GLOBAL ENERGY AND WATER CYCLE EXPERIMENT

WORLD CLIMATE RESEARCH PROGRAMME

GSWP-2

The Second Global Soil Wetness Project

Science and Implementation Plan

December 2002

IGPO Publication Series

No. 37

The Second Global Soil Wetness Project
GSWP-2

Compiled by

Paul Dirmeyer and Xiang Gao

Center for Ocean-Land-Atmosphere Studies

and

Taikan Oki

University of Tokyo

December 2002

IGPO Publication No. 37

EXECUTIVE SUMMARY

The Global Soil Wetness Project (GSWP) is an ongoing environmental modeling research activity of the Global Land-Atmosphere System Study (GLASS) and the International Satellite Land-Surface Climatology Project (ISLSCP), both contributing projects of the Global Energy and Water Cycle Experiment (GEWEX).

Its goals are to:

- Produce state-of-the-art global data sets of land surface fluxes, state variables, and related hydrologic quantities.
- Develop and test large-scale validation, calibration, and assimilation techniques over land.
- Provide a large-scale validation and quality check of the ISLSCP data sets.
- Compare Land Surface Schemes (LSSs), and conduct sensitivity studies of specific parameterizations and forcings, which should aid future model and data set development.

GSWP-2 is closely linked to the ISLSCP Initiative II data effort, and LSS simulations in GSWP-2 will encompass the same core 10-year period as ISLSCP Initiative II (1986-1995).

There are five basic categories of participants in GSWP-2: the operational centers, the land-surface modelers, validators of the model output, those involved in remote sensing applications, and other users of the model output. An Inter-Comparison Center will collect results from participating models, perform consistency checks, and basic comparisons.

A major product of GSWP-2 will be a multi-model land surface analysis for the ISLSCP Initiative II period. This will be a land surface analog to the atmospheric reanalyses, and will include estimates of uncertainties based on inter-model spread. The science plan also includes *in situ* validation with data from field campaigns, observational networks and long-term monitoring sites. Modeling sensitivity studies will involve re-integrating the LSSs over part or all of the global, 10-year domain to test the response of the models to changes in meteorological data (including choice of reanalysis products, impacts of bias correction, sensitivity to the range in observational estimates, and impacts of rain-gauge under-catch) and surface parameters. Comparisons to land models of simple and intermediate complexity will also be conducted.

A new thrust for GSWP-2 is a stronger connection to applications in remote sensing. In addition to the classical attempts to validate the typical land-surface state variables using satellite retrievals, GSWP-2 also intends to expand the validation and assimilation capabilities of current LSSs. This is to be done by the development of algorithms by which LSSs can directly report brightness temperatures, like those sensed by instruments in orbit.

All data sets will conform to the Assistance for Land-surface Modeling Activities (ALMA) standards set forth within GLASS. New Internet data server technologies will be used to distribute and analyze data, reducing archiving and data management burdens on participants.

TABLE OF CONTENTS

List Of Tables

LIST OF FIGURES

ACRONYMS

AGCM	Atmospheric General Circulation Model
AMIP	Atmospheric Model Inter-comparison Project
ALMA	Assistance for Land-Surface Modeling Activities
ARM-CART	Atmospheric Radiation Measurement Program – Cloud and Radiation Testbed
BATS	Biosphere-Atmosphere Transfer Scheme
CF	Climate and Forecast
COLA	Center for Ocean-Land-Atmosphere Studies
CREST	Core Research for Evolutional Science and Technology
CRU	Climate Research Unit
DIS	Data Sets Information Systems
DODS	Distributed Oceanographic Data System
DOE	Department of Energy
DVM	Dynamic Vegetation Models
ECMWF	European Centre for Medium-Range Weather Forecasts
EDC	EROS Data Center
EMC	Environmental Modeling Center
ENSO	El Niño/Southern Oscillation
ERA-40	ECMWF Re-analysis-40
EROS	Earth Resources Observation Systems
FIFE	First ISLSCP Field Experiment
FTP	File Transfer Protocol
GEWEX	Global Energy and Water Cycle Experiment
GCM	Global Atmospheric Climate Model
GDS	GrADS-DODS server
GLASS	Global Land-Atmosphere System Study
GPCC	Global Precipitation Climatology Centre
GPCP	Global Precipitation Climatology Project
GrADS	Grid Analysis and Display System
GSFC	Goddard Space Flight Center
GSWP	Global Soil Wetness Project
GTS	Global Telecommunication System
GVaP	GEWEX Water Vapor Project
HAPEX-Sahel	Hydrology-Atmosphere Pilot Experiment in the Sahel
ICC	Inter-Comparison Center
IGBP-DIS	International Geosphere-Biosphere Programme Data Information System
IIS	Institute for Industrial Studies
IMOP	Intensive Model Output Period
ISCCP	International Satellite Cloud Climatology Project
ISLSCP	International Satellite Land-Surface Climatology Project

JST	Japan Science and Technology Corporation
LAI	Leaf Area Index
LDAS	Land-Surface Data Assimilation
LMD	Laboratoire de Météorologie Dynamique du C.N.R.S.
LSSs	Land Surface Schemes
MODIS	Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NDVI	Normalized Difference Vegetation Index
NetCDF	Network Common Data Format
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
PILPS	Project for the Inter-Comparison of Land-Surface Parameter Schemes
PFT	Plant Functional Type
SCAN	Soil Climate Analysis Network
SiB	Simple Biosphere
SMMR	Scanning Multichannel Microwave Radiometer
SRB	Surface Radiation Balance
SSiB	Simplified Simple Biosphere
SSM/I	Special Sensor Microwave/Imager
UMBC	University of Maryland, Baltimore County
USAF	United States Air Force
USDA	United States Department of Agriculture
UTC	Coordinated Universal Time
WCRP	World Climate Research Programme

1.0 INTRODUCTION

1.1 Overview

The Global Soil Wetness Project (GSWP) is an ongoing environmental modeling research activity of the Global Land-Atmosphere System Study (GLASS) and the International Satellite Land-Surface Climatology Project (ISLSCP), both contributing projects of the Global Energy and Water Cycle Experiment (GEWEX) in the World Climate Research Programme (WCRP). GSWP is charged with producing as a community effort global estimates of soil moisture, temperature, snow water equivalent, and surface fluxes by integrating one-way uncoupled land surface schemes (LSSs) using externally specified surface forcings and standardized soil and vegetation distributions. GSWP-2 will produce the best model estimates of the land-surface water and energy cycles over a ten year period. This project will include an evaluation of the uncertainties linked to the LSSs, their parameters and the forcing variables.

The goals of GSWP are to:

- Produce state-of-the-art global data sets of land surface fluxes, state variables, and related hydrologic quantities.
- Perform large-scale model evaluation, validation and calibration over land.
- Provide a large-scale validation and quality check of the ISLSCP data sets.
- Compare LSSs, and conduct sensitivity studies of specific parameterizations and forcings which should aid future model and data set development.

GSWP-2 is closely linked to the ISLSCP Initiative II data effort (http://islscp2.sesda.gov/), and LSS simulations in GSWP-2 will encompass the same core 10-year period as ISLSCP Initiative II (1986-1995). Participation by land surface modelers, remote sensing scientists, field researchers, data collectors and others is voluntary. The basic operation of GSWP-2 is currently supported in the United States by the National Aeronautics and Space Administration (NASA), and in Japan by the Core Research for Evolutional Science and Technology (CREST) of the Japan Science and Technology Corporation (JST) under a project titled "Modeling Global Hydrological Cycles and World Water Resources Coupled with Human Activities."

1.2 Background

GSWP-2 is the follow-on project to GSWP-1, a 2-year pilot phase based on the ISLSCP Initiative I data set for 1987-1988. GSWP-1 was also an offline land-surface modeling and evaluation effort conducted at a spatial resolution of 1 degree.

The original motivation for GSWP stemmed from the paradox that soil wetness is an important component of the global energy and water balance, but it is unknown over most of the globe. Soil wetness is the reservoir for the land surface hydrologic cycle, it is a boundary condition for atmosphere, it controls the partitioning of land surface heat fluxes, affects the status of overlying vegetation, and modulates the thermal properties of the soil. Knowledge of the state of soil moisture is essential for climate predictability on seasonal-annual time scales. However, soil moisture is difficult to measure *in situ*, remote sensing techniques are only partially effective, and few long-term climatologies of any kind exist. The same problems exist for snow mass, soil heat content, and all of the vertical fluxes of water and heat between land and atmosphere. Even a consistent definition of soil wetness is elusive.

In GSWP-1, the ISLSCP Initiative I data set was used to supply boundary conditions, model parameters and meteorological forcing for more than one dozen LSSs integrated by members of the Production Group, which then reported a set of standard output data at thrice-monthly intervals to an Inter-Comparison Center, which performed consistency checks and basic comparisons. These data were then made available to the Validation Group, which performed *in situ* and remote sensing validation, as well as hydrologic validation of LSS runoff against observed streamflow. The Production Group members also individually performed assigned sensitivity tests to determine the impact of changes in model parameterizations and data sets on the results from the LSS simulations.

The pilot phase of the GSWP revealed that the quality of simulated land surface quantities, particularly in the hydrologic cycle, is a strong function of the availability of *in situ* observations feeding into the analysis stream of meteorological forcing data (Oki et al. 1999). Where forcing and parameter data are of good quality, the participating LSSs performed well. LSSs were found to have some variation in the partitioning of precipitation between runoff and evaporation, but much larger differences were found among the soil moistures simulated by the LSSs (Entin et al. 1999). The data sets have also been used in a number of coupled land-atmosphere climate modeling studies, which have shown the impact of high-quality soil moisture data, and land surface variability on climate simulations (Mocko et al. 1999; Douville and Chauvin 2000; Dirmeyer 2000, 2001; Douville et al., 2001; Douville 2002). Participation in GSWP-1 gave land surface modelers a global testbed for improving their LSSs, and many of the participants have used it for that purpose.

A special issue of the Journal of the Meteorological Society of Japan (Vol 77, No. 1B; 1999) was published containing the preliminary results of GSWP. An overview article was also published (Dirmeyer et al., 1999). Subsequent to that, other papers have been published — a complete bibliography is maintained on the GSWP web site (http://www.iges.org/gswp/).

More recently, the Rhône Aggregation Experiment (Rhône-AGG; Boone et al., 2001, 2002) has been completed. The entire Rhône model domain size contains the Rhône river basin in France, and is on the order of that of a coarse-resolution Global atmospheric Climate Model (GCM). However, the atmospheric forcing, the soil and vegetation parameters, and the observed river discharges are available at a significantly higher spatial resolution. This is accomplished through use of the Rhône modeling system, which is comprised of a distributed hydrological model, an analysis system to determine the near-surface atmospheric forcing from a combination high resolution monitoring and model assimilation, and a LSS interface.

It was the interest of Rhône-AGG to examine how the simulations from a wide range of LSSs (used in GCMs, Numerical Weather Prediction [NWP] models, mesoscale atmospheric models or hydrological models) are impacted by changing the spatial resolution over the domain (8km, 0.5°, and 1°). The main goals of the Rhône-AGG were to examine how various state of the art LSSs are able to simulate the river discharge over several annual cycles when inserted into the Rhône modeling system, and to explore the impact of the various scaling or aggregation methods on the simulation of certain components of the hydrological cycle (such as snow cover and surface runoff). A limited number of multi-year simulations (August 1985 through July 1989) were performed and studied by approximately 20 LSSs.

The conclusions from Rhône-AGG are that overall all LSSs simulated monthly discharge well, but only 9 of 15 simulated good statistics at the daily scale. Parameterization of surface (sub-grid) runoff was critical: at the gauging station at Viviers, seven of nine LSSs with sub-grid runoff parameterizations had significant skill, while only two of the six remaining LSSs showed skill. This same trend was found at other stations. The sub-grid runoff and drainage parameterizations were greatly impacted by scale. Surface runoff was generally reduced as resolution decreased. Drainage response was mixed among the LSSs, owing to different evaporation and baseflow responses. The most consistent response was in terms of snowpack: explicit snow schemes generally compare best with observations and discharge in Alpine basins. Evapotranspiration changes were generally offset by runoff differences as resolution was changed. LSSs tended to have a slightly wetter equilibrium soil moisture state at low resolution (less than 5% change for most LSSs). Grid resolution greatly impacted the snow simulation: snowmelt occurred too soon when the LSSs were run at low resolution, primarily owing to warmer conditions over the snow cover. One model, which included sub-grid altitude banding, showed results that were largely unaffected by changes in resolution. More information is available at the project website: http://www.cnrm.meteo.fr/mc2/projects/rhoneagg/.

1.3 Participants

There are five basic categories of participants in GSWP-2, and there may be some overlap as individuals or groups can be listed in more than one category. The categories are the operational centers for the project, the land-surface modelers, evaluators of the model output, those involved in remote sensing applications, and end users of the model output. The relationship among these groups is illustrated in Figure 1.

The members of the first category are the two operational centers for GSWP-2. These are the Center for Ocean-Land-Atmosphere Studies (COLA) in Calverton, Maryland, USA, and the Institute for Industrial Studies (IIS) at the University of Tokyo, Japan. These operational centers generate the forcing data and boundary conditions for the participating modelers, collect and compare the results of the models, and generally oversee the functioning of the project. Primary responsibility for data set production is at COLA. The IIS will maintain an Inter-Comparison Center (ICC) that will collect the model output and perform basic consistency checks on the results. The ICC will then redistribute the data to the validation and remote sensing groups, as well as generate a multi-model analysis, along with COLA, for broader distribution. There will be regional mirror sites of input data, and possibly also for model output submission, to ameliorate trans-oceanic Internet connectivity problems.

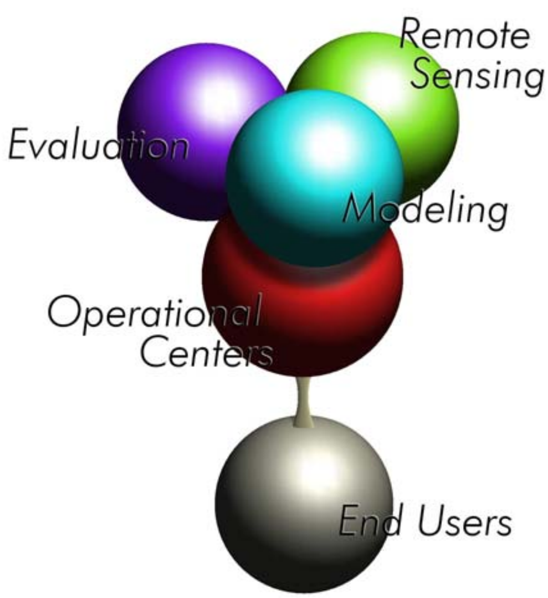

Figure 1. The five categories of participants in GSWP-2.

The second category includes the participating land surface modelers (called the "Production Group" during GSWP-1). The members of this group will run their LSSs with the provided forcing and boundary conditions, and furnish the results of this "baseline" integration to the ICC. The modeling group is open to anyone with a unique LSS and an interest in participating. No direct financial support can be provided from

GSWP to modeling groups for their participation. However, the access to data, analysis tools, validation, and expertise should make participation attractive, as has been the case in previous Project for the Inter-comparison of Land-surface Parameter Schemes (PILPS) and GSWP experiments. Once a modeling group has set up its LSS for participation in GSWP, it has a ready testbed for offline testing of changes and improvements to the LSS. Those who have previously participated in the Rhône-AGG or PILPS 2(e) experiments should find the transition to GSWP-2 very easy, because of the adherence to the ALMA data standards. The modeling groups will also participate in sensitivity studies that are designed to help elucidate the workings of both the models and the land surface component of the climate system.

The evaluation group (called the "Validation Group" in GSWP-1) is involved with comparison and validation of the model results with *in situ* observations. Largely this means local validation on time series of model output at a single grid point that corresponds to a field campaign or long-term monitoring site. For hydrologic validation, basin-scale runoff will be validated against stream gauge measurements. This group will also include a multi-model evaluation, that will feed back upon efforts of the operational group to produce multi-model analyses, and those involved in other research efforts within GSWP. Specifically, the other research efforts will be comparisons of the model results beyond that of the ICC, including comparisons to simple and intermediate models, and examination of any LSSs' ability to simulate other LSSs.

The remote sensing applications group is a new focus for GSWP, concerned with the large-scale use of satellite data for validation of LSSs, and ultimately assimilation of satellite data into LSSs. Studies of model performance that can be evaluated by satellite observations, such as snow coverage or radiative skin temperature, will be performed in this group. This will also involve the development and application of algorithms by which current LSSs can provide as output the brightness temperatures that are directly observed by space-based platforms. By performing validation and calibration of observable quantities, rather than converting observable quantities to model state variables, a more direct comparison can be achieved.

1.4 End Users

Finally, there is the group of users outside the project, a.k.a. the "End Users." This group includes those with the potential to take advantage of the GSWP-2 model output data, as well as the unique atmospheric forcing and boundary condition data sets. These people are, by definition, not contributors to GSWP-2 science goals, but are scientists, engineers, and other clients who may use the data for their own unique purposes.

There continues to be a strong need for global data sets, and the model products of GSWP-2 will expand upon those data sets provided by ISLSCP Initiative II by supplying enhanced versions of ISLSCP Initiative II data sets in ALMA format, as well as estimates of land surface fluxes and state variables with global coverage. End users may have interest in local, regional or global data, mean diurnal and annual cycles or synoptic, seasonal and interannual variations. Hydrologists, engineers, biogeochemists, agronomists, botanists, ecologists, geographers, climatologists, and educators may all have an interest in the products of GSWP-2.

2.0 SCIENCE PLAN

This section outlines the science plan for the global element of GSWP-2. Details of the execution of the project are described in Section 3 — the Implementation Plan. There are five parts to the science plan: surface model analysis, validation, sensitivity studies, remote sensing applications, and basic model investigations.

2.1 Global continuous multi-model analysis

A major product of GSWP-2 will be a multi-model land surface analysis for the ISLSCP Initiative II period. This will be a land surface analog to the atmospheric reanalyses, but encompassing an ensemble of different LSSs. There will be a monthly seasonal cycle data set of monthly values, and a larger data set for the entire series. Using the results of multiple LSSs will provide a model-independent result — compilation of a single analysis from an ensemble of LSSs may be a simple or complex exercise, and is discussed in Section 2.2.4. Of particular value, uncertainty estimates can be put on all of the fields, based on inter-model spread. Additional uncertainties regarding forcing data can be quantified, based on the sensitivity studies described later in this plan.

In addition to publication of the analysis data sets themselves, a journal article will be written describing the seasonal cycle, interannual variability and anomalies, including the signals of significant climate events. This effort will deliver on an unfulfilled goal of GSWP-1.

The multi-model analysis will be the principal data product of GSWP-2. Based on the reception of the GSWP-1 and ISLSCP Initiative I data sets, it is anticipated that a wide range of earth scientists, engineers, educators and social scientists will find the multi-model analysis useful in their work.

