A History and Review of the Global Soil Wetness Project (GSWP)

PAUL A. DIRMEYER
Center for Ocean–Land–Atmosphere Studies, Calverton, Maryland

(Manuscript received 29 October 2010, in final form 22 February 2011)

ABSTRACT

The Global Soil Wetness Project (GSWP) is an international land surface modeling research effort involving dataset production, validation, model comparison, and scientific investigation in the areas of land surface hydrology and climatology. GSWP is characterized by the integration of multiple land surface models on a latitude–longitude grid in a stand-alone uncoupled mode, driven by meteorological forcing data constructed by combining atmospheric analyses and gridded observed data products. The models produce time series of gridded estimates of land surface fluxes and state variables that are then studied and compared. Defining characteristics that have distinguished GSWP include its global scale, application of land surface models in the same gridded structure as they are used in weather and climate models, and the multimodel approach, which included production of a multimodel analysis in its second phase. This paper gives an overview of the history of GSWP beginning with its inception within the International Satellite Land Surface Climatology Project. Various phases of the project are described, and a review of scientific results stemming from the project is presented. Musings on future directions of research are also discussed.

1. Introduction

In the middle of the twentieth century, research in the earth sciences was concerned with delving into the workings of the individual components and how they behaved and functioned in isolation. Meteorology, oceanography, hydrology, ecology, etc., operated as highly independent disciplines, usually crossing only when they found they shared underlying principles of physics or techniques of mathematics. However, as these fields have evolved, their respective researchers increasingly found that the progress of their understanding was encountering barriers at the boundaries between the disciplines. Further progress would require consideration of the larger system and the interactions between the components that were defined early in the century. Often specific phenomena served as the bridge between disciplines. For example, study of El Niño brought the atmosphere and ocean communities together to an unprecedented degree. A similar convergence has been occurring at the land–atmosphere interface, between the meteorological and climatological communities on one side, and hydrology, ecology, geography, and related terrestrial sciences on the other.

This paper describes the history over nearly two decades of one research project at the land–atmosphere interface: the Global Soil Wetness Project (GSWP). GSWP was one of several initiatives that pioneered cross-disciplinary studies at the land–atmosphere interface. The niche of GSWP is designated in its acronym. It is global in scale, addressing the needs of global weather and climate models to have initial or boundary conditions of the land surface state. Soil wetness is the most important of these land states on intraseasonal-to-interannual time scales, although GSWP addresses all surface components of the energy and water budget. The central concept of GSWP is that soil wetness contains “memory” of weather and climate—a kind of inertia in the land surface analogous to the inertia that heat content in the ocean provides to the climate system, extending atmospheric predictability from 1–2 weeks to months or seasons.

The terms soil moisture and soil wetness are often used interchangeably (Dirmeyer 2004). In this review, we will be more precise. Soil moisture will refer to the water content of the soil expressed in terms of mass per unit area or a depth of liquid water contained in the soil column (because liquid water is essentially incompressible and has an essentially constant density). This may mean...
actual water, such as that baked out of the soil during the gravimetric measuring process, or the water that is the time-integrated residual of the water mass balance in a numerical model of the soil column. Soil wetness, on the other hand, is a dimensionless index and can be expressed in terms of a range defined arbitrarily between two extremes. The extremes may be bounded by zero water content, the wilting point, field capacity, saturation, or some range of radiances, conductances, or other thermal, electrical, electromagnetic, or nuclear scale. The conversion between soil moisture and soil wetness is often a simple linear relationship, but in some cases it is non-linear, or ill-defined—only slightly more precise than classifying as “wet” or “dry.” However, in the absence of the rather arduous process of directly weighing soil samples taken from the field to measure soil moisture, soil wetness estimates are the only viable approach to quantifying soil water content. Models also often quantify soil wetness, generally as a simple linear rescaling of soil moisture based on soil model parameters like porosity, wilting point, and field capacity. It should be noted that under the definitions proposed here, some usages of the terms—for example, in the name of the Soil Moisture Ocean Salinity (SMOS) satellite mission—are misnomers. However, we will stick to those definitions in this paper except where used disparately by others in their studies, projects, or measuring missions.

