastat/countries/mongolia/index.stm>, document accessed on January, 2008.
- Han Zaisheng *et al.*, 2006. Transboundary Aquifers in Asia with Special Emphasis to China. <http://unesdoc.unesco.org/images/0014/001483/148390E.pdf>, document accessed on January, 2008. 43.
- Lin Erda *et al.*, 2007. Advances in Climate Change Research, China's National Assessment Report on Climate Change (II): Climate Change Impacts and Adaptation, 6–11.
- Sorin Dumitrescu, 2006. The essence of life: UNESCO initiatives in the water sciences. In: Sixty Years of Science at UNESCO. UNESCO Publishing, 233–259.
- UN University, 1990. International Rivers and Lake Basins Management. <http://www.unu.edu/env/water/trans-boundary-water.html>, document accessed on January, 2008.
- UNDP, 2006. Human Development Report 2006. Beyond Scarcity: Power, Poverty and the Global Water Crisis. Palgrave Macmillan Publishing, 204–231.
- UNECE, 1992. Convention on the protection and use of transboundary watercourses and international lakes. <http://www.unece.org/env/water/text/text.htm>, document accessed on January 2008.
- United Nations, 2003. Water for People, Water for Life, the United Nations World Water Development Report 1, <http://www.unesco.org/water/wwap/wwdr/wwdr1/> document accessed on January, 2008.
- United Nations, 2006. Water: A shared responsibility, the United Nations World Water Development Report 2, <http://www.unesco.org/water/wwap/wwdr/wwdr2/> document accessed on January, 2008.
- UNESCO Office Beijing, 2007. Ground Water for Emergency Situation (GWES), <http://www.unescobeijing.org/projects/view.do?channelId=004002002001002004001>, document accessed on January, 2008.

The State Program for High-Tech Research and Development (863 Program) Key

Project of Earth Observation and Navigation Field

——Generation & Application of Global Products of Essential Land Variables

1. Introduction

In current research on global changes and R&D of earth system models, the potential of satellite observation has not been fully tapped. So far, as for the research on global land surface changes and R&D of land surface models, there is still lack of global land products of long-time series, high temporal and spatial resolution and high quality in the international land remote sensing field. Earth system model R&D and research on global changes in China are facing great development opportunities, and it is rather necessary to focus on the establishment of production system of global land products to support the research on global changes and R&D of new generation earth system models.

This key project implementation period is from July 2009 - July 2012 with total funding 35.39 million RMB. It will make comprehensive use of domestic and overseas remote sensing data sources, develop new inversion algorithm and production system for the following five land products: Leaf Area Index, Albedo, Emissivity, Downwelling Shortwave Radiation and Photosynthetically Active Radiation. It will further improve the parameterization scheme and assimilation technology of key land surface processes, advance its simulation capability and develop demonstrative application research of global land surface changes to provide data basis and technical support for China to make decisions in dealing with global changes.

2. Project Topics

No.	Sub-project Topics	Responsible Person	Supporting Unit
01	Integration of global multi-source land observation data	Wenping Yuan	Beijing Normal University
02	Inversion algorithm development	Qing Xiao	Institute of Remote Sensing Application Chinese Academy of Sciences
03	Production System Development and Product Generation	Shunlin Liang	Beijing Normal University
04	Land Surface Simulation and Data Assimilation System	Mingguo Ma	Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences
05	Demonstrative Research on Global Land Surface Change of Land Surface Products	Xiaodong Zeng	Institute of Atmospheric Physics Chinese Academy of Sciences

2.1. Integration of Global Multi-Source Land Observation Data

The project will comprehensively integrate domestic and overseas multiple satellite remote sensing data, land surface parameter products and other auxiliary data, and establish global land observation database through geometrical correction and radiometric calibration. It'll develop database management system, build highly efficient storage, retrieval and scheduling strategy of massive data, and realize visualization and batch distribution service of data products.

2.2. Development of Advanced Inversion Algorithms

For these five products (Leaf Area Index, Albedo, Emissivity, Downwelling Shortwave Radiation and Photosynthetically Active Radiation), the project will evaluate the best algorithm for single satellite data and integrate various advanced algorithms. It'll develop the inversion algorithms for domestic satellite data, study the generation method of land surface parameter products based on multi-source satellite data with the support of background field, and conduct validation and uncertainty analysis.