2.1.1 Mean seasonal cycle

The fundamental time interval for climate is the annual cycle. A multi-model climatology of the annual cycle at 1° resolution will be produced for all of the standard output variables, and will be served to the community at large. The purpose of this climatological data set will be to provide a best analysis of basic land surface state variables and fluxes at high spatial resolution. Many of the users cited above could make use of such a data set as a benchmark or boundary condition for other model simulations, or for ecological, geophysical, or economic calculations. For users who do not need information on interannual variations during the 10-year period of GSWP-2, this data set will be smaller and easier to use than the complete 10-year data set. It will consist of 12 monthly means for all of the output fluxes, surface state variables, and selected subsurface state variables. Because different LSSs have different vertical discretizations of soil layers, and different LSSs have different operating ranges for soil wetness (Koster and Milly 1997), soil wetness will be represented in terms of a few column integrated deviations from the annual mean calculated over standard depths.

2.1.2 Interannual variability/anomalies

A multi-model analysis for the complete 10-year period of GSWP-2 will serve as a land-surface analog to the atmospheric reanalyses. As the coverage in space and time of land surface observations is meager, an analysis using state-of-the-art models driven by the best-possible forcing data (based on atmospheric observations) offers the best proxy to a global observational network.

The analysis will contain 10 years of complete annual cycles of state variables and fluxes over the land surface at 1° resolution. As with the climatological analysis, monthly means will be provided. Soil moisture will be in terms of anomalies from the mean annual cycle. The data set can be used to drive ecological, biogeochemical, agricultural, meteorological or economic models. It can also be analyzed in studies of climate variability at the land surface on seasonal-interannual time scales. However, the 10-year time series will be too short to be used for the detection and analysis of trends, such as those associated with global warming. Thirty to forty years of data, at the very least, are necessary for such an analysis.

The period of GSWP-2 encompasses the 1986-1987 and 1991-1992 El Niño events, the 1994-1995 weak warm event, and the strong 1988-1989 La Niña. Figure 2 Shows the seasonal progression of the Niño 3.4 SST anomaly during the GSWP-2 period (http://www.cpc.ncep.noaa.gov/data/indices/index.html).

In addition to the cycle of El Niño/Southern Oscillation (ENSO), there were a number of significant regional hydrological anomalies during the GSWP-2 period. The late 1980s and early 1990s were a period of considerable variability in rainfall over many regions that rely heavily on seasonal precipitation regimes. For instance, 1986 and 1987 were consecutive years of very poor monsoon rainfall over India, as indicated in the all-India rainfall index (http://tao.atmos.washington.edu/data_sets/india/parthasarathy.html). Figure 3 (top panel) shows this, along with the very wet year in 1988. There are smaller but still relevant variations during the 1990s. Over the Sahel region of Africa, the decades-long drought that began during the 1960s persisted through the GSWP-2 period (http://www.cru.uea.ac.uk/tiempo/floor2/data/sahel.htm). Figure 3 (middle panel) shows the wet-season anomalies relative to the 30-year climatology period of 1961-1990 — itself a severely dry period during the century. Thus, the near-zero anomaly years during the later part of the period are in fact dry relative to the long-term record. The droughts shown in 1986, 1987 and 1990 were particularly severe. The only truly wet year in the period is 1994 – 1988 was actually near normal when compared to the rest of the 20th century.

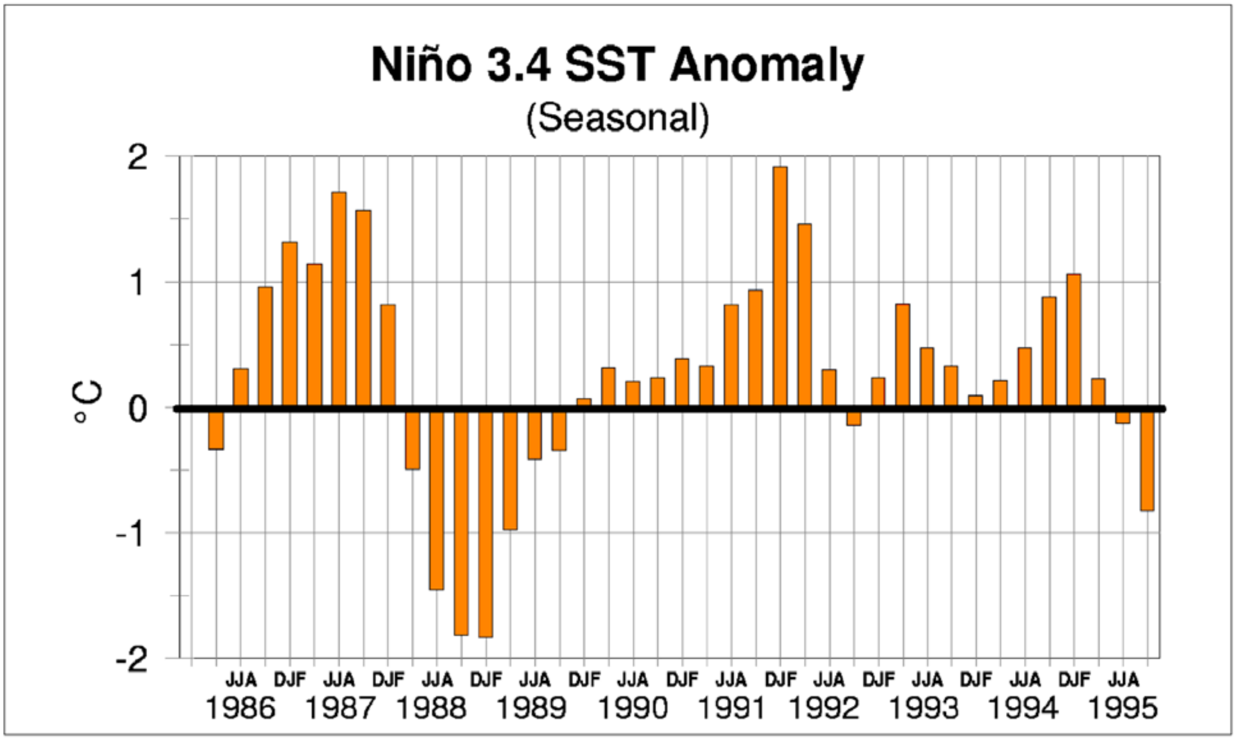

Figure 2. Seasonal eastern tropical Pacific sea surface temperature anomalies.

The final example shown in Figure 3 (lower panel) is for the wet-season over the Nordeste region of Brazil (http://tao.atmos.washington.edu/data_sets/brazil/). The span of the anomalies between the wet year of 1986 and the dry year of 1993 is nearly as large as any anomalies measured during the 20th century.

These examples show that there was indeed interesting variability in the ocean and atmospheric components of climate during the period of the GSWP-2 experiment. There is great potential for the GSWP-2 experiment to enhance our understanding of the variability of the land surface component associated with these and other climate variations.

2.1.3 Uncertainty in model-derived surface fluxes and state variables

There is no doubt that the different LSSs participating in GSWP-2 will not all produce identical estimates of surface fluxes and land surface state variables. There is no doubt that the different LSSs participating in GSWP-2 will not all produce identical estimates of surface fluxes and land surface state variables. The range of estimates calculated by the participating LSSs is one measure of our uncertainty in these terms, lacking complete coverage of observational data for validation. This is a part of the inter-comparison effort, although there will be less emphasis on comparison of standardized soil wetness indices than in GSWP-1, as that was found to be an unsuitable means of comparison (Saleem and Salvucci, 2002).

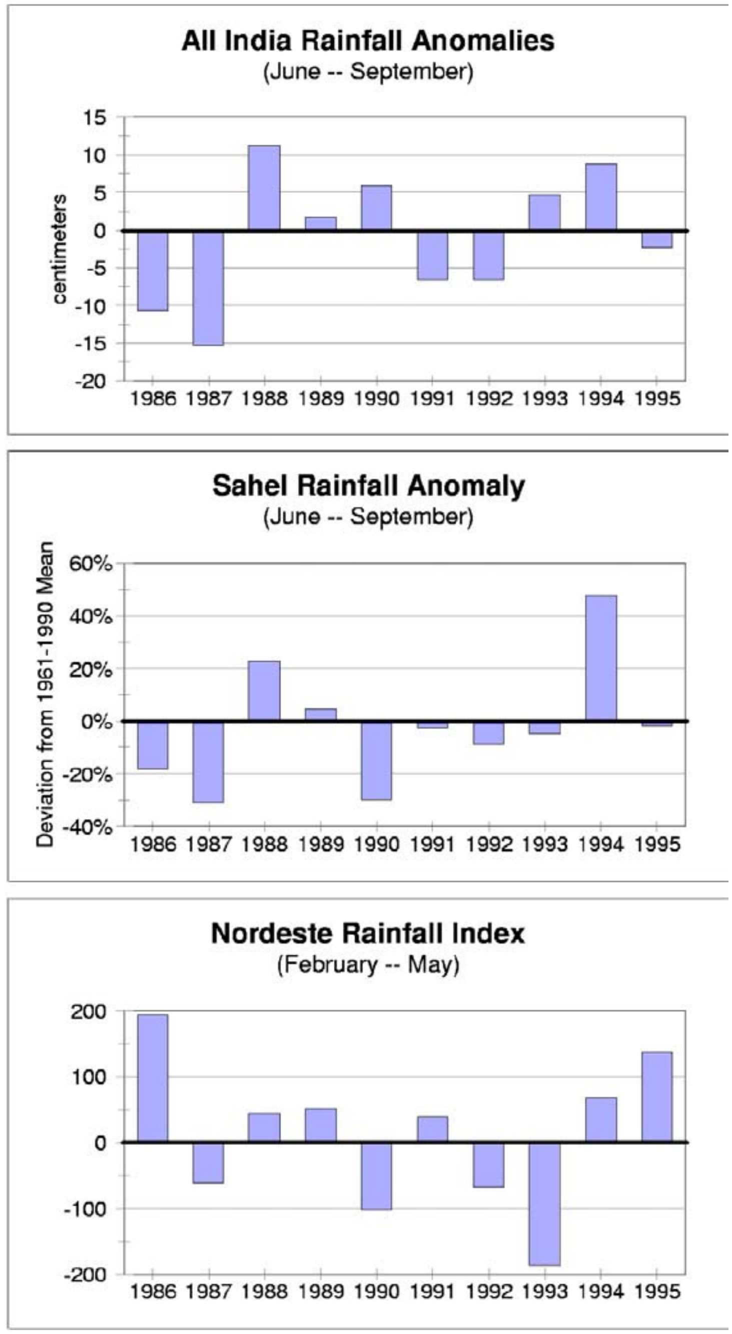

Figure 3. Regional rainfall indices

2.1.3.1 *Establish the multi-model envelope*

GSWP-2 will establish the envelope of model certainty for globally complete estimates of surface and sub-surface temperature and water storage, and fluxes of energy and water between land and atmosphere. This will be done for the ideal scenario that the forcing data are well known (i.e. identical for each LSS) and only the models themselves differ. The uncertainty in estimates of these quantities by uncoupled LSSs is measured by the spread of estimates among all LSSs (or, to mute the impact of problematic extreme outliers, some statistical measure based on standard deviations or interquartile ranges). Incertitude due to errors or uncertainty in forcing data will be addressed in the sensitivity experiments (Section 2.3).

2.1.3.2 *Uncertainty by variable*

It is quite likely that different variables will have different relative and absolute ranges of inter-model spread, and thus different levels of uncertainty associated with them. For instance, it was found that in GSWP-1, given the same precipitation forcing, different models produced nearly identical skin temperatures, similar time series of evaporation, but drastically different mean levels of soil moisture. The variables of the National Centers for Environmental Prediction/ National Center for Atmospheric Research (NCEP/NCAR) reanalysis were assigned letter designations based on how model-dependent, and thus potentially how observationally consistent, each variable was (Kalnay et al., 1996). This letter designation served as an indicator of confidence in each variable. A similar ranking may be applied for the multi-model analysis, based on the clustering of the values of each variable. Those that are strongly tied to the forcing data should show the least spread, and those that are the result of parameterizations removed several computational steps from the forcing data may show the most variation.

2.1.3.3 *Uncertainty by region*

Similarly, there may geographical dependence to the inter-model spread. For instance, in a hot desert region where moisture is in short supply but radiative energy is in abundance, one would expect that all models would partition nearly all of the specified precipitation to evaporation, and little or none to runoff, exhibiting a high level of inter-model consistency. Yet over a temperate mid-latitude region with seasonal snowpack, there may be a large range of simulated runoff and evaporation among the same models. Thus, a blanket assertion regarding the confidence interval for a given variable in the multi-model ensemble may be misleading. Spatial variability will be assessed as well.

2.2 Evaluation

PILPS, in its Phase 2, has conducted and continues to craft local land surface modeling experiments built around nearly-complete sets of forcing and validation data at a single location (Henderson-Sellers et al., 2002). It is not the intent of this *in situ* validation program to duplicate that effort. Rather, using the global forcing data sets, local validation may be performed when and where such data are available.

Observations of land-surface state variables are sparse; in several cases individuals have gone to great effort to collect related data sets from locations around the globe. These data are invaluable for LSS validation, and their participation in GSWP-1 helped make that pilot project a success. With a 10-year period of coverage, validation efforts can now span more observational data sets, and examine interannual variations.

Much of the data described here is not sufficiently complete (in temporal coverage, or in terms of the breadth of quantities measured) to be the basis of a stand-alone PILPS-2 experiment. With incomplete data it is not possible that closure of local and water energy balances can be obtained — it is unlikely that even all of the contributing terms to these balances would be measured in sufficient totality at a 1° scale. Furthermore, budgets from *in situ* measurements rarely come within a few percent of closure, due to instrument error, miscalibration, and other real-world hindrances. It is the goal of the *in situ* validation effort of GSWP-2 to provide measurements of key variables and fluxes of sufficient coverage and accuracy to provide an estimate of the ground truth, when and where available, for the LSS. Of course, this is far from ideal. The intent is to make the best use of the information available, as incomplete as it is.

2.2.1 Field campaigns

There have been a large number of field campaigns during the span of the GSWP-2 period that are of relevance to climate over the land surface. A partial list is presented in Table 1. These campaigns covered a broad range of ecological and climate regimes, and provide *in situ* observational data of a higher quality, better resolution (spatial and temporal), and greater range of variables than is otherwise available.

Table 1. Some field campaigns that overlap the GSWP-2 period.

Name	Location	Period	References
Boreal Ecosystem-Atmosphere Study (BOREAS)	Central Canada	1993-1996	Sellers et al. (1997), Hall (1999)
GEWEX Continental-Scale International Project (GCIP)	Mississippi River basin, USA	1995-2000	Coughlan and Avissar (1996), Lawford (1999)
Baltic Sea Experiment (BALTEX)	Baltic Sea basin	1994-2001	Raschke et al. (1998)
Mackenzie GEWEX Study (MAGS)	Mackenzie River basin, Canada	1994	Stewart et al. (1998)
Anglo-Brazilian Climate Observation Study (ABRACOS)	Manaus, Ji-Paraná and Marabá, Brazil	1990-1995	Gash and Nobre (1997)
European International Project on Climatic and Hydrological Interactions between Vegetation, Atmosphere and Land Surface (ECHIVAL) Field Experiment in Desertification Threatened Areas (EFEDA)	Southeastern Spain	1991-1995	Bolle et al. (1993)
First ISLSCP Field Experiment (FIFE)	Central Kansas, USA	1987-1989	Sellers et al. (1992), Hall and Sellers (1995)
Hydrological and Atmospheric Pilot Experiment - Modelisation du Bilan Hydrique (HAPEX-MOBILHY)	Southern France	1985-1987	Andre et al. (1989)
Hydrological and Atmospheric Pilot Experiment in the Sahel (HAPEX Sahel)	Western Niger	1991-1993	Goutorbe et al. (1994)
Hei Ho River Basin Field Experiment (HEIFE)	Gansu Province, China	1992-1993	Wang et al. (1993)
Northern Hemisphere Climate Processes Land Surface Experiment (NOPEX)	Central Sweden	1994-1996	Halldin et al. (1999)
Observation at Several Interacting Scales (OASIS)	Murray-Darling basin, Australia	1994-1995	http://www.clw.csiro.au/research/environment/interactions/oasis/
Monitoring the Usable Soil Reservoir Experimentally (MUREX)	Southwestern France	1995-1997	Calvet et al. (1999)

These data provide a unique validation opportunity for the various LSSs participating in GSWP-2. Many LSSs experience limited validation because of the effort necessary to collect and apply the relevant data sets from multiple observational campaigns. Thus, validation is often limited to one or two locations. With the help of ALMA, these data may be synthesized into a more useful standard and applied to the suite of LSSs (see Section 3.1.1.3).

2.2.1.1 *Global (ISLSCP) data sets versus local data*

Many of the field campaigns included a nearly-complete sampling of local near-surface meteorological variables. Some even include direct measurements of radiation. This information should be compared to the ISLSCP Initiative II based forcing data *before any validation* of the LSSs with other components of the *in situ* data. If there are significant differences between the NCEP or ECMWF reanalyses and the local meteorological measurements, then the interpretation of the results from any attempt at *in situ* validation must be altered accordingly.

To some extent, this comparison is a validation of the reanalyses themselves. However, differences, even systematic differences, do not necessarily reflect badly on the reanalysis products. Because of the differences in spatial scales, and the possible existence of regional or local microclimates in the domain of the field campaigns, one should not expect complete agreement between field campaign meteorological measurements and reanalysis data. Nonetheless, differences in the meteorological variables may go a long way toward explaining apparent failings in the simulations of the LSSs in these locations. Likewise, differences between GSWP and actual local soil and vegetation conditions should be considered when comparing results to local measurements.

2.2.1.2 *LSS output versus in situ measurements*

The principal validation of the LSSs will be between output variables from the LSS and the corresponding field measurements. The variables that can be validated will fluctuate among the field campaigns, but some measure of validation should be possible.

2.2.2 Observational networks and long-term monitoring

The broadest available archive for *in situ* soil moisture measurements is the Global Soil Moisture Data Bank (Robock et al., 2000). It includes station data from Russia, China, Mongolia, India and the United States covering between six and ten of the years of the GSWP-2 period. Some of these data formed the basis of the *in situ* evaluation for the pilot phase of GSWP, and it is expected that they will again play a major role in the validation effort.

Soil wetness data are also available for 19 of the USDA Soil Climate Analysis Network (SCAN) sites that typically start in October 1994. The data content includes soil moisture and temperature profiles (5, 10, 20, 50, 100 cm depth), meteorological data, and some flux data (surface water, radiation and heat), sampled at 6-hour intervals. LSS output for 1995 will be reported at a 3-hour interval (see Section 3.1.4.2), and can be compared to the high temporal resolution SCAN data.

Snow cover and depth data for the entire 10-year period are available over the Northern Hemisphere from the USAF Daily Snow Depth Analysis (Fennessy and Schlosser 2002). This analysis is a blend of *in situ* data and empirically-based estimates.

There exist other potential sources for *in situ* flux measurements (e.g., FluxNet, ARM-CART) and surface state variables (e.g., Oklahoma Mesonet, SnoTel).