Some measurements of soil wetness are made routinely, but they are sparse and unevenly distributed around the globe. Since soil wetness varies on spatial scales at least as fine as that of precipitation, and usually much finer because of the very small scale (as fine as a meter in the horizontal and centimeters in the vertical), variability in soil properties, and vegetation root structures (Rodriguez-Iturbe et al. 1998, 1999). It is difficult to capture the full spectrum of soil wetness with in situ measurements. Regular gravimetric measurements were made by state-operated soil moisture networks in the former Soviet Union, China, and Mongolia (Roboek et al. 2000). These measurements span decades for some stations, and provide sufficient coverage for some limited spatial and temporal analysis (Vinnikov and Yeserkepova 1991; Gao and Dirmeyer 2006). Measurements were taken several times a month, primarily in agricultural settings, but the stations are spaced dozens if not hundreds of kilometers apart, and cover only a small fraction of the globe.

More recently, automated methods of taking routine soil wetness (and temperature) observations have been deployed. The state of Illinois has been measuring soil wetness at depths down to 2 m at 19 stations since the early 1980s (Hollinger and Isard 1994). The Oklahoma Mesonet has deployed meteorological stations across nearly every county of the state of Oklahoma, and placed heat dissipation sensors at three or four depths to estimate soil water potential and volumetric water content (Illston et al. 2008). The U.S. Department of Agriculture has built a nationwide network of stations called the Soil Climate Analysis Network (SCAN) that uses electrical conductivity probes to measure soil wetness and temperature at depths down to 1 m (Seyfried et al. 2005). Recently, the global network of flux network sites (FLUXNET; Baldocchi et al. 2001) has added soil wetness sensors to its existing turbulent flux measurement towers so the water balance can be better monitored.

To provide more complete spatial coverage of soil wetness, remote sensing has been employed. Until recently, soil wetness signatures were extracted from data coming from instruments designed for other purposes. Early attempts relied on detecting the secondary effects of soil moisture on surface temperature via thermal sensors (e.g., Carlson et al. 1981; Wetzel et al. 1984). Because of the absorption and excitation properties of water in the microwave portion of the electromagnetic spectrum, attempts have largely focused in this band. The Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) is a passive-microwave radiometer that produces products that correlate with soil wetness in the top few centimeters of soil, but these products can give information only on temporal variations and not absolute values of soil moisture (Gruhier et al. 2008; Sahoo et al. 2008). Data from Special Sensor Microwave Imager (SMM/I) and Scanning Multichannel Microwave Radiometer (SMMR) orbital instruments have also been used to infer surface soil wetness (e.g., Paloscia et al. 2001).

Data from active microwave sensors from European Remote Sensing (ERS) scatterometers have been used to create an index of soil wetness (Wagner et al. 1999) that correlates well with in situ observations (e.g., Dirmeyer et al. 2004). Aircraft-borne passive-microwave sensors (e.g., Jackson et al. 2002), ground-based sensors (de Rosnay et al. 2006), and Tropical Rainfall Measuring Mission (TRMM) satellite measurements (Jackson and Hsu 2001; Bindlish et al. 2003) have also been used to estimate soil wetness. A dedicated soil wetness satellite mission has recently been launched (SMOS; Kerr et al. 2001; Drinkwater et al. 2009) and another is in the planning stage (Soil Moisture Active Passive (SMAP); Entekhabi et al. 2010).

All microwave-based remote sensing technologies for soil wetness suffer from the shortcoming that only a very thin surface layer can be measured, and then only when vegetation cover is sufficiently thin that water in the plant canopy does not obscure the ground from the sensor (Burke et al. 2003). An alternative approach, which can sample much more deeply into the soil, and also integrate across a larger footprint than other forms...
of in situ measurements, exploits properties of cosmic-ray neutron absorption and scattering by soil water (Zreda et al. 2008).

Each observational approach provides coverage that is incomplete in time or space (horizontal or vertical). Satellites can view the entire surface of the earth (between certain latitudes for polar-orbiting platforms) at a repeat frequency on the order of a few days, but they only detect soil wetness very near the surface. In situ measurements can effectively be made continuously and can sample much of the vadose zone, but the limited number of sites and the horizontal heterogeneity of precipitation and soil properties make these measurements only locally applicable.

The solution to this problem has been to use numerical models of the land surface, driven by global gridded meteorological analyses, to provide complete estimates in space and time. Independent observational datasets like those mentioned above are then used to validate the model estimates. This has been the paradigm of GSWP. This paper provides a historical review of GSWP, an overview of the scientific accomplishments of the project, and a sampling of other science and applications the project has enabled.

Section 2 describes research that led up to the development and implementation of GSWP. Section 3 gives an overview of the original pilot project. An interim regional project, the Rhône-Aggregation Intercomparison Project (Rhône-AGG), is described in section 4. Section 5 chronicles the Second Global Soil Wetness Project (GSWP-2). Section 6 contains a summary and brief discussion of the future.