2.3. Production System Development and Product Generation

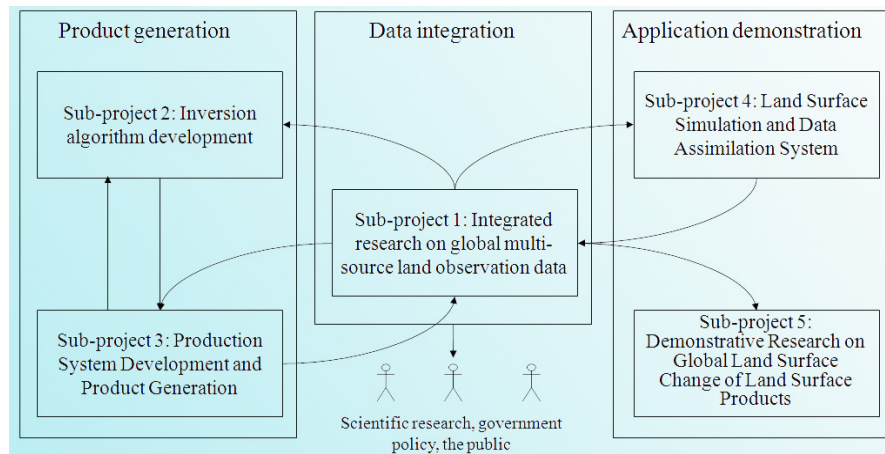
Based on various global remote sensing data, the project will develop a full automatic production system of land products. It is able to process multiple satellite data to meet the demand of global change and land surface model for data products. It'll produce global Leaf Area Index, Albedo and Emissivity products of 1985-2010, as well as global Downwelling Shortwave Radiation and Photosynthetically Active Radiation products of 2008-2010. Finally, it'll adopt the methods of automatic checking, visual inspection and field observation data validation to check the accuracy of quality and data products.

2.4. Improvement of Land Surface Simulation and Data Assimilation System

Based on global land surface parameter products produced by the project and land surface process model CoLM, the project will improve the canopy radiation transfer model and coupled dynamic vegetation model, and establish a better CoLM based land surface simulation system. It'll develop advanced land data assimilation method and high resolution forcing data sets, and establish land surface data assimilation system which can assimilate multi-source remote sensing data products produced by the project. The land surface data assimilation system will be demonstrated over China with the physics compliance and space-time compliance between land surface system simulation and forecast.

2.5. Demonstrative Research on Global Land Surface Change of Land Surface Products

Based on global land surface parameter sets, the project will assess their change trend, spatial and temporal characteristics, and leaf phenology to reveal the fact of global land surface energy, water and carbon cycles. It'll study the capability of land products in improving the simulation performance of the climate and Earth system models and distinguish the contribution and attribution of natural variations and human activities to global climate change.



3. Expected Outcomes

3.1. Integrated Global Multi-Source Land Observation Data

- (1) Collection of 11 kinds of moderate and low resolution satellite remote sensing data globally.
- (2) Pre-processing of Remote Sensing Data: Geometric correction and image registration precision is better than 2 pixels.
- (3) Data Management and Storage: It will have the capability of 200TB online massive remote sensing database management, and scalable and distributed storage.
- (4) Database Service Performance: It will provide quick database access interface and data access interface, and meet the online inquiries of over 100 users simultaneously.

3.2. Inversion algorithms

The project will integrate domestic and overseas existing algorithms, and develop inversion algorithm of global Leaf Area Index, Land Surface Albedo, Emissivity, Downwelling Shortwave Radiation and Photo-synthetically Active Radiation, with the overall accuracy of products better than that of current products.

3.3. Production System Development and Product Generation

(1) The project will integrate domestic and overseas existing algorithms, and develop inversion algorithm of global Leaf Area Index, Land Surface Albedo, Emissivity, Downwelling Shortwave Radiation and Photo-synthetically Active Radiation, with the Production System: It can process multiple global satellite data comprehensively with pre-processing functions of radiometric calibration, geometric correction, image registration, spatio-temporal filtering and cloud detection. It will integrate multiple algorithms for generating the same product.

(2) Production of five global land Products: Spatial resolution of Leaf Area Index, Land Surface Albedo and Emissivity products is 1km-5km, and temporal resolution is 8 days, with time span of 1985-2010; spatial resolution of Downwelling Shortwave Radiation and Photo-synthetically Active Radiation products is 5km, and temporal resolution is 3 hours, with time span of 2008-2010.

(3) These products will be validated with the validation data over 10 typical experimental areas of 5 continents globally, including over 20 typical land cover types.

3.4. Land Surface Simulation and Data Assimilation System

(1) Canopy Radiation Model: It can be used for earth system model and accurately estimate solar radiation and thermal radiation respectively absorbed by canopy and land surface; it can fully consider the differences of features, three-dimensional structure, underlying surface factor and heterogeneity of space, provide direct estimate of spectral land surface albedo corresponding with the number of sensor channel, output direction component and interpret albedo with visual angel restriction and direction land surface albedo.

(2) Dynamic Vegetation Model: Statistical dynamic vegetation model can distinguish different functions and types of vegetation, check and evaluate the simulation results of dynamic model. Dynamic vegetation model can simulate the process of vegetation growth, cycle and withering, and simulate vegetation distribution, seasonal and yearly change features based on function classification of vegetation.