2.2.3 Streamflow

Runoff fluxes from all participating LSSs will be routed with common river routing schemes to compare with streamflow measurements across a large portion of the globe, as an assessment of the simulation of annual, seasonal, and interannual variations in surface hydrology. Modeled streamflow will be assessed on several timescales. Climatological annual mean model-routed streamflow will be compared to observed discharge for basic water balance checks, bias audits, and consideration of impacts of irrigation withdrawals. The interannual variation of annual runoff can be used to assess the unbiased performance of the LSSs in simulating climate variability in the surface water cycle, as well as the drift in model state variables over the period. Monthly or seasonal discharge will be used to assess the basin-scale seasonal water balance, the simulation of snow melt, the monsoon signal, and the reddening of the precipitation spectrum in the LSSs. Daily discharge without missing data over the entire 10-year period exists for at least 25 stations, with basin sizes greater than 2,500 km^2, so it may be possible to examine more detailed hydrographs in these areas for flood and drought frequencies and extremes. Flow duration curves can be examined, and the ratio of surface/sub-surface runoff may be examined by “traditional” separation of the hydrograph both for LSS outputs and observations.

Similarly, large basin comparison of model water storage change with observed atmospheric moisture flux convergence minus discharge will be performed. This evaluation may also uncover problems in the forcing data and models at large basin scales.

2.2.4 Multi-model validation

Finally, we propose an effort for multi-model validation. That is, we propose a validation of the ensemble multi-model estimate of surface fluxes and state variables. It is the experience of the meteorological community that even simple averaging across models helps ameliorate the systematic errors of individual models, and usually produces a superior estimate to that given by any individual model, although more sophisticated ensembling methods may yield further improvements (Kharin and Zweirs, 2002; Krishnamurti et al., 1999; 2000). This will be pursued further, by trying to create an optimum ensemble, based on the assessed strengths and shortcomings of the results of the various LSSs.

This element would feed back on production of the multi-model analysis. If a clearly superior method of ensembling other than simple averaging of the results of the participating LSSs, the multi-model analysis described in Section 2.1 would be conducted with the better method.

2.2.4.1 *Simple ensembling*

The simplest approach to ensembling is to take an arithmetic mean of all ensemble members. This is commonly done for multiple simulations with the same atmospheric general circulation model (AGCM) in weather and climate prediction. It can be shown that for a reasonably large ensemble of simulations with similar statistics, the mean-square error of an ensemble mean forecast will converge to half the value of the mean-square error of an individual ensemble member. This does not guarantee a better simulation of anomalies by the ensemble mean, as measured, for instance, by a spatial anomaly correlation. Individual members may still outperform the ensemble mean in a given forecast. However, when the same model is used to produce each ensemble member in each forecast, there is no reason to expect one realization to consistently outperform the ensemble mean.

When we average together the results of different models, which may have fundamentally different statistics, it is mathematically possible that one or more members would consistently outperform or underperform the ensemble mean, in terms of mean-square error, correlation, or other metrics of skill. However, it is likely, based on experience with atmosphere and ocean models, and more importantly our experiences in the PILPS experiments, that no individual LSS will exhibit clear superiority of skill in the GSWP-2 simulation. Thus, our default "best analysis" for GSWP-2 will be a simple arithmetic mean of the participating LSSs.

2.2.4.2 *Optimal ensembling*

Kharin and Zweirs (2002) have found that for climate prediction of atmospheric 500 hPa geopotential height over the tropics, a bias-removed multi-model ensemble mean performs best when the ensemble size for AGCMs is larger than a half dozen or so, and a regression-improved ensemble mean performs best for small ensembles. Over the mid-latitudes, the bias-removed ensemble mean performs best, but only slightly better than climatology. This result was obtained after comparing seven different approaches, including regression-improved multi-model ensembling like that of Krishnamurti et al. (1999, 2000). A similar comparison can be made for LSSs in an offline analysis mode to determine whether there is method of ensembling that is superior to a simple average. However, the physics of land surface models is fundamentally different than the fluid dynamics of atmosphere or ocean models, so it is not clear whether such approaches will yield the same results in this context.

Even without global fields for validation and assessing skill, the exercise of multi-model ensembling with offline LSS simulations is a useful one, to assess, for instance, what may be gained by multi-model participation in land-surface data assimilation (LDAS). One participating LSS can be arbitrarily chosen to represent the truth, and the ensemble of the remaining models could be validated against the truth simulation. This would give a practical measure of the various ensembling techniques.

2.3 Sensitivity studies

The default meteorological forcing data set for GSWP-2 will be the ISLSCP Initiative II regridding of the NCEP-DOE reanalysis of 1986-1995 (Kanamitsu et al. 2002), with corrections to the systematic biases in the reanalysis fields made by hybridization of the 3-hourly analysis with global observationally-based gridded data sets at lower temporal resolution (see Section 3.1). The original ISLSCP Initiative II data sets will not have hybrid correction, but will contain only the uncorrected reanalysis products. Radiation forcing will come from the 3-hourly Surface Radiation Balance (SRB) product directly without hybridization with reanalyses estimates.

Modeling sensitivity studies will involve re-integrating the LSS over part or all of the global, 10-year domain to test the response of the models to changes in forcing data and surface parameters. Each participating modeling group will be encouraged to participate in some or all of the proposed studies. Several likely studies are described here — more may be conceived during the course of the project.

2.3.1 Sensitivity to atmospheric forcing

2.3.1.1 *Precipitation data*

Global observational data sets of quantities such as precipitation are probably superior to reanalysis estimates of precipitation, but they are far from perfect. Incomplete gauge coverage, difficulty in collecting observational data, and problems with the instruments themselves can lead to very uneven coverage and quality of observational data. Oki et al. (1999) showed that in the pilot phase of GSWP LSS runoff was systematically underestimated over high latitudes where snow is a significant contributor to annual precipitation. Motoya et al. (2002) has examined this issue in detail and provided an algorithm for gauge correction that is being applied in GSWP-2 (see Section 3.1.1.2).

Often remote sensing is employed to fill gaps and synthesize data across space. But remote sensing estimates of physical surface fluxes such as precipitation require the application of empirical algorithms to retrievals, and many such algorithms exist and are used (Adler et al. 2001), leading to variations in estimates. In addition, multiple spaceborne platforms are combined, sometimes to provide complete spatial coverage (e.g., in the case of geosynchronous satellites) or temporal coverage (e.g., polar orbiters with different timings of ascending and descending trajectories). Lastly, over long intervals like the 10-year ISLSCP Initiative II period, satellites come and go, each with different instruments that must be cross-calibrated, and few capable of performing usefully over such a long period. The instruments also drift in their calibration while in orbit.

All of these problems complicate the efforts to produce consistent long-term observational records with global coverage. As a result, there are many ways to go about producing such a global data set, each with its own strengths and weaknesses. In the ISLSCP Initiative II data set, there are at least five different interpretations of global precipitation, including the reanalysis model products. One set of sensitivity studies will compare the impact of the range of estimates of precipitation on simulation of the land surface climate, and in particular the hydrologic cycle. In some cases it may be possible to show precipitation estimates to be clearly incorrect (e.g., over a basin where observed streamflow is consistently greater than estimates of upstream rainfall). But in most cases, the range of estimates will be simply another degree of uncertainty, and its impact on the surface simulations can be appraised.

The following list describes the studies involving LSS sensitivity to uncertainties in the precipitation data. Tantamount to the precipitation sensitivity studies will be evaluation of the impact on simulated river discharge, by the same methodology described in Section 3.3.1. The baseline integration uses a hybrid precipitation product composed of NCEP/DOE reanalysis 3-hourly rainfall scaled to agree with gauge-corrected monthly means from GPCC, which are supplemented with merged data from GPCP where gauge density is low (see Section 3.1.1.2 for a complete description of the precipitation data set).

P1 *Hybrid precipitation derived using the ERA40 in place of the NCEP/DOE reanalysis*. This integration, compared to the baseline, will show the consequence of the choice of reanalysis rainfall (atmospheric model physics impact on temporal distribution of precipitation) on the simulation of the surface water balance. By the hybridization process, the monthly mean precipitation will be the same as in the baseline experiment, but the distribution of rainfall within the month, and particularly the diurnal cycle, may be quite different.

P2 *Hybrid precipitation as in the baseline simulation, but without the relaxation to GPCP (satellite-estimated) precipitation in regions of low gauge density*. This experiment will show the impact of satellite rainfall products on simulation of the terrestrial surface water balance in data-sparse regions.

P3 *Hybrid precipitation as in* **P2***, but without correction for gauge undercatch*. Comparison between this case and **P2** will show the hydrologic impact of neglecting the effect of wind on raingauge accuracy.

P4 *NCEP/DOE precipitation without hybridization.* This data set would be the pure NWP model product, without any observational data used for adjustment. Comparison to the baseline simulation would show the impact of adjustment of reanalysis precipitation to agree with observed time means on surface hydrologic simulation. Comparison to **P3** would show the impact of unadjusted gauge correction alone.

P5* *Using the experimental 0.5° GTS-based daily global precipitation product of Xie for hybridization of daily, rather than monthly, totals in the NCEP/DOE reanalysis.* Gauge undercatch adjustment would be performed as in the baseline case, but with daily wind analyses. Comparison to the baseline simulation would show the impact of potentially incorrect synoptic (sub-monthly) variability of precipitation on the simulation of the surface water balance.

* This sensitivity test may be dropped from the list, depending on interest and required effort.

2.3.1.2 *Radiation data*

The 3-hourly SRB radiation product will be used for the baseline simulations. But there exist 3-hourly downward shortwave and longwave radiation fields from each of the reanalyses. Sensitivity studies can be performed using in turn NCEP/DOE or ERA-40 estimates of downward radiation. The recommended studies are:

R1 *Using NCEP/DOE downward shortwave and longwave radiation instead of SRB*. Comparison to the baseline simulation will show the impact of the systematic errors in the reanalysis radiation on simulation of the surface energy and water balances, and provide an indication of the effect of representative downward radiation errors on the land surface in coupled land-atmosphere models.

RS *As in* **R1***, but substituting only for the downward shortwave radiation.* This sensitivity experiment will isolate the impact of shortwave radiation errors on the surface energy and water balance in LSSs.

RL *As in* **R1***, but substituting only for the downward longwave radiation.* This sensitivity experiment will isolate the impact of longwave radiation errors on the surface energy and water balance in LSSs.

The motivation for looking at the shortwave and longwave effects separately comes from the finding that the spatial and temporal structure of errors in these two terms in atmospheric general circulation models is typically very different, and may have very different effects on the land surface (Dirmeyer, 2002).

R2 *Using ERA40 downward shortwave and longwave radiation instead of SRB*. Comparison to the baseline simulation and **R1** will show the differences and similarities in the radiation errors in the reanalyses, and their impact on simulation of the surface energy and water balances.

2.3.1.3 *All meteorological data*

The final set of sensitivity studies proposed is simply the complete substitution of the NCEP/DOE based meteorological forcing data with the parallel product derived from the ERA-40 reanalysis. As with the NCEP/DOE reanalysis, the ECMWF reanalysis may also be provided as forcing data directly and in a hybridized form corrected by observations. The monthly means of the two hybrid products would be the same, but the different reanalyses would likely exhibit somewhat different characters in the diurnal cycle and synoptic variability, which could influence the simulation of the surface energy and water balances in offline LSSs. This experiment, **M1**, will be a test of how the different interpretations of the synoptic evolution of global weather impact the surface simulation of the energy and water balances. Because of the hybridization process for precipitation and temperature, the two forcing data sets will be somewhat constrained, and not as dissimilar as the original reanalyses on time scales of a month or longer.

2.3.1.4 *Implications of results*

Since reanalysis products are so widely used as a proxy for true atmospheric conditions, these sensitivity tests with forcing data have important implications for the certitude that should be applied to scientific results achieved using these data sets. Results from these experiments have broader implications for the impact of biases in any dynamical atmospheric model on land-surface simulation, particularly in an uncoupled mode. Such a comparison could provide useful feedback to the operational meteorological centers, particularly if land-surface validation reveals distinct differences in overall performance depending on the choice of reanalysis.

Results from these investigations may point to regions where increased or improved observations may have the biggest impact. Oki et al. (1999) showed a clear connection between a drop in the quality of simulated annual river discharge by LSSs in GSWP-1, and a lower threshold for rain gauge density within a river basin. Similar results maybe expected for other observable quantities.

2.3.2 Differences in parameters

There are also uncertainties in the parameters that describe the state of the land surface. Again, quality may vary spatially, as some nations have complete surveys of soil or vegetation, for instance, while others do not. Also, different analyses may reflect differing interpretations of the same basic information.

2.3.2.1 *Sensitivity to prescribed surface conditions*

Just as with meteorological forcing data, there is uncertainty in the data for surface conditions. For instance, the ISLSCP Initiative II includes global land cover classifications of IGBP-DIS from the EROS Data Center (EDC), the University of Maryland, and a MODIS-derived land cover product, which employ somewhat different divisions and definitions of vegetation types. This leads to ambiguity, and potentially different results for a LSS depending on which data set is used, and how it is mapped to the vegetation treatment within the LSS. Similar vagaries may exist for soil properties, albedo (for which there are also a half dozen versions with slightly different definitions in the ISLSCP Initiative II data set), Normalized Difference Vegetation Index (NDVI), leaf area index (LAI), and other vegetation properties.

Sensitivity to different interpretations of surface parameters in LSSs can be tested to determine the impact of uncertainty in these parameters on estimations of surface fluxes. Sensitivity test **V1** will be concerned with the global specification of vegetation type. The baseline simulations will be predicated on the IGBP vegetation. A global 1° map will be specified from one of the two other data sources, and used in place of the IGBP map. Impacts of this change on model simulations and errors will be assessed.

A second study will look at the impact of interannual variations of vegetation on the simulation of surface fluxes and state variables. In this sensitivity study, **I1**, a mean seasonal cycle of time-varying vegetation parameters (LAI, roughness length, displacement height, NDVI, greenness fraction) will be prescribed throughout the 10-year simulation. This study will elucidate the role of interannual variations on the land surface in an uncoupled mode (without atmospheric feedback). This experiment will be of particular interest to those modeling groups whose LSSs are also used in climate models, and who have an interest in land cover change studies, or interactive vegetation. Comparison of results from **I1** and baseline simulations to parallel integrations of the full climate model can isolate the role of land-atmosphere model coupling and feedbacks in the response of the land surface to phenology variability.

For LSSs that parameterize sub-grid variability in surface parameters, the ISLSCP Initiative II data sets include, in some cases, information of sub-grid variability at the 1° resolution, or explicit distributions of surface parameters at ½° or ¼° resolution. LSSs that use either a tiling approach or a statistical representation of sub-grid variability could conduct sensitivity experiments into the impact of aggregation/disaggregation on their simulations at the global scale.

2.3.2.2 *Implications for future ISLSCP efforts at data synthesis*

GSWP-2 may uncover problems in the ISLSCP Initiative II data sets, as they apply to land surface modeling. Hopefully GSWP-2 will find solutions as well. This information must be fed back to the ISLSCP data efforts for planning of future initiatives.

2.4 Remote sensing applications

One of the new thrusts for GSWP-2 is a stronger connection to applications in remote sensing. In addition to the classical attempts to validate the typical land-surface state variables using satellite retrievals, GSWP-2 also intends to expand the validation and assimilation capabilities of current LSSs. This is to be done by the development of algorithms by which LSSs can directly report brightness temperatures, like those sensed by instruments in orbit.

The principal goal of the effort in remote sensing applications is to expand validation beyond those few areas where *in situ* data are readily available. A secondary goal is to facilitate efforts to assimilate remotely sensed observations of the land surface into LSSs. Details are given below. Of course, remotely sensed surface state variables have their own inaccuracies, so in some sense this is a cross-validation exercise, where consistency between GSWP-2 model estimates and remote sensing may increase confidence in both products.

2.4.1 Prognostic brightness temperatures

The principal new remote sensing element is the direct reporting by LSSs of brightness temperatures in different radiative bands. Ideally, fairly basic and general algorithms can be developed and distributed that will allow participating LSSs to be modified without excessive duplication of developmental effort, or even to calculate brightness temperatures in a post-processing step given key model parameters and state variables. These algorithms may include options that could be implemented depending on the complexity of the scheme.

2.4.1.1 *Microwave/soil wetness*

Certain bands in the microwave spectrum are sensitive to moisture in the top few centimeters of soil. Microwave satellite instruments such as Scanning Multichannel Microwave Radiometer (SMMR) and Special Sensor Microwave/Imager (SSM/I), as well as aircraft-borne instruments deployed for limited-area field campaigns, can sense near-surface soil moisture. In our application within GSWP-2, we will pursue the more direct problem of estimating the observable brightness temperature from model soil wetness and vegetation properties. Relatively simple calculations exist to estimate microwave brightness temperatures, given surface characteristics such as the profile of near-surface soil moisture, vegetation cover, and surface topographic characteristics.

There are practical problems in this calculation resulting from different numbers and thicknesses of soil layers in the various LSSs. Microwave instruments are extremely sensitive to the soil wetness profile in the top few centimeters of soil, with little contribution from deeper layers. Most LSSs not developed for remote sensing applications have only one or perhaps two soil layers within the shallow depth range sampled by the microwave instruments. This presents a challenge — how should continuous or near-continuous profiles of soil wetness be interpolated (or between the center of the top soil layer and the surface; extrapolated) from the very coarse and discrete soil layers of each LSS? Two models may have the same near surface soil wetness, but because of differences in layer thicknesses they may yield different brightness temperature calculations. Some thought is also needed on parameterizations within the microwave emission models, particularly in terms of vegetation. Vegetation attenuates the soil wetness signal at a rate proportional to the canopy density, essentially contributing a vegetation optical depth through which the soil moisture signal must travel.

2.4.1.2 *Vegetation index/Dynamic Vegetation Models (DVM)*

Vegetation indices are a special case of shortwave remote sensing. These usually involve the ratio of brightness temperatures measured in different shortwave bands or "channels", and are used to enhance the detection of the state of vegetation. The NDVI is the most common of these indices.

Some LSSs take vegetation indices as an input parameter for determining greenness, LAI, etc. In fact, the global grids of vegetation cover fraction, greenness and LAI in the ISLSCP

Initiative II data set, which can be used as input parameters by many LSSs, are derived from remotely sensed vegetation indices.

For vegetation models that predict plant phenology, vegetation indices can be used to validate their simulation of seasonal and interannual variations in vegetation. Provided a sufficient number of the participating models in GSWP-2 have predictive vegetation components, GSWP-2 will also explore this area of remote sensing validation.

2.4.2 Validation of classical state variables

2.4.2.1 Skin temperature

Perhaps the most basic and straightforward quantity that is directly observable from space is the thermal skin temperature, or more properly surface upward longwave radiation. Of course, satellites can only measure this quantity during cloud free conditions, and atmospheric moisture and aerosols attenuate the signal. Nevertheless, most LSSs already report a radiative skin temperature that is used in the calculation of upward longwave radiation. Relatively robust retrieval algorithms exist to obtain surface skin temperature from remote sensing observations. These data can be used to validate the radiative skin temperature simulated by the LSSs. Model skin temperature is highly controlled by the near surface air temperature, which will be specified as the same for all LSSs. Nonetheless, there may be variations due to differences in the specification of heat capacity and thermal conductivity of vegetation and soils among the LSSs, as well as the simulated Bowen ratios, so identical results are not to be expected. There exist International Satellite Cloud Climatology Project (ISCCP) estimates of radiometric skin temperature from the surface that can be applied for validation over relatively cloud-free areas (less than 40% cloud cover). To represent surface radiative temperature, however, surface emissivity must be known. Emissivity is a function of soil properties, water content, and vegetation. Nevertheless, skin temperature is one of few quantities where global validation is possible.