2. Antecedents to GSWP

The need for gridded global fields of realistic soil wetness values has been recognized for many years. The primary customers for such data have been numerical modelers of climate and weather. Manabe (1969) was the first to predict soil moisture and snow cover as state variables within a general circulation model (GCM). He adopted the bucket model of land surface hydrology, treating each land surface grid box as a shallow pan 150 mm deep that could catch precipitation and allow it to evaporate back into the atmosphere. Precipitation that exceeded the capacity of the bucket became runoff and was lost to the land–atmosphere system. Once realistic treatment of the terrestrial water and energy cycles was incorporated into GCMs in the 1980s (e.g., Dickinson et al. 1986; Sellers et al. 1986), soil wetness became a full-fledged prognostic variable and needed to be initialized like other prognostic variables at the start of each simulation.

One of the earliest climatologies of soil wetness was the mean annual cycle estimated by Mintz and Serafini (1981) and Mintz and Walker (1993). They estimated soil wetness variations as the residual of a simple surface water balance relationship driven by global fields of precipitation and near-surface air temperature. Mintz and Walker (1993) cite two earlier attempts. Stone et al. (1977) used observed mean relative humidity over land as a proxy for “ground wetness,” specifying an ad hoc linear relationship between the two. Miyakoda et al. (1979), attempting to specify a more realistic land state than constant evaporation efficiency, derived a spatially varying soil wetness to use as β for the summer season. On the simple assumption that the only control on soil wetness is antecedent precipitation, an ad hoc relationship was derived relating summer soil wetness to mean precipitation during February–July. This relationship was found to underestimate soil wetness at high latitudes, where evaporative demand is low. However, it was an early attempt to incorporate a realistic distribution of land surface hydrologic states into a GCM.

The soil wetness estimates of Mintz and Serafini (1981) were used to constrain the soil wetness in the first multidecadal global reanalysis—the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (Kalnay et al. 1996). Subsequently, NCEP has produced long-term national and global soil wetness products (Huang et al. 1996; Fan and van den Dool 2004) using a “leaky bucket” model (runoff occurs during unsaturated conditions) driven by analyzed temperature and precipitation. There have been other multiyear gridded global datasets of soil wetness based on comparably simple treatments of the land surface (e.g., Schemm et al. 1992; Willmott and Matsuura 2001) or the application of a second-generation land surface model (LSM; with complete closure of surface energy and water budgets) driven by meteorological analyses (e.g., Schnur and Lettenmaier 1997; Dirmeyer and Tan 2001; Berg et al. 2003).

Liston et al. (1993) pioneered a key technique that was later used by the GSWP data production efforts as well as by some of the LSM-based soil wetness datasets mentioned above (e.g., Dirmeyer and Tan 2001; Berg et al. 2003). Lacking global observed daily precipitation, they combined 10 years of observed gridded monthly precipitation data with one year of daily precipitation data from the First Global Atmospheric Research Project (GARP) Global Experiment (FGGE). The daily data were used to provide frequency of occurrence and intensity information, but were scaled so that the monthly totals equaled the observed values.

A key motivation for the inception of GSWP came from the International Satellite Land Surface Climatology Project (ISLSCP). ISLSCP pioneered the synthesis of observations across many scales and platforms, from
in situ measurements to satellite remote sensing, in the First ISLSCP Field Experiment (FIFE; Hall and Sellers 1995). A groundbreaking data synthesis effort by ISLSCP was its Initiative I (Meeson et al. 1995; Sellers et al. 1996), which is a collection of global 1° latitude and longitude gridded datasets of soils, land cover, hydroclimate, and radiation covering the two-year period 1987/88. ISLSCP Initiative I data were distributed as a set of five CD-ROMs that was extremely successful, with over 13 000 copies of the data disseminated to scientists, educators, and others. However, after the conception of the data initiative, it was decided by the ISLSCP Science Panel that a global modeling effort should be designed to demonstrate the usefulness and applicability of the forthcoming data (International GEWEX Project Office 1996).

A series of workshops were convened that led to the establishment of the original GSWP. The first, in August 1994, was motivated from a different direction: the need to address the problems faced by weather and climate modelers in initializing soil wetness states for numerical predictions and simulations (cf. Dirmeyer 1995). These problems included those listed in the introduction of this paper, centering on the lack of reliable observations over much of the world, but also included issues of model spinup and climate drift in simulated soil wetness fields, the inconsistency among LSMs in how soil wetness is represented, and the scientific merit of statistical methods for initializing soil wetness. The workshop reached two principal conclusions. The first was that there should be a comparison study of LSMs as they are used in GCMs—on a global grid—with consistent realistic meteorological forcing. This was in contrast to the early multimodel experiments of the Project to Intercompare Land Surface Parameterization Schemes (PILPS; Henderson-Sellers et al. 1993), which until that time were conducted at a series of individual sites. The second recommendation was to produce a global soil wetness climatology using one or more LSMs driven by internally consistent gridded near-surface atmospheric data, like that beginning to be produced by operational reanalysis projects.