(3) Data Assimilation System: It will include the integrated assimilation method of Kalman filter, ensemble Kalman filter and various particle filters, with high performance parallel computing, and provide land surface data assimilation system software which can assimilate multi-source remote sensing data.

(4) Simulation accuracy of land surface model is better than that of the most advanced model of the same kind in the world. Data Assimilation Demonstration: It will generate land surface system assimilation data sets of China with temporal resolution of 1 hour and spatial resolution of 5km.

3.5. Demonstrative Research on Global Land Surface Change of Land Surface Products

(1) Generating global vegetation phenology data sets of 1985-2010 at spatial resolution of 5km.

(2) Providing quantitative analysis and evaluation of global land surface parameter change trend and space-time features.

(3) Providing the evaluation of the effect of land surface parameter remote sensing products on the improvement of earth system simulation performance.

4. Participating Institutes

The Principal Investigator of this project is Dr. Shunlin Liang, professor of Beijing Normal University, and the Department of Geography, University of Maryland.

The project is funded through the College of Global Change and Earth System Science, Beijing Normal University. All participating institutes include:

(1) Beijing Normal University

(2) Institute of Remote Sensing Application, Chinese Academy of Sciences (CAS)

(3) Cold and Arid Regions Environmental and Engineering Research Institute, CAS

(4) Institute of Atmospheric Physics, CAS

(5) Peking University

(6) University of Electronic Science and Technology of China

- (7) Wuhan University
- (8) Jiangxi Normal University
- (9) Sun Yat-sen University
- (10) Institute of Geographic Sciences and Natural Resources Research, CAS
- (11) Institute of Tibetan Plateau Research, CAS
- (12) National Climate Center, China Meteorological Administration

5. Contact Information

Address: 863 project office, College of Global Change and Earth System Science,
Beijing Normal University, 19 Xijiekouwai Street, Haidian District,
Beijing 1000875, China

Tel: +86-10-5880-9071

Fax: +86-10-5880-3002

E-mail: isl_863@163.com

Website: <http://www.glass-bnu.com>

Invited Experts for the Summer School

No.	Experts	Affiliation	E-mail
1	Prof. Shunlin Liang	Beijing Normal University and University of Maryland (USA).	sliang@umd.edu
2	Prof. Chunxiang Shi	Chinese Academy of Meteorological Sciences, China	cshi@mail.iap.ac.cn
3	Prof. Peng Zhang	National Satellite Meteorological Center, China	zhangp@cma.gov.cn
4	Dr. Suhung Shen	George Mason University and NASA/GSFC, USA	Suhung.Shen@nasa.gov
5	Prof. Jiancheng Shi	University of California, Santa Barbara (UCSB)	shi@icess.ucsb.edu
6	Prof. Yongkang Xue	University of California, Los Angeles, USA	yxue@geog.ucla.edu
7	Prof. Xu Liang	University of Pittsburgh, USA	xuliang@pitt.edu
8	Prof. Kun Yang	Institute of Tibetan Plateau Research, Chinese Academy of Sciences (CAS), China	yangk@itpcas.ac.cn
9	Prof. Soroosh Sorooshian	University of California, Irvine, USA	soroosh@uci.edu
10	Prof. Qingyun Duan	Beijing Normal University, China	qyduan@bnu.edu.cn
11	Dr. Dennis S. Ojima	The H John Heinz III Center for Science, Economics and the Environment, USA	ojima@heinzctr.org
12	Prof. Xin Li	Cold and Arid Regions Environmental and Engineering Research Institute, CAS, China	lixin@lzb.ac.cn
13	Prof. Xiaogu Zheng	Beijing Normal University, China	x.zheng@bnu.edu.cn
14	Prof. Jun Qin	Tibetan Plateau Research, CAS	shuairenqin@yahoo.com.cn
15	Prof. Jeffrey P. Walker	University of Melbourne, Australia	j.walker@unimelb.edu.au
16	Prof. Lixin Lu	University of Colorado, USA	lixin@atmos.colostate.edu
17	Prof. Hongliang Fang	State Key Laboratory of Resources and Environmental Information System, CAS, China	fanghl@lreis.ac.cn
18	Dr. Youlong Xia	NCEP/NOAA, USA	youlong.Xia@noaa.gov
19	Dr. Xianhong Xie	Beijing Normal University, China	xiexh@pku.edu.cn
20	Dr. Bhanu Neupane	UNESCO Office New Delhi	b.neupane@unesco.org
21	Dr. Ramasamy, Jayakumar	UNESCO Office Beijing	r.jayakumar@unesco.org
22	Dr. Anil Mishra	UNESCO Office Paris	a.mishra@unesco.org