There are complicating factors, however. The satellite instrument does not sample the entire infrared band, and different instruments have somewhat different sensitivity spectra (this is true for all types of radiative instruments). Also, most LSSs produce a single value per grid box, or a small number of values for LSSs that include sub-grid tiling. At 1° resolution, this model grid box corresponds to hundreds or thousands of individually sensed satellite pixels, each of which represents the sensor's synthesis of patterns of radiation at even finer scales that the instrument cannot resolve. How should this gap in spatial scales be bridged? What methods of aggregation are appropriate? Also, the instant of sampling of a location by a satellite may not correspond to the time that data are reported by the LSS. Some sort of windowing must be applied to assure the model and satellite data are adequately co-temporal. GSWP-2 will have to address these issues.

2.4.2.2 *Albedo*

Another fairly direct quantity that is both observed from space and reported by LSSs is the surface reflection of visible light, or albedo. As with infrared, visible radiation is usually sensed by orbital instruments in rather discrete bands, so a total shortwave albedo is not directly observed. Much of the shortwave spectrum is not truly visible but in the near-infrared. Many of the same problems that plague retrievals of infrared brightness temperatures also affect albedo — cloud contamination, clear-sky attenuation, frequency band mismatch, and spatial aggregation. In addition, many LSSs take snow-free albedo as an input parameter, not as a prognostic variable. These models may only alter the specified albedo in the presence of snow cover, or in cases of high soil moisture (soil albedo). Models that perform a two-stream calculation for shortwave radiation, such as the Simple Biosphere (SiB) family of models, do calculate surface albedo in all situations. Also, models with dynamic vegetation cover calculate greenness, LAI, and other parameters which then affect surface albedo. For these schemes, direct validation of albedo against remote sensing products could be a useful tool.

2.4.2.3 *Snow cover*

Areal snow coverage is routinely deduced from remote sensing, and historical archives for Northern Hemisphere snow cover exist on daily to monthly time intervals during the GSWP-2 period. This data can be used to assess the large-scale performance of the LSSs in simulating snow coverage and the timing and rate of snowpack loss due to melt during the spring.

The daily product comes from the US Air Force (USAF) Daily Snow Depth Analysis, available through National Oceanic and Atmospheric Administration (NOAA) (K. Mitchell) with additional quality control for climate applications (M. Fennessy, Center for Ocean-Land-Atmosphere Studies [COLA]; C. A. Schlosser, University of Maryland, Baltimore County [UMBC]). These analyses are a blend of *in situ* and empirically-based estimates of snow coverage and depth, and cover the entire GSWP-2 period. Microwave-based retrievals of snow are also available (Chang, D. Hall, J. Foster, and R. Kelly, National Aeronautics and Space Administration [NASA]/Goddard Space Flight Center [GSFC]) from SMMR up to 1987, and are currently being produced for 1987-1995 using SSM/I. Surface freeze-thaw cycles can also be detected by remote sensing, and compared to the LSS simulations.

2.4.3 Assimilation techniques

As mentioned before, some models already take remotely sensed data as input parameters (e.g., albedo or NDVI). Assimilation implies taking as input observed data that corresponds to one or more model state variables. Thus, LSSs that predict albedo within their shortwave radiation calculations could assimilate observed albedo. Already some DVMs are designed to assimilate remotely sensed vegetation indices.

Assimilation of remotely sensed surface infrared information is already being conducted (Houser et al., 1998) and could be expanded to other LSSs. There is tremendous interest in the assimilation of soil moisture information (e.g., Salvucci, 1997; Margulis and Entekhabi, 2001; Crow and Wood, 2002) as soil moisture is the most important land surface state variable for seasonal-interannual climate variations throughout most of the globe (Dirmeyer and Shukla, 1993). In particular, since remote sensing cannot penetrate the surface layers to detect soil moisture in the vadose zone, a combination of modeling and remote sensing provides the best solution for diagnosing soil moisture.

Methods of assimilation can be simple (e.g., relaxation with a fixed damping time scale) or complex (variational or Kalman-filtering methods). It is not the purview of GSWP-2 to explore the assimilation techniques themselves. Rather, it is to provide a starting point for independent efforts to probe this area further (e.g., LDAS; http://ldas.gsfc.nasa.gov/).

2.5 Other science efforts

2.5.1 Comparison to simple and intermediate models

Can simple or intermediate models reproduce the behavior of more complex LSSs? The aim of this investigation is to determine how much of the “signal” in complex LSSs, integrated at time steps of one hour or less, can be reproduced by simple or intermediate models.

The inclusion of the so-called “Bucket” model in previous PILPS and GSWP simulations is a form of this comparison, although historically the Bucket has been treated like a benchmark that other LSSs should be able to outperform. There are other approaches to approximating the behavior of complex LSSs with systems that have a much more limited number of free parameters. Koster and Milly (1997) showed for a particular PILPS experiment that the fundamental behavior of the surface water balance of all of the participating LSSs on monthly time scales could be described almost entirely by a two-parameter model. Koster et al. (2001) showed that for coupled land-atmosphere models, the behavior of the partitioning of precipitation between evapotranspiration and runoff on annual time scales conformed to Budyko’s (1974) expectation, based on knowledge of only net radiation and total precipitation.

Several LSSs of intermediate complexity have been developed to study specific phenomena such as tropical deforestation (e.g., Eltahir and Bras, 1993; Zeng 1998; Neelin and Zeng, 2000).

How well do simple and intermediate models perform in the GSWP framework, when compared to the complex LSS? This comparison will quantify this both in terms of skill (where validation is available), and complexity of response. For example, a simple model with no diurnal cycle, that must be driven by daily mean data, cannot reproduce variations on those short time scales, whereas a complex LSS can. This is an extreme example, but there may be other modes of variability (both in time and space) where there are discernable differences. Perhaps more significant will be the identification and explanation of the situations where there is no real difference in performance, particularly when the complexity built into one or more LSSs is justified by the need to better simulate a given phenomenon (e.g., snow cover). The simple and intermediate models are thus diagnostic tools for understanding the working of both processes in the physical climate, and the more complex LSSs.

Another possibility is to revisit the study of Koster et al. (1999), which compared GSWP-1 runoff rates against those produced with the simple Budyko model. In that study, the complexity of the LSSs did not lead to a systematically better simulation of the annual water balance. Perhaps LSSs have since improved.

2.5.2 Uniqueness

The question may be asked, “Do we need so many different LSSs?” This is a valid question, particularly if it can be shown that there is not a significant variation among LSSs in their performance. Another diagnostic project in GSWP-2 will be to see how well each LSS can “predict” the behavior of the others.

This investigation of uniqueness among LSSs amounts to an exercise in the construction of transfer functions among LSSs. For each pair of LSSs, a training period is defined (e.g., the first 5 years of the baseline GSWP-2 period), over which state variables or fluxes from one model are regressed upon those of another at each grid point to produce transfer coefficients. Then, for the test period (the last 5 years), the results of the predictor model are run through the transfer functions to produce a simulation of the predictand model — a forecast of one LSS by another LSS. This can be performed for some or all of the participating models, on one or more time scales (e.g., for annual, monthly, or daily) using the output data reported to the Inter-Comparison Center (ICC).

The null hypothesis is that any LSS can perfectly predict any other. In the absence of differences in forcing data, variations from perfect skill must be due to differences among the models. The patterns of differences (in space, time, and within the matrix of participating LSSs) may reveal something about capabilities of the models, reconfirm their original design applications, and perhaps reveal something new about the family of LSSs. If two or more LSSs are found to be essentially undifferentiatable, that would also be very revealing and raise the issue of redundancy in the land surface modeling community. Key to isolating model differences will be determining whether the models specified their surface parameters in a novel fashion. This is conveyed by the reporting of ancillary information with the submission of model output from each LSS integration (see Section 3.1.4.5).

2.5.3 Global Water and Energy Cycles

One of the main research objectives of GEWEX is to "determine the hydrological cycle and energy fluxes by means of global measurements of atmospheric and surface properties." To first order, the "rate" of the global water cycle can be quantified by the global fluxes of precipitation and evaporation. On an annual basis, these fluxes should nearly balance, as the capacity for the atmosphere to store the residual difference between precipitation and evaporation is quite small. Currently, there exists no global-wide capacity to directly measure the fluxes of water and energy over the continental surfaces, and therefore we must rely on the highest-quality estimates based on model simulations. The GSWP-2 experiment will provide the most representative collection to date of continental fluxes of water and energy from "state-of-the-art" land models used by the international climate research community.

Therefore, based on the results from the analyses discussed in Sections 2.2.4 and 2.5.2, the consensus (and scatter) of the GSWP-2 model simulations will be used to supplement a global water cycle synthesis. The GSWP-2 outputs will be combined with global precipitation products (e.g., the GPCP data of Huffman et al., 1997) and ocean flux estimates (e.g., Chou et al., 1997) to assess our scientific accounting of the global water cycle. Using this synthesis, not only can the relative roles of the land and ocean in the global water cycle be quantified, but the 10-year outputs (1986-1995) of the GSWP-2 (which overlap with the global precipitation and ocean flux data) will allow for interannual variations (and trends) to be diagnosed and checked for consistency against the global variations of precipitation and ocean flux estimates. The residual of these global precipitation and evaporation fluxes will also be cross-verified against the GEWEX Global Water Vapor Project (GVaP) data to provide a further assessment of consistency and confidence in our global observations of key hydrologic states and fluxes. In a similar manner, the output of radiation and heat fluxes from the GSWP 2 model suite can be used to further supplement (and update) our current depiction of the global energy cycle. Leveraging off of GEWEX radiation projects, such as the International Satellite Cloud

Climatology Project (ISCCP), surface, atmospheric and top-of-the-atmosphere radiation and heat fluxes can be combined to assess and quantify variations and trends in the global energy cycle, and whether consistent linkages exist between these water and energy cycle assessments (i.e., do variations/trends in observed cloud cover show a consistent association with our observed/estimated water and energy budgets?).

3.0 IMPLEMENTATION PLAN

In order to conduct the proposed scientific program of GSWP-2, a well-defined sequence of actions must be undertaken. Figure 4 shows a flowchart for the actions, and Figure 5 shows the proposed timeline. All discussion in this section centers around these diagrams, and they should be used for reference throughout. Herewith, a practical discussion of the conduct of the GSWP-2 experiment is presented.

3.1 Production of the GSWP-2 data sets

The starting point for GSWP-2 is the ISLSCP Initiative II data set. There are two fundamental categories of ISLSCP data that are needed by the LSSs that participate in GSWP-2: land surface parameters and meteorological forcing. Land surface parameters include vegetation classification maps, vegetation properties (which may vary from month-to-month), and soil properties. There are two parallel versions of the meteorological data in the ISLSCP data set: the National Centers for Environmental Prediction/Department of Energy (NCEP/DOE) and European Centre for Medium-range Weather Forecasts (ECMWF) (ERA-40) reanalyses. The meteorological data are provided at a 3-hourly time step for a period of 13½ years (1982-1995), although the complete set from ECMWF will not be available until early 2003.

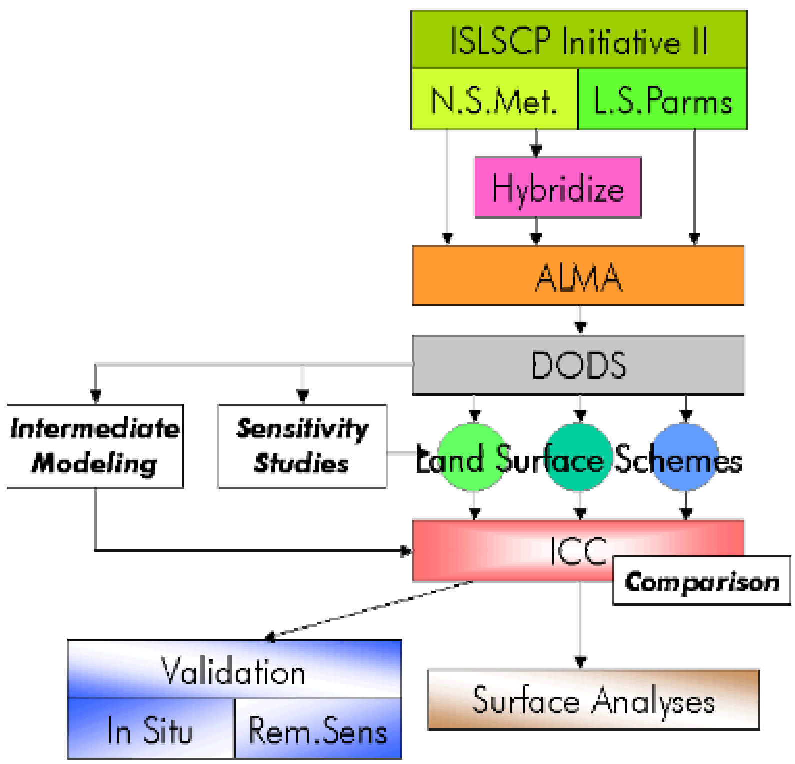

Figure 4. Implementation flowchart for GSWP-2.

3.1.1 Input fields

3.1.1.1 *Land surface data*

Land surface parameters for participating LSSs will be specified from the ISLSCP Initiative II data set. The ISLSCP Initiative II data set includes land cover data from the EROS

Data Center (EDC) in several forms, including vegetation types for Biosphere-Atmosphere Transfer Scheme (BATS) and Simple Biosphere Model (SiB). In addition, information on the fractional distribution of vegetation for each type in each 1° grid box is provided, for LSSs that have a mosaic or tile approach to sub-grid variations in land cover type. Time-varying information on biophysical parameters (e.g., Leaf Area Index [LAI], Normalized Difference Vegetation Index [NDVI], greenness fraction) are also included. These monthly fields should be interpolated linearly from mid-month to mid-month. Data spanning 1982-1998 will be included for these fields. Each participating LSS should use those data sets that are appropriate to its formulation. Dynamic vegetation models (DVMs) will need to specify initial plant functional types (PFTs) rather than the vegetation types described here. It will be the responsibility of DVM groups to derive their own set of PFTs consistent with the baseline GSWP vegetation distribution. Another vegetation distribution data from University of Maryland, as well as a MODIS-derived land cover product, may be used for sensitivity studies. If necessary data for a LSS is not provided, the modeler should work out an alternative solution with GSWP and report the choices made when submitting his output (see Section 3.1.4.5).

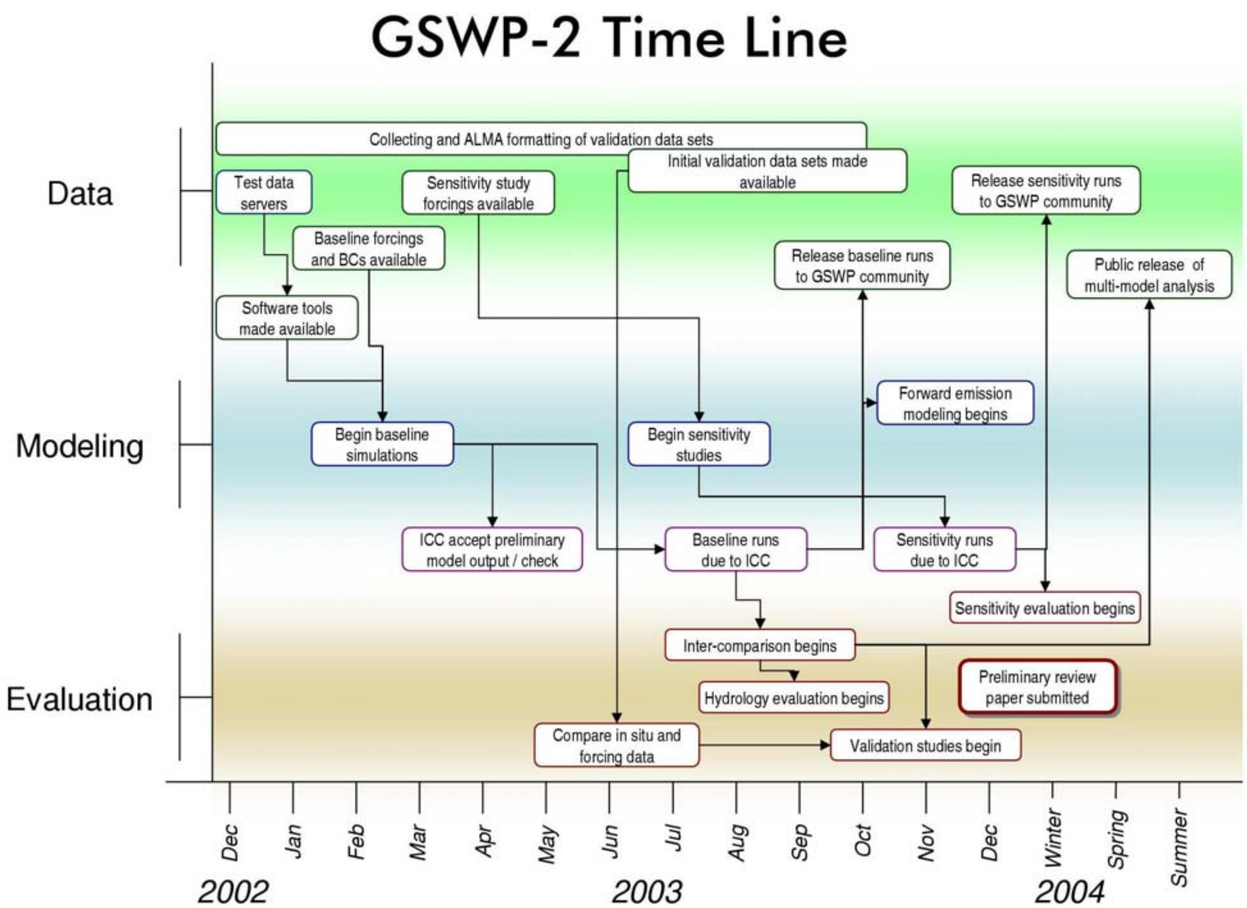

Figure 5. Timeline for GSWP-2

ISLSCP Initiative II soils data come from the International Geosphere-Biosphere Programme Data Information System (IGBP-DIS) soils CD-ROM. LSSs use a variety of methods to specify soil parameters; GSWP-2 will try to accommodate them. Global 1° maps of sand, clay and silt fractions are provided for models that derive soil hydrologic and thermal properties from soil structure. The fractions add to unity. There also exist data on the fraction of organic content in the soil. Models that use this information need to rescale the soil fractions, as the sum of organic plus mineral content will exceed unity. In the original IGBP soils data provided by ISLSCP Initiative II, there exists points with missing data over some desert areas, rocky mountainous regions, and at points with deep icepack. In these areas, reasonable values have been interpolated by GSWP-2 from surrounding points: the missing values of fractions of clay, sand, silt, and organic (548 points, mainly over Greenland) are also filled by the averaged value from surrounding land points; the missing values of W_fieldcap, W_wilt, W_sat, and W_sat_hydc (548 points) are filled by averaging the values from surrounding land points; and the elevation and slope data are aggregated from 0.5° to 1°. Note that the values of field capacity over Greenland for soil texture class 2 are unusually high. Soil Depth, Albedo_vi, and Albedo_ir are obtained from ISLSCP-I by using ISLSCP-II land-sea mask to remap them. Method CEA84 was used to determined soil hydrological parameters from Cosby et al., 1984.