The two recommended actions share many of the same steps in order to execute. This was realized at a workshop on global soil wetness three months later in Longmont, Colorado, which was sponsored by the Global Energy and Water Cycle Experiment (GEWEX). The recommendations served as the basis for a proposed research project to develop multiyear soil wetness time series at several climate modeling centers and academic institutions in conjunction with meteorological reanalysis efforts. An intercomparison of the time series was proposed for both the output from the LSMs and the input data. The name Global Soil Wetness Project was also suggested at this workshop.

3. The GSWP pilot project

The pilot project of GSWP, also known now as GSWP-1, was designed to use the ISLSCP Initiative I datasets to supply boundary conditions (near-surface temperature, humidity, winds, precipitation, downward shortwave and longwave radiation) and model parameters (vegetation-type distributions and soil and vegetation properties) to a selection of LSMs from operational centers and research institutions around the world. A thorough description of GSWP-1 can be found in Dirmeyer et al. (1999). A brief summary is presented here.

a. Project structure

Organizationally, GSWP was considered an initiative of ISLSCP and a part of GEWEX. However, it was also supported by, and closely aligned with, the Biospheric Aspects of the Hydrologic Cycle (BAHC; IGBP 1993) project of the International Geosphere Biosphere Programme (IGBP). GSWP-1 consisted of three components: the Production Group, the Validation Group, and the Intercomparison Center. The Production Group was composed mainly of modelers and had two main tasks: creation of a two-year global forcing dataset over land at 1° resolution from the ISLSCP Initiative I data and integration of LSMs driven by this forcing data to produce global 1° datasets of land surface fluxes and state variables (most notably soil wetness). The Validation Group was tasked with comparing the model results with available observations to assess model performance. The Intercomparison Center served as a collection and redistribution point for the model output, and performed output consistency checks and direct comparisons between models.

A major accomplishment of the Production Group was the creation of a global 6-hourly meteorological forcing dataset at 1° resolution. Much of the near-surface meteorology came from the 6-hourly operational analysis archive of the European Centre for Medium-Range Weather Forecasts (ECMWF). However, precipitation and radiation fields were produced by carefully combining low-time-resolution observationally based data with 6-hourly model diagnostic fields from reanalyses. Convective and large-scale precipitation fields from NCEP–NCAR reanalysis were rescaled so that the total precipitation for each month agreed with the monthly gridded observed precipitation from the Global Precipitation Climatology Project (GPCP; Huffman et al. 1997). NCEP–NCAR reanalyses were used because precipitation fields were not available from operational
ECMWF analyses. Furthermore, a rainfall frequency screening was applied following Liston et al. (1993) in order to improve upon the reanalysis statistics for days with measurable rainfall. Surface radiation fields were also hybridized, combining the time series of ECMWF estimates with monthly means from satellite-based monthly radiation fluxes (Darnell et al. 1992).

Ten modeling groups participated in GSWP-1. They are listed in Table 1. Computational power and data storage capabilities in the latter 1990s were limited enough that a set of proposed sensitivity studies were divided among the modeling groups, with each model used to investigate one or two specific issues.

The Validation Group’s efforts were conducted on three fronts. First, model-simulated fluxes and state variables were compared to in situ measurements, namely from field campaigns over single grid points. Of course, there were scaling issues, as the models represented the land surface at a resolution of about 10^4 km^2 while observations were representative of a much smaller area. Also, there were a limited number of observations available for the years 1987 and 1988. The second validation approach was more integrative, and took the simulated runoff from the various LSMs as input to a common global river routing model (Oki and Sud 1998), whose streamflow could then be validated against gauging stations. Finally, the predicted land surface states were used as initial or specified boundary conditions in GCM simulations to assess whether potentially more realistic soil wetness states improve climate simulations.

The Intercomparison Center devised a uniform format for LSM data submission and established the set of variables and output frequency in cooperation with the Production Group. Model output was checked to see whether conservation of energy or water was violated. Intermodel comparisons were made of the simulated global means and annual cycles of many variables.

A key aspect of GSWP-1 was that it covered two years that were marked by great contrasts in hydrology over many parts of the globe. Nineteen eighty-eight was a notorious drought year over the central United States, and had a failed monsoon over India. In contrast, 1987 was extremely wet over India and a rare “normal” year within a period of extended drought over the Sahel. There were also strong contrasts, with 1987 being much drier than 1988 during March–May over Alaska, South Africa, and Australia. Over parts of South America during austral winter, 1987 was abnormally wet and 1988 dry.

**b. Results**

Dirmeyer (1999, 2000), in boreal summer GCM simulations with prescribed observed SSTs, found that specifying GSWP soil wetness, compared to control simulations with predicted soil wetness initialized from operational analyses, improved the spatial correlations of rainfall over land. Simulation of the differences in rainfall over land between 1987 and 1988 were also improved. Furthermore, specifying soil wetness from the “wrong” year—that is, 1988 soil wetness with 1987 SSTs or vice versa—seriously degraded the simulations below the skill levels of the control simulations (Fig. 1). Each point in the figure represents the changes for a pair of ensemble members. This showed that the improvements were not merely coming from the correction of biases in the model soil wetness, but that there was useful information in the interannual variations of soil wetness. Douville et al. (2001) and Douville (2002) examined the impact of both global and regionally specified soil wetness from GSWP in a GCM, focusing on the African and Asian monsoon regions, and found different sensitivities.
to land surface states in the two areas. These studies hinted at the potential importance that a long-term soil wetness dataset could have for climate studies and even climate prediction if near-real-time analyses were made available. They also gave an early glimpse into the regional nature of climate sensitivity to soil wetness anomalies.

Validation of LSM output was performed in several ways. Entin et al. (1999) used in situ observations from what would later be called the Global Soil Moisture Data Bank (GSMDB; Robock et al. 2000) to validate all the LSMs over a wide swath of the former Soviet Union, Mongolia, and China, as well as India and Illinois. They found that there were large differences in the mean soil moisture, or preferred “operating range” of soil moisture among models across all locations, generalizing the earlier findings of Koster and Milly (1997), but when comparing only anomalies, models behaved much more similarly, reflecting some convergence in their simulations of evaporation and runoff. In fact, the coefficient of variation among models in terms of their global mean evaporation was 0.086 and for runoff was 0.173 (Dirmeyer et al. 1999). The NCEP LSM was validated over Illinois (Hollinger and Isard 1994) with an eye toward future operational hydrologic applications of the model by Chen and Mitchell (1999). Matsuyama and Nishimura (1999) validated both the forcing data and Japan Meteorological Agency (JMA) Simple Biosphere Model (SiB) output over the FIFE site, elucidating the problems involved with area-averaged versus point precipitation as an input to an LSM. Zhang et al. (1999) tested the SiB2 model in this framework.

As mentioned earlier, a key validation variable is runoff. Koster et al. (1999) compared the annual runoff from the LSMs over selected river basins to observations and a semiempirical formulation by Budyko (1974), finding LSM errors to be of the same range as the simple estimate. The unimpressive performance by the relatively sophisticated LSMs could be a result of shortcomings in the models themselves, or errors in the meteorological data and parameters supplied to the models. Oki et al. (1999) was able to zero in on error sources by routing each LSM’s runoff through a common river network and comparing the simulated annual cycles for river discharge with observations. Errors were found to be well correlated with the density of the rain gauge network used by GPCP to formulate their gridded observed monthly precipitation estimates (Fig. 2). River basins with a high density of gauges also showed the lowest errors for LSM runoff.

---

**Fig. 1.** Change from GCM control simulation (prognostic soil wetness) in the global (land only, north of 60°S) spatial correlation of 1988 minus June–August (JJA) 1987 precipitation between each pair of ensemble members (4 for each of the 2 years—16 pairs); x axis is the change from runs where soil wetness is specified from the GSWP-1 simulation by the same LSS, and y axis is change from runs where soil wetness from the opposite year was specified. Rays to the right of the origin indicate better skill when the correct soil wetness is specified throughout JJA, and rays downward indicate a decrease in skill when the wrong soil wetness is specified. Values inside the shaded square centered on the origin indicate changes that did not pass a two-tailed significance test at the 95% confidence level for either test case (after Dirmeyer et al. 1999).
simulations. Also evident in Fig. 2, individual LSMs consistently showed higher or lower errors in runoff despite gauge density, suggesting model parameterizations also play a role in the accuracy of the simulations. Finally, runoff biases were found to become increasingly negative with latitude. This result suggested precipitation, specifically snow, was underestimated in the GPCP dataset. Rodell et al. (2005) and Chen (2005) have also assessed the quality and shortcomings of the GSWP-1 model output.