Some LSSs use a soil texture classification scheme to specify soil parameters based on texture classes from the USDA soil triangle. GSWP-2 provides a global map of texture classes with a 12-class soil texture table (Table 2). Missing values of the soil texture class are determined from the calculated values of percent of sand and clay, as described above. In the

Table 2. Soil properties as a function of texture class.
Cosby values for silt are estimated, as they were not provided in the original RhôneAGG data set.

Texture Class		*USDA* Sand	Silt	Clay	*Cosby (RhoneAGG) USDA* Wfc	Wwilt	Wsat	b	PHIsat	Ksat
1	**Sand**	92%	5%	3%	0.132	0.033	0.373	3.30	-0.05	2.45E-05
2	**Loamy Sand**	82%	12%	6%	0.156	0.051	0.386	3.80	-0.07	1.75E-05
3	**Sandy Loam**	58%	32%	10%	0.196	0.086	0.419	4.34	-0.16	8.35E-06
4	**Loam**	17%	70%	13%	0.270	0.169	0.476	5.25	-0.65	2.36E-06
5	**Silt Loam**	10%	85%	5%	*0.361*	*0.045*	*0.471*	*3.63*	*-0.84*	*1.10E-06*
6	**Silt**	43%	39%	18%	0.250	0.148	0.437	5.96	-0.24	4.66E-06
7	**Sandy Clay Loam**	58%	15%	27%	0.253	0.156	0.412	7.32	-0.12	6.31E-06
8	**Clay Loam**	10%	56%	34%	0.334	0.249	0.478	8.41	-0.63	1.44E-06
9	**Silty Clay Loam**	32%	34%	34%	0.301	0.211	0.447	8.34	-0.28	2.72E-06
10	**Sandy Clay**	52%	6%	42%	0.288	0.199	0.415	9.70	-0.12	4.25E-06
11	**Silty Clay**	6%	47%	47%	0.363	0.286	0.478	10.78	-0.58	1.02E-06
12	**Clay**	22%	20%	58%	0.353	0.276	0.450	12.93	-0.27	1.33E-06

ISLSCP Initiative II soil map, there are no points categorized as silt. Table 2 closely corresponds to the RhôneAGG table, and the approach is similar to the Zobler classifications used in GSWP-1.

ISLSCP Initiative II also provides maps of soil properties calculated from pedon attributes on the IGBP CD-ROM. These derived parameters may not be consistent with the assumptions built into each LSS (e.g., wilting point and field capacity are derived using the van Genuchten (1980) relationship, which would not be appropriate to use in a model based on the Clapp and Hornberger (1978) approach). GSWP-2 will provide these gridded data, but we caution modelers to carefully check the IGBP documentation (which will be provided on the GSWP-2 website) before using these fields. Using a look-up table would be safer.

Topographic data comes from the EDC, derived from the ETOPO30 and Hydro1K data sets. Fields include mean elevation, surface slope, and sub-grid elevation statistics that may be useful for land surface modeling.

3.1.1.2 *Atmospheric forcing data*

Atmospheric forcing data will use information from global gridded observational data sets when possible. In some cases, no adequate global observational data exist, so a pure model reanalysis product will be used. In most cases, observational data are available globally, but not at the high time resolution needed to resolve the diurnal cycle, as is necessary to force LSSs. In these instances, observational data will be combined with data from model-based reanalysis products.

The 3-hourly meteorological data provided in ISLSCP Initiative II are pure reanalysis products, and have not been amended by “hybridizing” with observational data, as was done for ISLSCP Initiative I. GSWP-2 is undertaking this step, using the observationally-based precipitation, surface radiation, and near-surface meteorology data also included in ISLSCP Initiative II. The hybridization process has been developed, tested, applied, and documented for the NCEP/NCAR reanalysis data (Dirmeyer and Tan, 2001), and can be applied to the other reanalysis products. The hybridized products will be the primary forcing data sets for GSWP-2; the original ISLSCP products will be used in the modeling sensitivity experiments. Because the full 10-year span of the NCEP/DOE reanalysis data will be available before ECMWF’s reanalysis, we choose the hybridized version of the NCEP/DOE data set for the baseline simulation by the LSSs. The 3-hourly data from the NCEP/DOE reanalyses have been processed for inclusion in the ISLSCP Initiative II data set by COLA from hourly data provided by W. Ebisuzaki at the NCEP Environmental Modeling Center (EMC).

Precipitation

The precipitation product for the baseline simulations in GSWP-2 will be a hybrid product. In the hybridization process for precipitation and radiation, the errors are removed via a multiplicative scaling factor that is based on the ratio of observed monthly rainfall to reanalysis estimates, rather than by subtraction of the error:

$$[P]_{Y,M,D,T} = \frac{[P_{OBS}]_M}{[P_{NCEP}]_M}[P_{NCEP}]_{Y,M,D,T}$$

To get the adjusted forcing data for precipitation, for instance, the value at a grid box of one of the reanalysis precipitation terms (total or convective) at a given year, month, day and 3-hour time interval $[P_{NCEP}]_{Y,M,D,T}$ is scaled by the ratio of the monthly mean observed precipitation to the corresponding mean value from the reanalysis for that month. This approach avoids problems of negative values in positive definite quantities with frequent zeroes, such as precipitation. It provides the best attainable improvement in the reanalysis estimates given the lack a long-term sub-monthly global observationally-based data set.

No attempt will be made to adjust the monthly storm frequency (Liston et al., 1993), as was done for the 6-hourly precipitation estimates in the ISLSCP Initiative I data set (Mitchell and Lin, 1994). Nor is any attempt made to adjust the diurnal cycle, which is known to be in error over some regions. The main constraint is that the monthly mean precipitation should agree with the observation data, with some small differences introduced as a result of spatial interpolation. This preservation of observed monthly means is also in effect for all other hybridized variables.

Several observational precipitation data sets will be available from ISLSCP Initiative II. The Climate Research Unit (CRU) (New et al., 1999; 2000) data set from University of East Anglia is a high-resolution (0.5°) gauge-only product, but relies on only operational data sources, does not correct for gauge undercatch, and relaxes the data to a mean annual cycle climatology when *in situ* data are scarce. The Global Precipitation Climatology Centre (GPCC) (Rudolf et al., 1994) maintains a gridded gauge analysis that contains more stations than the CRU analysis; these data are also provided to ISLSCP on a 0.5° grid. They do provide a separate monthly correction factor to adjust for wind-caused gauge undercatch. The Global Precipitation Climatology Project (GPCP) (Huffman et al., 1997) also provide monthly analyses, that blend corrected gauge and satellite estimates. This data set may prove to be the best for interannually-varying precipitation data, although it has the lowest native spatial resolution (2.5°), although a 1° version is provided to ISLSCP. GSWP-2 will use GPCC gauge data for the baseline period, and CRU for the spin-up period, applying the gauge correction of Motoya et al. (2002) from source code supplied by the author. Where the gauge density is low (Figure 6), the GPCP product is blended in. In this way, gauge correction can be applied at a higher spatial resolution, while maintaining the benefit of satellite data where there are no gauges. The next step is hybridization with the reanalysis rainfall estimates, as described above, to produce a 3-hourly precipitation product from GPCC, GPCP and CRU data, which can then be combined and blended to produce the final product.

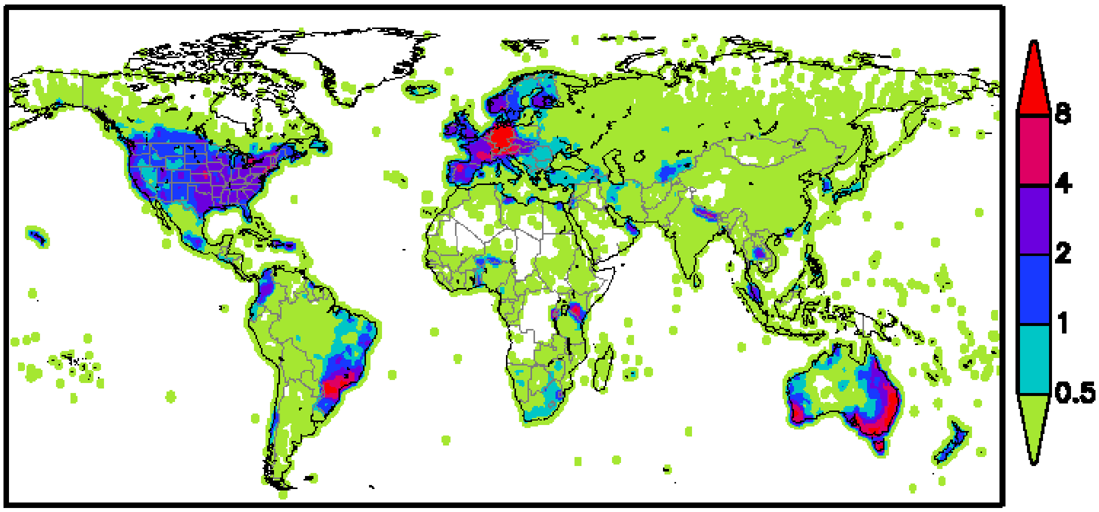

Figure 6. Example of gauge density for CRU stations for January 1986. Unshaded areas have no stations within 2° of grid box.

The process is as follows. Aggregated monthly GPCC, or CRU data are calculated, transforming from 0.5° to 1° (the GPCP data provided by ISLSCP Initiative II is available at 1°). Due to wind-caused gauge undercatch for precipitation, the Motoya et al. (2002) wind correction is applied to the unadjusted GPCC or CRU data. First, an "uncorrected gauge" for the reanalysis precipitation is estimated based on the catch ratio (C_R) correction factor calculated by the Motoya algorithm using the daily mean reanalysis wind:

$$P_{NCEP\_Gauge} = \frac{P_{NCEP}}{C_R}$$

This is necessary because the precipitation reported by the analysis model is unaffected by wind, so the corresponding undercatch error must be introduced into the model estimate before adjustment. This step ensures that the final wind-corrected precipitation estimates maintain the same relative storm-to-storm totals as the NCEP/DOE reanalysis.

Secondly, we hybridize this "NCEP gauge" with GPCC or CRU gauge data as indicated in the equation at the beginning of this subsection on precipitation, so that the monthly total agrees with the uncorrected gauge data. Then we reapply Motoya's wind correction to the hybrid data:

$$P_{Wind\_corrected} = C_R P_{Hybrid\_uncorrected}$$

In regions of low gauge density, this result is combined with the hybridized 3-hourly version of the GPCP data:

$$P_{GSWP2} = aP_{Wind\_corrected} + (1-a)P_{GPCP}$$

where

a = 1, GPCC or CRU gauge density ≥ 2

a = 0.5, GPCC or CRU gauge density = 1

a = 0, GPCC or CRU gauge density = 0

The final step is the separation of the precipitation components into rainfall and snowfall, convective and large-scale. The NCEP/DOE reanalysis reports total precipitation rate, as well as a snowfall rate that is diagnosed at each model time step from the 850 hPa air temperature (or lowest model air layer temperature if the surface pressure at a given grid point is lower than 850 hPa). If this temperature is equal to or less than 0.0C, then snowfall is designated; otherwise, rainfall is chosen (K. Mitchell, personal commmunication). Since the snowfall criterion is based on the atmospheric model state aloft, and not surface conditions, no re-estimation of snowfall will be conducted based on the hybrid near surface air temperature. Rainfall is assumed to be total precipitation minus snowfall.

The NCEP/DOE reanalysis also reports a convective precipitation rate. For GSWP-2, to conform to ALMA standards, a convective rainfall rate is given where:

$$Rain_{Conv} = \frac{P_{Conv}}{P_{Total}}(P_{Total} - Snow)$$

Large-scale rain would be $Rain$ - $Rain_{Conv}$, and if necessary, convective snowfall can be estimated assuming the same ratio as for total precipitation in the equation above.

Some decisions had to be made where inconsistencies were apparent. For instance, in the bi-linear interpolation of the precipitation data from the reanalysis grid to the ISLSCP grid, there were instances near coastlines where the interpolated snowfall rate exceeded the total precipitation rate. In these cases, the rainfall rate was set to zero. Also, during hybridization, there were instances where one but not both of the observed monthly precipitation or the reanalysis precipitation was equal to zero. In these situations, it was assumed that precipitation for the month was zero.

Temperature

There is also more than one choice for near-surface temperature data, although only the CRU data set from University of East Anglia will be included in ISLSCP Initiative II. The monthly CRU temperature data are calculated as anomalies from a 12-month climatology that is derived relative to a fixed elevation model at 0.5°. In order to make a consistent adjustment of near-surface air temperature to the ISLSCP grid, the CRU temperatures will be corrected for the altitude difference between the CRU grid and the ISLSCP Initiative II mean altitude derived from the GTOPO30 data set (see Section 3.1.1.1). First, the CRU elevation data are recreated by aggregating the GTOPO5 elevation data from 5' to 0.5°. The monthly CRU temperature data are then corrected to the ISLSCP elevation:

$$T_1 = T_{CRU} - \frac{6.5}{1000}(Z_{ISLSCP2} - Z_{CRU})$$

The monthly CRU data are then aggregated from 0.5° to 1°, and hybridized with the 3-hourly NCEP/DOE reanalyses 2-meter air temperature data by correcting the differences of monthly diurnal range and mean:

$$T_{air} = \delta(T_{NCEP} - \overline{T_{NCEP}}) + T_1$$

where:

$$\delta = \frac{D_{CRU}}{D_{NCEP}}$$

constrained so that:

$$0.5 \le \delta \le 2.0$$

D is the monthly mean diurnal range of the temperature, which is reported for the CRU data, and has been calculated from the original hourly data from the NCEP/DOE reanalyses. Limits are placed on the diurnal scaling factor δ to prevent unreasonable extreme temperatures.

Surface Pressure

An altitude correction is applied to the surface pressure data, to adjust from the reanalysis model grid elevations to the ISLSCP Initiative II mean altitude:

$$P_{S\_Corr} = P_{S\_NCEP} e^{\frac{-g}{R\overline{T}}(Z_{ISLSCP2} - Z_{NCEP})}$$

Z_{NCEP} and $Z_{ISLSCP2}$ are the grid box mean altitudes for the reanalysis and ISLSCP Initiative II respectively, g is the acceleration due to gravity, and R is the gas constant. $\overline{T}$ is the mean temperature between the two altitudes, calculated using the same lapse rate used to adjust temperature:

$$\overline{T} = T_{air} - \frac{1}{2}\left(\frac{6.5}{1000}(Z_{ISLSCP2} - Z_{NCEP})\right)$$

Specific Humidity

These adjustments to temperature also affect the estimated saturation specific humidity. Thus, it is also necessary to adjust the estimates of near surface specific humidity from the reanalysis to avoid incidents of super-saturation. This is done by assuming the same relative humidity before and after the temperature correction, and then adjusting the specific humidity accordingly to agree with the adjusted temperature.

Radiation

No hybridization will be performed on the radiation data for the 1986-1995 period, because three-hourly Surface Radiation Budget (SRB) data produced at NASA/Langley Research Center will be available through ISLSCP Initiative II. In the SRB data set there are some "undetermined" points for high latitude bands where the solar zenith angle approaches 90E. These have been arbitrarily set equal to a small value: 10 W m^{-2}.

For the spin-up period, a hybrid product will have to be produced for the surface downward shortwave and longwave radiation. To adjust downward radiation, the diurnal cycle is particularly important. Studies by Dirmeyer and Tan (2001) have shown that the systematic errors in the NCEP/National Center for Atmospheric Research (NCAR) reanalysis is very systematic from year to year, but varies substantially across both the seasonal and diurnal cycles. We will create a hybrid radiation forcing data set from reanalysis estimates by removing the climatological monthly mean diurnal cycle systematic errors calculated from the SRB and reanalysis data over the 1986-1995 period:

$$[R]_{Y,M,D,T} = \frac{[R_{SRB}]_{Y,M,T}}{[R_{NCEP}]_{Y,M,T}}[R_{NCEP}]_{Y,M,D,T}$$

As with precipitation, a multiplicative scaling is used to adjust the reanalysis.

Wind

The reanalysis wind products will be used as is, with the 10 meter wind speed provided in the forcing data set.

3.1.1.3 *ALMA conventions*

All of the gridded surface data to be used by GSWP-2 will be converted to the ALMA data convention, including compression by gathering to reduce the data set sizes by removing ocean and land-ice points. Some extensions to ALMA are recommended, especially concerning the formats for land surface properties, to create an updated ALMA version for GSWP-2.

Tables 3 and 4 shows the ALMA conventions for the land surface parameters that are being supplied to the modeling groups. These include vegetation, soils, and topographic information at 1° resolution for all land points excluding Antarctica, essentially cutting off the

Table 3. Soil parameter data.

Name	Description	Units	Range	Source	Time scale
SoilClass	Soil texture class	-	Max = 12 Min = 0	ISLSCP-II	fixed
SoilDepth	Depth of active soil column	m	Max = 30.0 Min = 0.0	ISLSCP-I	fixed
Clay	Clay fraction	-	Max = 1.0 Min = 0.0	ISLSCP-II	fixed
Sand	Sand fraction	-	Max = 1.0 Min = 0.0	ISLSCP-II	fixed
Silt	Silt fraction	-	Max = 1.0 Min = 0.0	ISLSCP-II	fixed
Organic	Organic fraction	-	Max = 1.0 Min = 0.0	ISLSCP-II	fixed
Elevation	Mean grid elevation	m	Max = 9000.0 Min = -400.0	ISLSCP-II	fixed
Slope	Mean slope	$m\ m^{-1}$	Max = 1.0 Min = 0.01	ISLSCP-II	fixed
CTI	Compound topographic index	-	Min = 0.0 Max = 13	ISLSCP-II	fixed
W_fieldcap	Field capacity	$m^3\ m^{-3}$	Max = 1.0 Min = 0.0	ISLSCP-II	fixed
W_wilt	Wilting point	$m^3\ m^{-3}$	Max = 1.0 Min = 0.0	ISLSCP-II	fixed
W_sat	Saturated water content	$m^3\ m^{-3}$	Max = 1.0 Min = 0.0	ISLSCP-II CEA84	fixed
W_bpower	B exponent	-	Max = 15.0 Min = 1.0	CEA84	fixed
W_sat_hydc	Saturated hydraulic conductivity	$m\ s^{-1}$	Max = 0.0001 Min = 0.0	ISLSCP-II CEA84	fixed
W_sat_matp	Saturated matric potential	m	Max = -0.0001 Min = -3.0	CEA84	fixed
Albedo_vi	Visible albedo of soil (snow free)	-	Max = 1.0 Min = 0.0	ISLSCP-I	fixed
Albedo_ir	Near-infrared albedo of soil (snow free)	-	Max = 1.0 Min = 0.0	ISLSCP-I	fixed

global grid at 60°S. Time-varying vegetation parameters are provided as monthly values. The IGBP vegetation type data are available in three different styles; IGBP categories, SiB, and BATS. Table 5 lists the classifications for each variety. Table 6 is a similar listing for the near surface meteorological forcing fields, all of which are provided on a 3-hour time interval synchronized with Coordinated Universal Time (UTC; also known as Greenwich Mean Time, Zulu, or "Z"). A description of the ALMA standards is available at http://www.lmd.jussieu.fr/ALMA/.