The GSWP framework allowed several modeling groups to diagnose LSM deficiencies and develop improved parameterizations. Douville et al. (1999) found they could improve model runoff by a series of changes to the handling of the water cycle in the Interaction Sol–Biosphère–Atmosphère (ISBA) model. Problems in the simulated timing of spring snowmelt led to development and implementation of a new snow physics package for Simple SiB (SSiB) (Sud and Mocko 1999; Mocko et al. 1999). Boone and Wetzel (1999) found evidence that the use of aggregated soil hydraulic parameters at 1° resolution may lead to significant underestimation of runoff and overestimation of evapotranspiration. They implemented a subgrid heterogeneous treatment of soil parameters in the Parameterization for Land–Atmosphere–Cloud Exchange (PLACE) LSM to try to account for the strong nonlinearities in soil hydraulic formulations common to single-column soil schemes.

Pitman et al. (1999b) performed a sensitivity study comparing the contemporary modeling practice of assigning vegetation parameters [in this case leaf area index (LAI)] as a characteristic solely of vegetation type to instead using the ISLSCP Initiative I spatially varying properties derived from satellite observations. They found that at that time, the impact on the surface energy and water balance was well within the range of uncertainty for the global LAI dataset. Morrill et al. (1999) identified a problem in the forcing data—a phase error in downward longwave radiation of 6 h. Both Dirmeyer and Zeng (1999) and Douville et al. (1999) focused on the translation of gridscale precipitation into realistic infiltration and runoff within LSMs, finding that as parameterized subgrid variability of precipitation or soil increases, so does runoff, providing a tuning parameter for LSMs used in hydrologic applications.

Douville et al. (1999) examined a prototype land data assimilation system based on the technique of sequential optimal interpolation using iterative correction of simulated near-surface temperature and humidity to produce a soil moisture analysis. A form of this method was then applied at ECMWF for operational analysis and as part of their reanalysis procedure. Dirmeyer et al. (2000) used the GSWP-1 model output to assess the relationship between surface heat fluxes and soil wetness across several LSMs. GSWP-1 also illuminated the extent to which the soil moisture relaxation in the NCEP–NCAR reanalysis was leading to unrealistic soil moisture variability (Chen and Mitchell 1999)—one of several factors leading to the reexecution by NCEP of their reanalysis for intercomparison applications (Kanamitsu et al. 2002). Dirmeyer (2005, 2006) used GSWP-1 states to initialize seasonal climate model studies evaluating the role of land surface feedbacks on climate.

c. Science enabled by GSWP-1

GSWP-1 parameters and forcing data were used outside of the original Production Group to aid in the development and validation of LSMs (Douville 1998; Ducharne et al. 2000; Kanae et al. 2001; Mocko and Sud 2001; Chapelon et al. 2002; Liang et al. 2003; Devonec and Barros 2003; Gedney and Cox 2003; Gedney et al. 2004; Chen 2005; Slater et al. 2007). The methodology of GSWP-1—that is, the production of hybrid forcing datasets and integration of LSMs in an uncoupled or offline configuration—has been widely applied (Pitman et al. 1999a; Nijssen et al. 2001; Dirmeyer and Tan 2001; Milly and Shmakin 2002; Saleem and Salvucci 2002; Berg et al. 2003, 2005; Hirabayashi et al. 2005; Ngo-Duc et al. 2005; Qian et al. 2007; Weedon et al. 2010) and became the basis for the first Land Data Assimilation Systems (LDASs; e.g., van den Hurk et al. 2002; Cosgrove et al. 2003; Mitchell et al. 2004; Rodell et al. 2004). Furthermore, GSWP-1 data were used to help develop algorithms for the retrieval of large-scale fluctuations in subsurface hydrology from satellite gravity measurements (Rodell and Famiglietti 1999, 2002) as part of global ocean mass assessment (Cazenave et al. 2001), crustal motion projections (Mangiarotti et al. 2001), and as a land surface analysis used for comparison and validation in other.
studies (e.g., Georgakakos and Smith 2001; Nakaegawa et al. 2003).

In summary, the pilot phase of GSWP proved the viability of producing global land surface analyses, confirmed the value of information on year-to-year variations of the land surface state for climate, demonstrated the importance of high-quality observations to produce realistic forcing for land surface models, and showed that calibration, validation, and model development need not only be conducted at individual stations, but should also be done across large spatial scales. Furthermore, innovations in datasets and modeling were tested that have led to routine remote sensing of land surface parameters, near-real-time land surface analyses, and interest in the potential predictability of climate that can be contributed by knowledge of the land surface state.

4. The Rhône Aggregation experiment

GSWP-1 introduced an experimental design that is useful for comparing, testing, and validating LSMs in a spatially distributed framework. Experiments in this structure at the river basin scale were conducted as part of PILPS (Wood et al. 1998; Liang et al. 1998; Lohmann et al. 1998; Bowling et al. 2003). The PILPS study over the Red-Arkansas River basin of the United States, Phase 2(c), was conducted at the same 1° resolution as GSWP. Phase 2(e), over the Torne and Kalix basins of Scandinavia, was at 1/8°. This raised a question—what is the effect of spatial resolution on the uncoupled simulation of land surface state variables and fluxes with LSMs?

Rhône-AGG (Boone et al. 2004) was designed to investigate how the simulation of the surface hydrologic cycle varied among LSMs and across spatial resolutions. The project took advantage of high-resolution atmospheric forcings, parameters, observations, and a hydrologic model for the Rhône basin of southern France (Habets et al. 1999). Ultimately, 19 land surface schemes (LSSs) participated (Table 2), simulating three years of water and energy budgets at horizontal resolutions ranging from 1° to 8 km. The Rhône basin is interesting because it contains a range of altitudes, climates, and vegetation, ranging from alpine to semiarid and coniferous and deciduous forests to several types of crops. The basin contains over 1500 precipitation gauges for generating forcing data. Validation was against daily river discharge measurements from over 145 river gauges and snow depth measurements at 24 sites. As with GSWP-1, the runoff from all LSSs was routed through a common routing scheme—in this case the Modél Couplé de l’École des Mines de Paris (MODCOU; Ledoux et al. 1989).

As with GSWP-1, models simulated a range of mean soil wetness for any specific location, but with the mean removed, the annual cycles were in good agreement.

### Table 2. Land surface models participating in Rhône-AGG.

<table>
<thead>
<tr>
<th>Modeling group</th>
<th>Model(s)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meteorological Service of Canada</td>
<td>CLASS</td>
<td>Canadian Land Surface Scheme</td>
</tr>
<tr>
<td>University of Washington</td>
<td>VIC</td>
<td>Variable Infiltration Capacity model</td>
</tr>
<tr>
<td>Centre National de Recherches Météorologiques (CNRM) and Météo-France</td>
<td>ISBA</td>
<td>Interaction Sol–Biosphère–Atmosphère</td>
</tr>
<tr>
<td>Russian Academy of Sciences</td>
<td>SPONSOR</td>
<td>Semi-Distributed Parameterization Scheme of the Orography-Induced Hydrology</td>
</tr>
<tr>
<td>COLA</td>
<td>SSIB</td>
<td>Simple SiB (COLA version)</td>
</tr>
<tr>
<td>Institute of Water Problems, Moscow</td>
<td>SWAP</td>
<td>Soil–Water–Atmosphere–Plants model</td>
</tr>
<tr>
<td>Centre for Ecology and Hydrology (CEH)</td>
<td>MOSES-PDM</td>
<td>Met Office Surface Exchange Scheme</td>
</tr>
<tr>
<td>University of Arizona</td>
<td>VISA</td>
<td>Probability-Distributed Moisture model</td>
</tr>
<tr>
<td>NCEP–EMC</td>
<td>Noah</td>
<td>CLM with the Versatile Integrator of Surface and Atmospheric processes</td>
</tr>
<tr>
<td>Macquarie University</td>
<td>CHASM</td>
<td>Chameleon Surface Model</td>
</tr>
<tr>
<td>ECMWF</td>
<td>ECMWF</td>
<td>ECMWF land surface scheme</td>
</tr>
<tr>
<td>Koninklijk Nederlands Meteorologisch Instituut (KNMI)</td>
<td>MECMW</td>
<td>Modified ECMWF land surface scheme</td>
</tr>
<tr>
<td>Institut Pierre Simon Laplace (IPSL)</td>
<td>ORCHIDEE</td>
<td>Organizing Carbon and Hydrology in Dynamic Ecosystems</td>
</tr>
<tr>
<td>NASA Seasonal–Interannual Prediction Project (NSIPP)</td>
<td>Catchment</td>
<td>Mesoscale Analysis and Prediction Scheme LSM</td>
</tr>
<tr>
<td>NOAA Forecast Systems Laboratory (FSL)</td>
<td>MAPS</td>
<td>Land Surface Process/Radiobrightness model</td>
</tr>
<tr>
<td>University of Michigan</td>
<td>LSP/R</td>
<td>SiB with explicit representation of urban landscapes</td>
</tr>
<tr>
<td>Kyoto University</td>
<td>SiBUC</td>
<td>ISBA as used in the High-Resolution Local Area Model (HIRLAM)</td>
</tr>
<tr>
<td>Instituto Nacional de Meteorología (INM), Numerical Weather Prediction Department, Madrid</td>
<td>ISBA-HIRLAM</td>
<td>Probability-Distributed Moisture model</td>
</tr>
<tr>
<td>Deutscher Wetterdienst (DWD)</td>
<td>DWD-MLSM</td>
<td>DWD Multilayer Soil Model</td>
</tr>
</tbody>
</table>
Basin-average partitioning of precipitation between evaporation and runoff showed 80% of the models fell within ±10% of observations. Most LSMS performed reasonably well when evaluated in terms of their routed runoff as compared to monthly observed discharge. However, there was a large variation in skill when the LSMS were validated at the daily time scale, suggesting that partitioning between surface and subsurface runoff and the treatment of subgrid runoff can have significant impacts on the simulation of daily river hydrographs.