Table 4. Vegetation parameter data.

Name	Description	Units	Range	Source	Time scale	Period
VegClass	Vegetation class	-	Max = 21 Min = 0	IGBP, U. Maryland	fixed	-
LAI	Leaf area index	$m^2\ m^{-2}$	Max = 8.08 Min = 0.0	UK Univ. of Wales	monthly	1982-1995
vegFrac	Fraction of vegetation cover	-	Max = 1.0 Min = 0.0	UK	fixed	-
grnFrac	Greenness fraction *(green LAI / Total LAI)*	-	Max = 1.0 Min = 0.0	Calculated from Univ. of Wales data	monthly	1982-1995
classFrac	Fraction of each VegClass *(not in ALMA variable list)*	-	Max = 1.0 Min = 0.0	IGBP; U. Maryland	fixed	-
NDVI	Normalized difference vegetation index	-	Max = 1.0 Min = 0.0	UK	monthly	1982-1995
FPAR	Fraction of photosynthetically active radiation	-	Max = 1.0 Min = 0.0	UK	monthly	1982-1995
Z0Surf	Roughness length	m	Max = 10.0 Min = 0.0	UK	monthly	1982-1995
DisplH	Zero plane displacement height	m	Max = 50.0 Min = 0.0	UK	monthly	1982-1995
Albedo	Snow-free albedo	-	Max = 1.0 Min = 0.0	CSU	monthly	1982-1995
RootDepth	Root depth *(mean 50% and 95% ecosystem rooting depth)*	m	Max = 30.0 Min = 0.1	ISLSCP-II	fixed	-
Rs_min	Minimum stomatal resistance	$s\ m^{-1}$	Max = 1000.0 Min = 10.0	Look-up table	fixed	-

3.1.1.4 *Serving of data*

The final input data sets for the participating GSWP-2 LSSs will be served over the Internet via one or more Distributed Oceanographic Data System (DODS) servers. DODS provides a protocol for remotely accessing and subsetting large data sets (see: http://www.unidata.ucar.edu/packages/dods/). DODS-enabled clients exist via the NetCDF protocol for FORTRAN and C, and a growing number of popular graphics and analysis packages are DODS-capable. In a nutshell, the “open” statement, which typically accesses a file on local disk via its pathname, instead accesses the DODS server via its URL using HTTP addressing. DODS includes metadata, so that served data sets are self-describing. DODS reduces strains on local disk resources, and, when properly used, is not taxing on the network or the server hardware. In order to reduce trans-oceanic networking delays, we anticipate serving the data from multiple DODS servers, operating in mirror mode, in North America (COLA, USA), Asia (University of Tokyo, Japan), and Europe (LMD, France). If this approach is intractable to one or more modelers, we will consider alternate solutions, probably involving shipping of data on magnetic media.

Assistance will also be provided to the modelers to access the forcing and parameter data sets on the DODS servers via the ALMA Software Bazaar and the information systems

Table 5. Vegetation categories for the IGBP-derived land cover types.

	IGBP Land Cover		Simple Biosphere (SiB) Model Land Cover		Biosphere Atmosphere Transfer Scheme (BATS) Land Cover
1	Evergreen Needleleaf Forest	1	Evergreen Broadleaf Trees	1	Crops, Mixed Farming
2	Evergreen Broadleaf Forest	2	Broadleaf Deciduous Trees	2	Short Grass
3	Deciduous Needleleaf Forest	3	Deciduous and Evergreen Trees	3	Evergreen Needleleaf Trees
4	Deciduous Broadleaf Forest	4	Evergreen Needleleaf Trees	4	Deciduous Needleleaf Tree
5	Mixed Forest	5	Deciduous Needleleaf Trees	5	Deciduous Broadleaf Trees
6	Closed Shrublands	6	Ground Cover with Trees and Shrubs	6	Evergreen Broadleaf Trees
7	Open Shrublands	7	Groundcover Only	7	Tall Grass
8	Woody Savannas	8	Broadleaf Shrubs with Perennial Ground Cover	8	Desert
9	Savannas	9	Broadleaf Shrubs with Bare Soil	9	Tundra
10	Grasslands	10	Groundcover with Dwarf Trees and Shrubs	10	Irrigated Crops
11	Permanent Wetlands	11	Bare Soil	11	Semi-desert
12	Croplands	12	Agriculture or C3 Grassland	12	Ice Caps and Glaciers
13	Urban and Built-Up	13	Persistent Wetland	13	Bogs and Marshes
14	Cropland & Natural Vegetation	14	Water	14	Inland Water
15	Snow and Ice	15	Ice Cap and Glacier	15	Ocean
16	Barren or Sparsely Vegetated	16	Missing Data	16	Evergreen Shrubs
17	Water Bodies			17	Deciduous Shrubs
18	Missing Data			18	Mixed Forest
				19	Forest/Field Mosaic
				20	Water and Land Mixtures
				21	Missing Data

support at the DODS server sites. COLA has tested the capability of running LSS integrations driven by data served over the Internet via DODS in an environment of typical network connection speeds. Accessing the boundary conditions and forcing data from a DODS server across the local network, increased run time of the SSiB LSS by 15% relative accessing the data on a local NFS-mounted disk. Accessing the data across the Internet from an identical DODS server at NCAR in Boulder, Colorado added another 8% to the wall-clock execution time.

The GrADS-DODS server (GDS) software allows for dual-mode access to compressed NetCDF data in the CF convention (compression by gathering) — either as the original one-dimensional spatial vector of data (for access by LSSs), or as a repopulated grid (for easy display purposes).

Table 6. Meteorological forcing data (July 1982 - December 1995).

Name	**Description**	**Units**	**Range**	**Source**	**Time Scale**
Tair	Near surface air temperature at 2m	K	Max = 350 Min = 190	NCEP, CRU	3 hourly
Qair	Near surface specific humidity at 2m	kg kg^{-1}	Max = 0.07 Min = 0	NCEP, CRU	3 hourly
Wind	Near surface wind speed at 10m	m s^{-1}	Max = 75 Min = 0	NCEP	3 hourly
SWdown	Surface incident shortwave radiation	W m^{-2}	Max = 1360 Min = 0	SRB (with NCEP for spin-up period)	3 hourly
LWdown	Surface incident longwave radiation	W m^{-2}	Max = 750 Min = 0	As above	3 hourly
Psurf	Surface pressure	Pa	Max = 113000 Min = 50000	NCEP, EDC	3 hourly
Rainf	Rainfall rate	kg m^{-2} s^{-1}	Max = 0.03 Min = 0	NCEP, GPCC GPCP (and CRU for spin-up period)	3 hourly
Rainf_C	Convective rainfall rate	kg m^{-2} s^{-1}	Max = 0.03 Min = 0	As above	3 hourly
Snowf	Snowfall rate	kg m^{-2} s^{-1}	Max = 0.003 Min = 0	As above	3 hourly

3.1.2 Initial conditions

Once the DODS servers are populated and the input data documented, modeling groups can begin their simulations. An issue that must be resolved is the method of spin-up that the modelers should follow. GSWP-1 provided a starting point for land-surface state variables (soil moisture at 75% of saturation, no snow cover, specified soil temperature), and asked modelers to loop through the first year (1987) until some degree of "stability" in soil moisture was attained. This method can unduly amplify a year's anomalies at the beginning of the free integration period. A balance needs to be found between having the maximum number of useable years in the simulation, and having an adequate spin-up period. For the NCEP/DOE reanalysis, the option exists to use data from earlier years for the spin-up. Thus, spin-up will be performed using data beginning 0300UTC 1 July 1982. LSS integrations will loop through the first 12 months of forcing data until the modeler is satisfied that soil moisture has spun up and sufficiently equilibrated. A lesson from the GSWP pilot project was that this spin-up process overly amplifies the impact of climate anomalies from that year on the land surface state variables. Therefore, the models will then proceed with their integrations forward from July 1983 – December 1985 so as to converge to a realistic "land climate" at the start of the evaluation period. The 10-year baseline integration, which will be evaluated within the group of GSWP participants and later released to the community at large, covers the 10-year period from 0000UTC 1 January 1986 up to 0000UTC 1 January 1996.

3.1.3 Execution and production

After spin-up, each LSS will be integrated globally for the 10-year period 1986-1995. The forcing data are provided at a 3-hour time interval, which adequately resolves the diurnal cycle. For modelers that prefer to run their LSS at a shorter time step, the data will have to be interpolated in time. Sample routines for doing the temporal disaggregation will be supplied (see the GSWP web site), but it will be up to each modeler to choose and report the method he/she prefers.

There should be no adjustment to the spatial grid — all participating LSSs should be run on the same 1° grid. However, though the data subsetting capabilities provided by DODS, it should be possible for the modeler to choose how he proceeds with the integration (e.g., running the entire global grid at each time step, or running each grid point individually from start to finish). The forcing data can be served to the modeler as a sequence of 2-D (or 3-D) global or regional grids, as a set of time series at specific points, or as spatial vectors (e.g. latitude bands). This provides added flexibility to the modeler to tailor the method of integration to best utilize local resources.

Table 7 lists the files on the DODS server that correspond to the input data sets listed in the previous three tables. This should aid modeling groups in accessing the data interactively, or downloading the data to local file systems. The GSWP data sets on the COLA DODS server are available at http://www.monsoondata.org:9191/dods/gswp and can be accessed by browser or via any DODS-enabled client. There are two subdirectories: grid and vector. The grid directory serves full 360x150 grids with the ocean points set to the missing data value. See any

Table 7. GDS file names of input fields for the baseline integration (**B0**).

Fixed Fields	File	Grid only	File
SoilClass	soilclass	**Landmask**	landmask
SoilDepth	soildepth	**Monthly Fields**	**File**
Clay	clay	**LAI**	lai
Sand	sand	**grnFrac**	grnfrac
Silt	silt	**NDVI**	ndvi
Organic	organic	**FPAR**	fpar
Elevation	elevation	**Z0Surf**	z0surf
Slope	slope	**DisplH**	displh
CTI	cti	**Albedo**	albedo_csu
W_fieldcap	w_fieldcap		
W_wilt	w_wilt	**3-hourly**	**File**
W_sat	w_sat	**Tair**	tair_cru
W_bpower	w_bpower_cea84	**Qair**	qair_cru
W_sat_hydc	w_sat_hydc	**Wind**	wind_ncep
W_sat_matp	w_sat_matp	**SWdown**	swdown_srb
Albedo_vi	albedo_vi	**LWdown**	lwdown_srb
Albedo_ir	albedo_ir	**Psurf**	psurf_eds
VegClass	vegclass_igbp, _sib, _bats	**Rainf**	rainf_gswp
classFrac	classfrac	**Rainf_C**	rainf_c_gswp
RootDepth	rootdepth_50, _95	**Snowf**	snowf_gswp
vegFrac	vegfrac		

of the links labeled “info” to see a description of the data set. The vector directory serves the same fields as land-only vectors (compressed by gathering in the NetCDF CF convention) of length 15238. The vectors are more compact, and suitable for input into a LSS. The grids are directly displayable as maps. On the GDS, only one type of data set, the vector data, are actually stored. The GDS automatically repopulates the grid before serving through the grid directory. Note the land-sea mask is served only as a grid file.

The time-invariant fields are in a subdirectory called fixed. So, for instance, to access the vector form of the soil porosity data, one would access the URL: http://www.monsoondata.org:9191/dods/gswp/vector/fixed/w_sat from one’s DODS-enabled client. We encourage modeling groups who need input files that are not supplied to generate their own data sets, and to make them available to the GSWP operational centers, so that we may share them with others who might also need those data sets.

3.1.4 Output fields

Output data can quickly grow to unmanageable proportions if careful choices regarding variables and time intervals are not made. Table 8 shows the projected storage requirements per variable for 10-years of 1° data (ice-free land points only), uncompressed single-precision binary, saved at various frequencies. While monthly data for about 100 variables could be saved on a single CD-ROM, 3-hourly data for a single variable exceeds the capacity of two CD-ROMs.

Table 8. Potential space requirements per output variable.

Frequency	10 Years of Output
Monthly	6.4 MB
Decad	19 MB
Pentad	39 MB
Monthly 3-Hourly	51 MB
Daily	195 MB
3-Hourly	1.6 GB

Monthly data may be adequate for many purposes and most variables. For hydrologic validation over major river basins, for instance, monthly data may be adequate. Pentad frequency adds information on synoptic-subseasonal variations without an inordinate increase in storage demands, and is rather convenient in that there are exactly 73 pentads in a year, excluding leap years. During leap years, data for 29 February are typically included in the 12th pentad, making it effectively a six-day pentad. For GSWP-1, decads were used. Decads are constructed to occur thrice per month, divided at 0000UTC on the 1st, 11th, and 21st of each month. This means the last decad of a month may have 8, 9, 10 or 11 days. The decad interval is slightly more awkward to handle than the pentad, but can be easily accumulated into monthly data.

The daily interval makes sense because it is the only natural interval shorter than yearly. However, it requires 30 times more storage space than monthly. The pentad and decad intervals represent a compromise between daily and monthly. A balance must be struck between maximum useful information and minimum data volume.

The highest temporal resolution, 3-hourly, corresponds to the frequency of the meteorological forcing data. The 3-hourly frequency resolves the diurnal cycle, providing information that the afore-mentioned intervals cannot. This interval, along with daily, was used for the Rhône-AGG experiment. Although at much higher spatial resolution, the Rhône-AGG experiment covered a very small fraction of the globe, and spanned only 3 years (for required reporting of model output). This frequency may be impractical for a global experiment. An alternative is the monthly-3-hourly, which represents a monthly-mean diurnal cycle. Much of the data in the ISLSCP Initiative II collection is reported as monthly 3-hourly. This interval provides basic information on the diurnal cycle, and its variation on seasonal-interannual time scales, and represents a compromise over full 3-hourly resolution.

3.1.4.1 *Daily fields*

The full compliment of ALMA output variables are requested from each model on a daily interval. Fluxes will be reported as means for the day. Surface state variables are daily means, and subsurface state variables are instantaneous unless noted otherwise. The reporting time is 0000UTC, with the averaging period as the 24 hours preceding the reporting time. For example, the daily flux data for 0000UTC 10 November 1993 will represent the mean for 9 November 1993 (UTC). Table 9 gives the ALMA list of output variables to be reported on the daily interval. The sign convention shown is traditional meteorological convention. ALMA also allows for the mathematical sign convention, see their website, listed in the table below, for details.

Table 9. ALMA standard output variables for GSWP-2
(see http://www.lmd.jussieu.fr/ALMA/ for a detailed discussion of these variables).

Variable	Description	Units	Sign (+)
O.1) General energy balance components			
Swnet	Net shortwave radiation	W m^{-2}	Down
Lwnet	Net longwave radiation	W m^{-2}	Down
Qle	Latent heat flux	W m^{-2}	Up
Qh	Sensible heat flux	W m^{-2}	Up
Qg	Ground heat flux	W m^{-2}	Down
Qf	Energy of fusion	W m^{-2}	Sol.<Liq.
Qv	Energy of sublimation	W m^{-2}	Sol.<Vap.
Qa	Heat transferred to snowpack by rainfall	W m^{-2}	Down
DelSurfHeat	Change in surface heat storage	J m-2	Incr.Heat
DelColdCont	Change in snow cold content	J m-2	Decr.Heat

Table 9 (continued)

Variable	Description	Units	Sign (+)
O.2) General water balance components			
Snowf	Snowfall rate	kg m^{-2} s^{-1}	Down
Rainf	Rainfall rate	kg m^{-2} s^{-1}	Down
Evap	Total evapotranspiration (all terms)	kg m^{-2} s^{-1}	Up
Qs	Surface runoff	kg m^{-2} s^{-1}	Out
Qsb	Subsurface runoff	kg m^{-2} s^{-1}	Out
Qsm	Snowmelt	kg m^{-2} s^{-1}	Sol.<Liq.
Qfz	Refreezing of water in the snowpack	kg m^{-2} s^{-1}	Liq.<Sol.
Qst	Water flowing out of snowpack	kg m^{-2} s^{-1}	Out
DelSoilMoist	Change in column soil moisture	kg m^{-2}	Increase
DelSWE	Change in snow water equivalent	kg m^{-2}	Increase
DelSurfStor	Change in surface liquid water storage	kg m^{-2}	Increase
DelIntercept	Change in canopy interception storage	kg m^{-2}	Increase
O.3) Surface state variables			
SnowT	Snow surface temperature	K	Pos.Def.
VegT	Vegetation canopy temperature	K	Pos.Def.
BaresoilT	Bare soil surface temperature	K	Pos.Def.
AvgSurfT	Area-weighted average surface temperature	K	Pos.Def.
RadT	Effective surface radiative temperature	K	Pos.Def.
Albedo	Surface albedo	-	Pos.Def.
SWE	Snow water equivalent on ground (**3-D**)	kg m^{-2}	Pos.Def.
SWEVeg	Snow water equivalent in canopy interception	kg m^{-2}	Pos.Def.
SurfStor	Surface water storage	kg m^{-2}	Pos.Def.
O.4) Subsurface state variables			
SoilMoist	Average layer soil moisture (**3-D**)	kg m^{-2}	Pos.Def.
SoilTemp	Average layer soil temperature (**3-D**)	K	Pos.Def.
SoilWet	Total column soil wetness (wilting point = 0, saturation = 1)	-	> wilting point
O.5) Evaporation components			
PotEvap	Potential evapotranspiration	kg m^{-2} s^{-1}	Up
Ecanop	Evaporation of canopy interception	kg m^{-2} s^{-1}	Up
Tveg	Vegetation transpiration	kg m^{-2} s^{-1}	Up

Table 9 (continued)			
Variable	**Description**	**Units**	**Sign (+)**
Esoil	Bare soil evaporation	kg m^{-2} s^{-1}	Up
Ewater	Open water evaporation	kg m^{-2} s^{-1}	Up
RootMoist	Root zone soil moisture	kg m^{-2}	Pos.Def.
CanopInt	Total canopy water storage	kg m^{-2}	Pos.Def.
EvapSnow	Evaporation of liquid water from snowpack	kg m^{-2} s^{-1}	Up
SubSnow	Snow sublimation	kg m^{-2} s^{-1}	Up
SubSurf	Sublimation of ice from soil and canopy interception	kg m^{-2} s^{-1}	Up
Acond	Aerodynamic conductance	m s^{-1}	Pos.Def.
CCond	Canopy (stomatal) conductance	m s^{-1}	Pos.Def.
O.6) Other hydrologic variables			
WaterTableD	Depth to water table	m	Pos.Def.
O.7) Cold season processes			
SnowFrac	Snow cover fraction	-	Pos.Def.
SAlbedo	Snow albedo	-	Pos.Def.
SnowDepth	Depth of snow layers (**3-D**)	m	Pos.Def.
O.8) Variables to be compared with remotely sensed data			
RadTmax	Maximum daily radiative surface temperature	K	Pos.Def.
RadTmin	Minimum daily radiative surface temperature	K	Pos.Def.

3.1.4.2 *Global 3-hourly fields*

For the final year of the integrations (1995), a subset of the required ALMA output variables will be reported globally at the 3-hourly forcing interval. This is requested for two main reasons. First, this 1-year data set will provide a global picture of the simulation of the diurnal cycle across all seasons for all models. Second, this set of output data will provide sufficient archives for the remote sensing evaluations, as 3-hourly data give a closer rendition of the instantaneous picture that satellite platforms provide. Reporting of data only once a day spatially limits the regions of satellite validation, due to the sun-synchronous orbits of most polar-orbiting platforms. As with the daily data, state variables will be instantaneous and fluxes will be reported as the average rate over the preceding interval (3 hours). Variables to be reported during this intensive model output period (IMOP) are indicated in Table 10. In this case, all state variables are surface variables, and should be averaged over the previous 3-hour interval.

Table 10. ALMA variables to be reported at 3-hourly intervals during the IMOP.			
O1	**O2**	**O3**	**O5**
Swnet	Snowf	SnowT	PotEvap
Lwnet	Rainf	VegT	Ecanop
Qle	Evap	BaresoilT	Tveg
Qh	Qs	AvgSurfT	Esoil
Qg	Qsb	RadT	RootMoist
Qf	Qsm		Acond
Qv	Qfz		CCond
Qa	Qst		
DelSurfHeat	DelSoilMoist		
DelColdCont	DelSWE		
	DelSurfStor		
	DelIntercept		

The simplest way to generate this data set will be for each modeling group to save a restart file for 0000UTC 1 January 1995 during the baseline integration. After completing the baseline run, one can change the output table for the LSS, and rerun the final year with the new reporting frequency and variable list.

3.1.4.3 *Fixed fields*

To aid in the calculation of soil wetness indices, and the comparison of soil wetness and temperatures among models or between models and observations, we ask that each model report a one-time file of soil properties as listed in Table 11. These may be redundant with the input soil parameter data in Table 3, but there are so many different approaches to soil modeling, that it will be clearer if each LSS simply reports the three-dimensional grid of global soil properties used.

Table 11. ALMA fixed fields to report with output data.		
Name	**Description**	**Units**
SoilDepth	Depth of each soil layer in the column (**3D**)	M
W_fieldcap	Field capacity (**3D**)	M
W_wilt	Wilting point (**3D**)	M
W_sat	Saturated water content (porosity × layer depth) (**3D**)	M

3.1.4.4 *Local 3-hourly fields*

It will be impractical to request many global 3-hourly fields, and it is certainly not within current storage and data transmission capabilities to do so for the entire 10-year period. However, it may be tractable to request 3-hourly data over specific grid boxes for more extended periods, corresponding to locations where comparable measurements have been collected that can be used for validation. In particular, locations of field campaigns (such as those listed in Section 3.3.2) or long-term monitoring sites, such as the ARM-CART site, FluxNET sites, or Valdai. Again, capabilities of the DODS server technology should facilitate reintegration of the LSSs over single grid points or small subsets of the global domain. Details of these point integrations will be decided later in the project.

3.1.4.5 *Recording output data*

All output data should conform to the ALMA standards, and be written as land-only vectors of length 15238 single-precision floating-point values. Assistance and sample software are available from the GSWP and ALMA websites.

Sample FORTRAN code will be made available on the GSWP website to aid in the reporting of requested ALMA output variables from a LSS. The basis of the code is two subroutines. The first maps a LSS's variables to the variables in the ALMA output table through a series of equations laid out in the same structure as the ALMA table. The example given is for SSiB, where the LSS variables are passed to the subroutine via a FORTRAN-90 data module. It would be straightforward for another modeler to simply re-map the variables of another LSS into those ALMA variables. The second set of code includes a subroutine for writing the ALMA output variables to a file in NetCDF format, and an easily edited ASCII table, structured on the ALMA output table, that indicates which variables should be reported and the reporting frequency. Again, after an initial investment of time to implement such code in a LSS, it thereafter becomes trivial to re-run the model with different output requirements.

3.1.4.6 *Ancillary information*

Each modeler should document all non-standard choices made in regard to the integration of their model and submit that information to the ICC with their final output. This will help to ensure clean comparisons of the products from the various models and may affect the way multi-model products are assembled and interpreted. Standard reports should also include specific information about each model, including references to relevant published work (as in the RhôneAGG experiment). Table 12 lists the information requested in the report.

Table 12. Requested ancillary information about the LSS and its integrations in GSWP

Model description:

1. List the most recent/important reference(s) for your LSSs.
2. Give the name of your model, as used in the output file names (see Section 3.2.1)
3. What time step did you use for your LSS?
4. Describe the vertical structure (number and depth of layers, or other methods of discretization) of your LSS for each of the following:
    a. Soil moisture
    b. Soil temperature
    c. Snow
5. How were soil properties (wilting point, field capacity, porosity, hydraulic conductivity, thermal conductivity, etc.) assigned?
    a. From provided fields (Table 3)
    b. Other (describe)
6. Does your LSS include frozen soil properties:
    a. No change in soil moisture treatment below freezing
    b. Binary frozen soil moisture (all or none)
    c. Fractional water and ice content in the soil
7. Does your LSS include liquid water in the snowpack?
8. What type of sub-grid parameterizations are used by your LSS (what best describes your LSS)?
    a. Effective parameters (and which parameters)
    b. Tile/mosaic approaches (structure: i.e. multiple columns or many surfaces over a single soil, etc.)
    c. Integration using probability distributions (gamma, normal, etc...and for which variables)
    d. Some combination of the above (describe)
9. What sub-grid parameterizations are used in your LSS? (none, one or several of the below)
    a. Sub-grid runoff scheme
    b. Sub-grid precipitation interception
    c. Sub-grid saturated fraction for evaporation
    d. Fractional snow cover area parameterization and how it influences the overall surface energy budget
    e. Other (describe)
10. Is there recharge of soil moisture from below from the water table?
11. What is the surface energy budget structure of your LSS?
    a. Single soil/snow/vegetation composite surface energy budget
    b. Distinct canopy and soil with composite snow
    c. Distinct soil, canopy and snow
    d. Other (describe)

Continued on next page

Table 12 (continued)

Data issues:

12. Which soil parameters are used in your LSS?
    a. Sand, silt, clay fractions directly (and what method for deriving hydraulic and thermal parameters?)
    b. Other derived parameters from Table 3 (which ones?)
    c. Provided Cosby parameters based on soil class (Table 2)
    d. Parameters derived from provided data (describe how and for what method; e.g., Van Genuchten)
    e. Other (describe)
13. What vegetation type data are used by your LSS?
    a. IGBP land cover types (Table 5)
    b. SiB vegetation types (Table 5)
    c. BATS vegetation types (Table 5)
    d. Plant Functional Types derived from IGBP data
    e. Other (describe)
14. Which of the supplied vegetation data sets from Table 4 did you use?
15. Which forcing data from Table 6 did your LSS require?
16. How did you interpolate the 3-hourly forcing data?
    a. Used the provided interpolation code from the GSWP web site
    b. No interpolation (model time step is 3 hours or longer)
    c. Other (describe)
17. What data sets does your LSS require that were not supplied by GSWP? How did you supply these data?

3.2 Inter-comparison Center

3.2.1 Data submission

Each participating modeling group will submit their model output to the Inter-Comparison Center (ICC) at the University of Tokyo for quality control, and basic comparisons. Again, data may simply be served to the ICC via a DODS server at the modeling group's home site, or by FTP or HTTP. Otherwise, the data will be submitted by magnetic or optical media to be determined.

A website has been set up at the ICC for delivery of output data (http://www.tkl.iis.u-tokyo.ac.jp:8080/gswp2/). The web side allows for data to be submitted through a web browser from local disk to the ICC (pushed), or posted on the modeler's own FTP site or DODS server to be retrieved by the ICC (pulled). To facilitate this semi-automated data retrieval, a convention for file names should be followed. Daily output data from LSSs should be structured as:

MODEL_expNAME_VERSION_dYEARMONTH.nc

where MODEL: model name, NAME: experiment name identifier (such as GSWP2 for baseline simulations; a list of names for sensitivity experiments is given in Table 13), VERSION: version number of the run by 8 digits of the run (probably year, month, and date of completion of the integration), and the year and month of the contents. The letter `d` indicated daily data, `h` is used for 3-hourly data, and `f` for the fixed fields.

The 10-year data set will be divided into 120 monthly files containing all of the recommended variables in one file. For example, the baseline run by the gSiB model from IIS for October 1992 completed in mid-April would be:

```
gSiBIIS_expB0_20030415_d199210.nc
```

The version number is crucial since many models will likely re-run and submit their updated outputs.

3.2.2 Quality control

The ICC will analyze the multi-model envelope and quantify the uncertainty in simulated land-surface state variables and fluxes as a function of variable and location. It is advised that before completing the entire baseline simulation, each modeling group should submit a sample one-month output data set produced by the LSS. A check can then be performed to ensure units are not obviously incorrect and that all required variables are reported in a defined and readable manner. ALMA also provides code to perform basic water and energy balances are satisfied. All modeling groups should check their output before sending it to the ICC. While the ICC has primary responsibility for basic model comparison, novel ideas for comparisons by other researchers will be welcome. The establishment of a multi-model consensus or model-independent analysis will be pursued among the two operational centers (see Section 2.1).

3.2.3 Data redistribution

Once LSS data are submitted to the ICC and are checked and accepted, they will be made available for evaluation by other GSWP-2 participants. The primary means of data distribution will be online access. Submission of LSS output for the baseline simulation is taken as an agreement by that modeling group to the release of the data to other GSWP-2 participants. These data will be available not only for validation, evaluation, and remote sensing applications, but also among the modeling groups themselves for cross-comparison if they so desire. The timeline in Figure 5 shows the working schedule for data release.

The first product to be released to the general public will be the multi-model analysis. This will occur after the first overview paper, describing the model comparison and production of the multi-model analysis, has been written and submitted to a peer-reviewed journal. Release of data from individual model simulations to the general public will be made at a later date, after a critical amount of validation and evaluation has been publicly presented and/or submitted for publication. Participation in the baseline simulation is taken as an agreement by each modeling group to the eventual public release of that data, *unless expressly stated otherwise* by the representatives of that modeling group.

Data from sensitivity studies will be released within the set GSWP-2 participants as those sensitivity studies are completed. Schedules for individual sensitivity studies will be set and posted on the website as appropriate. Each sensitivity study will have a primary investigator who will lead the evaluation, and who will determine if and when individual model results should be made available to general public, with the opportunity for exception by individual modeling groups as spelled out above.

3.3 Evaluation

The main thrust of external analysis of the LSS results will be in the form of model evaluation. There are several types of evaluation that can be performed, but they fall into two basic categories: *in situ* validation and remote sensing evaluation. This section is concerned with *in situ* validation — remote sensing evaluation is discussed separately in Section 3.5.

In situ validation means direct validation against observations. We envisage two primary types of *in situ* validation: hydrologic validation (be conducted at the IIS, University of Tokyo) and local validation (led from the Hydrologic Sciences Branch of the Goddard Space Flight Center of NASA).

3.3.1 Hydrologic validation

3.3.1.1 *Streamflow*

Hydrologic validation is performed on basin-integrated model runoff against observed streamflow and discharge. Thus, it is an integrated validation of the models' water balances over basins, smoothed in time. Each LSS runoff will be routed through common routing schemes for validation. For this validation effort, only the runoff terms from the LSSs will be used.

3.3.1.2 *Hydrology standards for ALMA*

Currently the ALMA data protocol does not contain variables associated with sub-surface hydrology (e.g., parameters such as topographic index or state variables such as water table depth) or lateral transport of surface water (parameters such as river flow direction, slope, or channel width, nor fluxes such as streamflow out of a grid box). ALMA does have a surface water storage variable that can be used for river or lake water mass. GSWP-2 offers an opportunity to expand ALMA to include a complete set of hydrologic quantities.

3.3.2 Field campaign data

The local validation will have two thrusts - validation against data from long-running observed networks, and validation against data from intensely monitored but short-duration field programs. The ISLSCP Initiative II period overlaps part or all of a number of field campaigns (see Table 1).

3.3.2.1 *Access*

Comparison to measurements at locations in these field campaigns would be very useful, but presents practical difficulties as the data are widely scattered, under various usage limitations, and in a number of different formats. Many of these campaigns included data information systems to collect and archive the measurements that were gathered. In some cases (e.g. First ISLSCP Field Experiment [FIFE]), a certain degree of data synthesis was performed to put the observational record in a more useful form for land surface modelers (Betts and Ball, 1998).

It would be a task of value to GLASS and ALMA, as well as to GSWP, to make an effort to gain access to these data. This effort is being pursued in parallel with the GSWP-2 experiment. Data sets that are otherwise publicly available, and of use for calibration and validation for GSWP-2, will be collected and served along with the other data on the DODS servers. To date, data from FIFE, HAPEX-Sahel, Global Soil Moisture Data Bank and the Soil Climate Analysis Network (SCAN) have been collected.

3.3.2.2 *ALMA formatting*

GLASS and ALMA will further add value by synthesizing field measurements into the ALMA standard. This consolidation and standardization will be pursued in conjunction with the data information systems (DIS) personnel of the field campaigns. Surface observational data sets that are already publicly available and deemed of potential use to GSWP-2 will be converted to ALMA format and made available to the GSWP-2 participants online for calibration, evaluation and validation exercises. With consent of the producers and managers of each data set, the ALMA-format version of the data may be made publicly available, either through GLASS, or from the originating DIS. Whether served by ALMA or the individual DISs, the standardized data should be made accessible to the land surface modeling community. We recommend that this non-trivial task be heartily pursued with substantial assistance from ALMA.

3.3.3 Observational networks

Long-running observational networks that monitor land surface state variables or fluxes are not as plentiful as stations that monitor near-surface meteorology. They include the various soil moisture monitoring stations in the Soil Moisture Data Bank, national soil temperature monitoring networks (e.g. Germany), snow monitoring networks such as SnowTel, and potentially a few of the original FluxNet flux monitoring sites that were in place before the end of the ISLSCP Initiative II period.

3.4 Sensitivity studies

Two categories of sensitivity studies are proposed, that will involve re-integration of the participating LSSs. The first category involves runs to examine the role of uncertainty in meteorological forcing data in the estimates of surface state variables and fluxes. This will involve using alternative meteorological analyses (namely, the ECMWF reanalysis, of which the complete 10-year data set should be available from ISLSCP in early 2003), and alternative observational data sets for hybrid correction of the analyses.

The second category involves uncertainty in the specification of land surface properties. There are three different maps of global land cover classification. The impact of suppression of the interannual variability of vegetation phenology will be investigated. For LSSs with sub-grid tiling of surface properties, we may examine the impact of parameter aggregation across those models on the global scale. These tests will determine the sensitivity of LSSs to the specification of some key observable parameters on a global scale, and thus establish uncertainties caused by parameter choices.

Participation in the sensitivity studies is optional, and it is expected to be beyond the capabilities and interests of most modeling groups to participate in all of them. However each modeling group will hopefully find one or more of the studies of sufficient interest to agree to participate.

The precipitation (**P**) series is a key study for GSWP-2, and would benefit from participation of at least 6 modeling groups. From this series, the impact of remote sensing data on estimates of global surface hydrologic cycle, the impact of gauge errors, reanalysis errors, and the relative merits of the precipitation statistics of the two reanalysis products can be estimated. With more LSSs participating, the range of inter-model sensitivity to uncertainties in precipitation can be determined.

The radiation (**R**) series will provide a similar estimate for the impact of radiation errors and uncertainties on the LSSs. For modelers with a particular interest in the relative impacts of shortwave and longwave radiation errors, such as those whose LSS is coupled into an atmospheric GCM where parameterization development is occurring, the **RS** and **RL** will be particularly interesting and useful.

The all-meteorological (**M1**) study gives the broadest assessment as to the impact of differences from the two reanalyses. This study could be useful for anyone thinking of expanding their LSS simulations for other research or applications, and wanting to ascertain the best reanalysis product to use for their endeavor.

Similarly, the study using different vegetation data sets (**V1**) may help a modeling group decide which global vegetation data set to use with their model. The interannual vegetation (**I1**) experiment is useful for any modeling group that is not currently using dynamic vegetation or vegetation phenology parameterizations, but is considering using or developing one for their LSS. This study will show the sensitivity of the surface energy and water budgets to interannual variations in vegetation, which may be a factor in deciding the usefulness and approach toward modeling vegetation phenology. For models that have predicted vegetation phenology, this study can be used along with both specified and predicted time-varying observed vegetation to validate that component of the models.

3.4.1 Optional input data sets

Alternate data sets of meteorological forcing will be generated and supplied for the sensitivity experiments. These will include non-hybrid versions of the NCEP/DOE reanalysis, hybrid and non-hybrid versions of the ECMWF ERA40 reanalysis, and alternate hybridizations of those reanalysis data. The alternate hybrid data sets would use different observationally-based precipitation or near-surface temperature data sets for the hybridization process. These data sets will all be in ALMA format, and could easily be swapped in and out. Sample software to query and parse multiple input ALMA files will be supplied on the GSWP-2 website and in the ALMA software bazaar.

Alternate data sets for land surface parameters will be supplied in a similar manner. There may be inconsistencies among the ISLSCP Initiative II data sets that will have to be addressed (e.g., snow-free albedo inconsistent with land cover type). These inconsistencies may exist for the control case data sets as well.

Table 13 lists the alternative data sets that will be made available for each sensitivity study. Files not listed are the same as for the baseline simulation. Again, the latest information on changes will be posted on the project web site.

Table 13. Alternate input files for the sensitivity studies.

Exp	Description	Field	Files
B0	Baseline integration		As given in Table 7
P1	Hybrid ERA-40 precipitation (instead of NCEP/DOE)	**Rainf** **Snowf** **Rainf_C**	rainf_era snowf_era rainf_c_era
P2	NCEP/DOE hybrid with GPCC corrected for gauge undercatch (no satellite data)	*as above*	rainf_gage snowf_gage, rainf_c_gage
P3	NCEP/DOE hybrid with GPCC (no undercatch correction)	*as above*	rainf_gpcc, snowf_gpcc, rainf_c_gpcc
P4	NCEP/DOE precipitation (no observational data)	*as above*	rainf_ncep, snowf_ncep, rainf_c_ncep
P5	NCEP/DOE hybrid with Xie daily gauge precipitation	*as above*	rainf_xie, snowf_xie, rainf_c_xie
R1	NCEP/DOE radiation	**LWdown** **SWdown**	lwdown_ncep swdown_ncep
RS	NCEP/DOE shortwave only	**SWdown**	swdown_ncep
RL	NCEP/DOE longwave only	**LWdown**	lwdown_ncep
R2	ERA-40 radiation	**LWdown** **SWdown**	lwdown_era swdown_era
M1	All NCEP meteorological data (no hybridization with observational data)	**Tair, Qair, LWdown, SWdown, Rainf, Snowf, Rainf_C**	*_ncep
V1	UMCP vegetation class data	**VegClass**	vegclass_umcp
I1	Climatological vegetation	**LAI, vegFrac, grnFrac, NDVI, FPAR, Z0Surf, DisplH**	*_clim

3.4.2 Data submission/distributed analysis

Most likely, the ICC will not be responsible for collecting data from all of the sensitivity studies. It will likely be the responsibility of one or more participants to propose, arrange, oversee and analyze the results from each sensitivity study, similar to the sub-projects in AMIP. This is a situation where the DODS server/client technology could be extremely useful in the sharing and analysis of the results from multiple modeling groups, provided those groups choose to install and operate a DODS server at their sites. DODS would allow distributed analysis of the results of the control and sensitivity studies, potentially making analysis of results much easier and swifter than previously possible.

3.5 Remote sensing applications

Remote sensing offers the only opportunity for large spatial scale validation of the LSS state variables. Validation of radiative skin temperature, snow coverage extent, and surface soil moisture should be possible. There also is a plan to develop code for LSSs to predict surface microwave radiances as observed from space. This effort ties into the general remote sensing thrust in GSWP-2 (see Section 2.4).

Forward retrievals for microwave channels sensitive to soil moisture will be calculated *a posteriori* using the 3-hourly model output fields from each LSS. LSSs that have poor vertical resolution in the soil column near the surface may not perform well in this test, unless a reliable means for interpolating soil moisture near the surface can be applied. Potential observational data sources include 6.6 GHz SMMR products from the 1987-1988 period and 19 GHz SSM/I measurements.

The application of backwards (i.e., retrieval) algorithms to remote sensing observations allows for the possibility of direct comparisons between LSS and remote estimates of soil moisture. The ERS global soil moisture data set - derived from 5.3 GHz active radar measurements from 1992-2000 - provides one potential point of comparisons for LSS surface soil moisture predictions during the GSWP-2 period (Wagner et al. 1999).

Developing high-quality retrieval algorithms for surface soil moisture often requires a great deal of ancillary information concerning land surface states. One challenge for interpreting high frequency (> 5 GHz) microwave and radar observations of the land surface is isolating temporal soil moisture signals from simultaneous seasonal variations in vegetation cover and surface conditions. Independent predictions of seasonal soil moisture trends from LSSs during the GSWP-2 period may aid in the filtering of vegetation and surface roughness effects from SSM/I- and ERS-derived soil moisture products.

Radiometric skin temperature will be compared to remote sensing estimates of skin temperature. Daily maximum and minimum radiative temperature are to be reported from each model for the 10-year period, along with daily mean skin temperature, and the radiative temperature based on the daily mean upward longwave radiation. This will allow sufficient coverage to compare to available remote sensing data.

Snow and freeze/thaw coverage comparisons will also be performed using the daily snow cover and LSS predictions of soil thermal states.

Models which predict vegetation phenology will asked to report NDVI as an output. This will be compared to the monthly observed NDVI (which is provided as an input for those LSSs that require it.

3.6 Contact Information

The latest information on GSWP-2 and the latest version of this plan can be found on the project website: http://www.iges.org.gswp/

Email inquiries should be addressed to: gswp@cola.iges.org.

At the GSWP operational centers:

Paul A. Dirmeyer or **Xiang Gao**
Center for Ocean-Land-Atmosphere Studies
4041 Powder Mill Road, Suite 302
Calverton, Maryland 20705-3106 , USA
Phone: +1-301-902-1254
Fax: +1-301-595-9793
E-mail: dirmeyer@cola.iges.org

Taikan Oki

Research Institute for Humanity and Nature
Kyoto 602-0878, Japan
Phone: +81-075-229-6180
Fax.: +81-075-229-6150
E-mail: oki@chikyu.ac.jp

Institute of Industrial Science, University of Tokyo
4-6-1 Meguro-ku, Komaba, Tokyo 153-8505, Japan
Phone: +81-03-5452-6382
Fax.: +81-03-5452-6383
E-mail: taikan@iis.u-tokyo.ac.jp

ACKNOWLEDGEMENTS:

Many pairs of eyes have looked over this plan, and their owners have provided many useful comments and suggestions – we thank them all. In particular, substantive contributions have been provided by Randy Koster, Mei (Maggie) Zhao, Wade Crow, Adam Schlosser, and Jan Polcher.

We would also like to than everyone who participated in the GSWP-2 kickoff meeting in Calverton, Maryland, USA on 30 September – 1 October 2002 for their help in crystallizing this project. Support for this project is provided in the United States by National Aeronautics and Space Administration Grant NAG5-11579, and in Japan by the Core Research for Evolutional Science and Technology (CREST) of the Japan Science and Technology Corporation (JST).

REFERENCES

Adler, R. F., C. Kidd, G. Petty, M Morissey, and H. M. Goodman, 2001: Intercomparison of global precipitation products: The third precipitation intercomparison project (PIP-3). *Bull. Amer. Meteor. Soc.*, **82**, 1377-1396.

Andre, J.-C., J.P. Goutorbe, T. Schmugge, and A. Perrier, 1989b: HAPEX-MOBILHY: Results from a large-scale field experiment. In: *Remote sensing and large-scale global processes*, A. Rango, Ed., Wallingford, UK, International Association of Hydrological Sciences, 13-20.

Bolle, H.J., André, J.C., Arrue, J.L., Barth, H.K., Bessemoulin, P., Brasa, A., de Bruin, H.A.R., Cruces, J., Dugdale, G., Engman, E.T., Evans, D.L., Fantechi, R., Fiedler, F., van de Griend, A., Imeson, A.C., Jochum, A., Kabat, P., Kratzsch, T., Lagouarde, J.-P., Langer, I., Llamas, R., Lopez-Baeza, E., Melia Miralles, J., Muniosguren, L.S., Nerry, F., Noilhan, J., Oliver, H.R., Roth, R., Saatchi, S.S., Sanchez Diaz, J., de Santa Olalla, M., Shuttleworth, W.J., Søgaard, H., Stricker, H., Thornes, J., Vauclin, M. and Wickland, D., 1993: EFEDA: European field experiment in a desertification threatened area. *Annales Geophysicae*, **11**, 173-189.

Betts, A. K., and J. H. Ball, 1998: FIFE surface climate and site-average dataset 1987-1989. *J. Atmos. Sci.*, **55**, 1091-1108.

Boone, A., F. Habets, and J. Noilhan, 2001: The Rhône-aggregation experiment. *GEWEX News*, **11(3)**, 3-5.

Boone, A., F. Habets, J. Noilhan, E. Blyth, D. Clark, P. Dirmeyer, Y. Gusev, I. Haddeland, R. Koster, D. Lohmann, S. Mahanama, K. Mitchell, O. Nasanova, G.-Y. Niu, A. Pitman, J. Polcher, A. B. Shmakin, K. Tanaka, B. van den Hurk, S. Verant, D. Verseghy, and P. Viterbo, 2002: The Rhône-aggregation land surface scheme intercomparison project: An overview. *J. Climate*, (in review).

Budyko, M. I., 1974: *Climate and Life,* Academic Press, New York, 508 pp.

Calvet, J.-C., P. Bessemoulin, J. Noilhan, C. Berne, I. Braud, D. Courault, N. Fritz, E. Gonzalez-Sosa, J.-P. Goutorbe, R. Haverkamp, G. Jaubert, L. Kergoat, G. Lachaud, J.-P. Laurent, P. Mordelet, A. Olioso, P. Péris, J.-L. Roujean, J.-L. Thony, C. Tosca, M. Vauclin, and D. Vignes, 1999: MUREX: a land-surface field experiment to study the annual cycle of the energy and water budgets. *Annales Geophysicae*, **17**, 838-854.

Chou S. H., Shie CL, Atlas RM, Ardizzone J, 1997: Air-sea fluxes retrieved from special sensor microwave imager data , *J. Geophys. Res*., **102**, 12705-12726.

Clapp, R., and G. Hornberger, 1978: Empirical equations for some soil hydrologic properties. Water Resour. Res., 14, 601 604.

Cosby, B. J., G. M. Hornberger, R. B. Clapp, and T. R. Ginn, 1984: A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. Water Resour. Res., 20, 682 690.

Coughlan, M., and R. Avissar, 1996: The Global Energy and Water Cycle Experiment (GEWEX) Continental-Scale International Project (GCIP): An overview. *J. Geophys. Res.,* **101**, 7139-7148.

Crow, W. T., and E. F. Wood, 2002: 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 Res.*, (in press).

Dirmeyer, P. A., and J. Shukla, 1993: Observational and modeling studies of the influence of soil moisture anomalies on the atmospheric circulation. *Predictions of Interannual Climate Variations* (J. Shukla, ed.), NATO Series I, Vol. 6, Springer-Verlag, 1-23.

Dirmeyer, P. A., A. J. Dolman, and N. Sato, 1999: The Global Soil Wetness Project: A pilot project for global land surface modeling and validation. *Bull. Amer. Meteor. Soc.*, **80**, 851-878.

Dirmeyer, P. A., 2000: Using a global soil wetness data set to improve seasonal climate simulation. *J. Climate*, **13**, 2900-2922.

Dirmeyer, P. A., 2001: Climate drift in a coupled land-atmosphere model. *J. Hydrometeor.*, **2**, 89-100.

Dirmeyer, P. A., and L. Tan, 2001: A multi-decadal global land-surface data set of state variables and fluxes. *COLA Technical Report 102* [Available from the Center for Ocean-Land-Atmosphere Studies, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705 USA], 43 pp.

Dirmeyer, P. A., 2002: The role of the land surface background state in climate predictability. J. Hydrometeor, (in review).

Douville, H., and F. Chauvin, 2000: Relevance of soil moisture for seasonal climate predictions: A preliminary study. *Climate Dyn.*, **16**, 719-736.

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. Climate*, **14**, 2381-2403.

Douville, H., 2002: Influence of soil moisture on the Asian and African monsoons. Part II: Interannual variability. *J. Climate*, **15**, 701-720.

Eltahir, E. A. B., and R. L. Bras, 1993: On the response of the tropical atmosphere to large-scale deforestation. *Center for Global Change Science, Report No. 19* [Available from the Ralph M. Parsons Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA], 28pp.

Entin, J. K., A. Robock, K. Ya. Vinnikov, V. Zabelin, S. Liu, and A. Namkhai, 1999: Evaluation of Global Soil Wetness Project soil moisture simulations. *J. Meteor. Soc. Japan*, **77**, 183-198.

Fennessy, M. J., and C. A. Schlosser, 2002: The USAF 47 Km Daily Snow Depth Analysis suitable for climate research. COLA Technical Report, (in preparation).

Gash, J. H. C., and C. A. Nobre, 1997: Climatic effects of Amazonian deforestation: Some results from ABRACOS. *Bull. Amer. Meteor. Soc.*, **78**, 823-83.

Goutorbe, J-P, T. Lebel, A. Tinga, P. Bessemoulin, J. Brouwer, A. J. Dolman, E. T. Engman, J. H. C. Gash, M. Hoepffner, P. Kabat, Y. H. Kerr, B. Monteny, S. Prince, F. Said, P. Sellers, J. S. Wallace. 1994. HAPEX-Sahel: A Large Scale Study of Land-Atmosphere Interactions in the Semi-Arid Tropics. *Ann. Geophys.*, **12**, pp. 53-64.

Hall, F. G., and P. J. Sellers, 1995: First International Satellite Land Surface Climatology Project (ISLSCP) Field Experiment (FIFE). *J. Geophys. Res.*, **100**, 25383-25396.

Hall, F. G., 1999: Introduction to special section: BOREAS in 1999: Experiment and science overview. *J. Geophys. Res.*, **104**, 27627-27639.

Halldin, S., S-E. Gryning, L. Gottschalk, A. Jochum, L-C. Lundin, and A.A. Van de Griend, 1999: Energy, water and carbon exchange in a boreal forest landscape --- NOPEX experiences. *Agricultural And Forest Meteorology*, **98-99**, 5-29.

Henderson-Sellers, A., A. J. Pitman, P. Irannejad, and K. McGuffie, 2002: Land surface simulations improve atmospheric modeling. *EOS, Tansactions, American Geophysical Union*, **83**, 145, 152.

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 hydrologic modeling using data assimilation. Water Resour. Res., 34, 3405 3420.

Huffman, G. J., R. F. Adler, P. Arkin, A. Chang, R. Ferraro, A. Gruber, J. Janowiak, A. McNab, B. Rudolf, and U. Schneider, 1997: The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull. Amer. Meteor. Soc.*, **78**, 5-20.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, & D. Joseph, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437-471.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP-DOE AMIP-II reanalysis (R-2). Bull*. Amer. Meteor. Soc.,* **83**, 1631-1648.

Kharin, V. V., and F. W. Zweirs, 2002: Climate predictions with multimodel ensembles. *J. Climate*, **15**, 793-799.

Koster, R. D., and P. C. D. Milly, 1997: The interplay between transpiration and runoff formulations in land surface schemes used with atmospheric models. *J. Climate*, **10**, 1578-1591.

Koster, R. D., T. Oki, and M. J. Suarez, 1999: Assessing success in the offline validation of land surface models. J. Meteor. Soc. Japan, 77, 257 263.

Koster, R. D., P. A. Dirmeyer, P. C. D. Milly, and G. L. Russell, 2001: Comparing GCM-generated land surface water budgets using a simple common framework. *Land Surface Hydrology, Meteorology and Climate: Observations and Modeling, V. Lakshmi, J. Albertson, and J. Schaake, eds., Water Science and Applications series, American Geophysical Union, Washington DC*, **3**, 95-105.

Krishnamurti, T. N., C. M. Kishtawal, T. E. LaRow, D. R. Bachiochi, Z. Zhang, C. E. Willifor, S. Gadgil, S. Surendran, 1999: Improved weather and seasonal climate forecasts from multimodel superensemble. *Science*, **285**, 1548-1550.

Krishnamurti, T. N., C. M. Kishtawal, Z. Zhang, T. LaRow, D. Bachiochi, E. Williford, S. Gadgil, and S. Surendran, 2000: Multimodel ensemble forecasts for weather and seasonal climate. *J. Climate*, **13**, 4196-4216.

Lawford, R. G., 1999: A midterm report on the GEWEX Continental-scale International Project (GCIP). *J. Geophys. Res.*, **104**, 19279-19292.

Liston, G. E., Y. C. Sud, and G. K. Walker, 1993: Design of a global soil moisture initialization procedure for the Simple Biosphere model. *NASA Tech. Memo. 104590,* Goddard Space Flight Center, Greenbelt, Maryland, 138 pp.

Margulis, S. A., and D. Entekhabi, 2001: A coupled land surface-boundary layer model and its adjoint. *J. Hydrometeorol.*, **2**, 274-296.

Mitchell, K. E., and Y. Lin, 1994: Production of 6-hourly continental precipitation data sets for 1987 and 1988 for ISLSCP Initiative I. Development Division, National Meteorological Center, Washington, DC.

Mocko, D. M., G. K. Walker, and Y. C. Sud, 1999: New snow-physics to compliment SSiB Part II: Effects on soil moisture initialization and simulated surface fluxes, precipitation and hydrology of GEOS II GCM. *J. Meteor. Soc. Japan*, **77**, 349-366.

Motoya, K., K. Takata, K. Masuda, and T. Oki, 2002: Sensitivity of precipitation gauge correction for the estimation of the global water balance. *J. Hydrometeor.,* (submitted).

Neelin, J. D., and N. Zeng, 2000: A quasi-equilibrium tropical circulation model - formulation. *J. Atmos. Sci.*, **57**, 1741-1766.

New, M., M. Hulme, and P. Jones, 1999: Representing twentieth-century space-time climate variability. Part II: Development of 1901-96 mean monthly terrestrial climatology. *J. Climate*, **12**, 829-856.

New, M., M. Hulme, and P. Jones, 2000: Representing twentieth-century space-time climate variability. Part II: Development of 1901-96 monthly grids of terrestrial surface climate. *J. Climate*, **13**, 2217-2238.

Oki, T., and Y. Sud, 1998: Design of the global river channel network for Total Runoff Integrating Pathways (TRIP). *Earth Interactions*, **2**, 001.

Oki, T., T. Nishimura, and P. Dirmeyer, 1999: Assessment of annual runoff from land surface models using Total Runoff Integrating Pathways (TRIP). *J. Meteor. Soc. Japan*, **78**, 235-255.

Raschke, E, U. Karstens, R. Nolte-Holube, R. Brandt, H.-J. Isemer, D. Lohmann, M. Lobmeyr, B. Rockel, and R. Stuhlmann, 1998: The Baltic sea experiment BALTEX: A brief overview and some selected results of the Authors. *Surveys in Geophysics*, **19**, 1-22.

Robock, A., K. Ya. Vinnikov, G. Srinivasan, J. K. Entin, S. E. Hollinger, N. A. Speranskaya, S. Liu, and A. Namkhai, 2000: The Global Soil Moisture Data Bank. *Bull. Amer. Meteor. Soc.*, **81**, 1281-1299.

Ropelewski, C. F., J. E. Janowiak, and M. F. Halpert, 1985: The analysis and display of real time surface climate data. *Mon. Wea. Rev.*, **113**, 1101-1107.

Rudolf, B., H. Hauschild, W. Reuth, and U. Schneider, 1994: Terrestrial precipitation analysis: Operational method and required density of point measurements. *Global Precipitation and Climate Change* (M. Desbois, F. Desalmond, eds.), NATO ASI Series I, Vol 26, Springer-Verlag, 173-186.

Saleem, J. A., and G. D. Salvucci, 2002: Comparison of soil wetness indices for inducing functional similarlty of hydrologic response across sites in Illinois. *J. Hydrometeor.*, **3**, 80-91.

Salvucci, G. D., 1997: Soil and moisture independent estimation of stage-two evaporation from potential evaporation and albedo or surface temperature. *Water Resour. Res.*, **33**, 111-122.

Sellers, P. J., and F. G. Hall, 1992: FIFE in 1992: Results, scientific gains, and future research directions. *J. Geophys. Res.*, **97**, 19,091-19,109.

Sellers, P. J., F. Hall, H. Margolis, B. Kelly, D. Baldocchi, G. den Hartog, J. Cihlar, M. G. Ryan, B. Goodison, P. Crill, K. J. Ranson, D. Lettenmaier, and D. E. Wickland, 1995: The Boreal Ecosystem-Atmosphere Study (BOREAS): An overview and early results from the 1994 field year. *Bull. Amer. Meteor. Soc.*, **76**, 1549-1577.

Shuttleworth, W. J., Z.-L. Yang, M. A. Arain, 1997: Aggregation rules for surface parameters in global models. *Hydrol. Earth Sys. Sci.*, **2**, 217-226.

Stewart, R.E., H.G. Leighton, P. Marsh, G.W.K. Moore, H. Ritchie, W.R. Rouse, E.D. Soulis, G.S. Strong, R.W. Crawford, and B. Kochtubajda, 1998: The Mackenzie GEWEX Study: The water and energy cycles of a major North American river basin. *Bull. Amer. Meteor. Soc.*, **79**, 2665-2684.

van Genuchten, M. T., 1980: A closed form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J., 44, 892 898.

Wagner, W., G. Lemoine, and H. Rott, 1999: A method for estimating soil moisture from ERS scatterometer and soil data, *Remote Sens. Environ.*, **70**, 191-207.

Wang, J., Y. Gao, and Y. Hu, 1993: An overview of the HEIFE experiment in the People's Republic of China. Exchange Processes at the Land Surface for a Range of Space and Time Scales. Proceedings of a symposium held during the joint meeting of the International Association of Meteorology and Atmospheric Physics and IAHS at Yokohama, July 1993, edited by H.-J. Bolle, R. A. Feddes & J. D. Kalma, IAHS Pub. 212, 397-406.

Zeng, N., 1998: Understanding climate sensitivity to tropical deforestation in a mechanistic model. *J. Climate*, **11**, 1969-1975.