An interesting separation was found among models in their ability to simulate snow cover, melt, and the runoff generated from the snowpack. The LSMS were categorized as having either a composite formulation for snow (the simplest), a bulk formulation (one layer but with full thermodynamic treatment), or as multilayer snow schemes. Errors decreased as the snow schemes became more complex (Fig. 3). Temporal correlations to observed snow and discharge were lower for composite schemes than for bulk or multilayer schemes, which were comparable. Composite schemes tended to have a negative bias in snow cover, and bulk schemes a slight positive bias, while the multilayer schemes had almost no bias. Spatial resolution was found to have a profound impact on simulations of snow in mountainous terrain, where resolving the peaks and valleys of the topography has major consequences for the ability of models to simulate the accumulation and melt of snow. The only model unaffected by resolution was a version of the Variable Infiltration Capacity (VIC) model that parameterized subgrid altitude bands (Hamlet et al. 2005), essentially capturing the critical orographic characteristic that was otherwise lost as horizontal resolution was reduced from 8 km to 1°.

5. GSWP-2

The pilot phase of GSWP contributed to many interesting scientific results and served as a “proof of concept” for integrating LSMS as stand-alone models on a global grid, but there was room for improvement. There were found with the forcing data (Morrill et al. 1999; Oki et al. 1999; Chapelon et al. 2002). Two years’ simulation showed the potential for study of interannual variations, and a desire grew to run for a longer period. A comparison of longer global soil wetness datasets spanning 8–20 years taken from reanalyses, offline land model simulations, and remote sensing retrievals verified that there was useful statistically independent information among the datasets (Dirmeyer et al. 2004). In GSWP-1 there was much discussion among the Production Team members about creating a single dataset synthesized from all model runs, but the consensus at the time was that it was not prudent. This would be revisited in GSWP-2.

FIG. 3. (top) River discharge error, (middle) correlation squared, and (bottom) mean bias for the Rhône-AGG participating models, organized by the complexity of their snow schemes (composite schemes have a single energy budget for all components in the grid box, bulk schemes have a separate energy balance for a single snow layer, and multilayer schemes have between 2 and 3 snow layers). Horizontal dotted bars indicate category means.
Planning for a second phase of GSWP began even before analysis efforts for GSWP-1 were completed. The idea was discussed at ISLSCP Science Panel meetings and BAHC workshops. Planning began in earnest at a meeting in Norman, Oklahoma in 2000 that was part of a more general workshop on soil wetness (Leese et al. 2001). Early goals included performing integrations over a longer period of time, removing variability among models caused by different land surface parameters, improving the quality of the meteorological forcing data, including alternative datasets for surface parameters and forcing for the performance of sensitivity studies, improving the methods of data collection and distribution, and producing an official “product” of land surface states and fluxes that would represent the best possible estimate of these quantities. Several other ideas were considered and eventually discarded, mainly for reasons of practicality. These included full consideration of the carbon cycle alongside the energy and water cycles and provision of meteorological data above the surface layer into the planetary boundary layer so that LSM fluxes would not be overconstrained.

a. Project structure

Again, GSWP was to be tied to a data action of ISLSCP—Initiative II (Hall et al. 2003). GSWP-2 would cover a 10-year period (1986–95) at the same 1° spatial resolution as in GSWP-1. All near-surface meteorology, soil, and vegetation parameters came from ISLSCP Initiative II. An essential quality of ISLSCP Initiative II was the production of multiple datasets for most gridded fields. This would allow for a myriad of sensitivity studies in GSWP-2. The science and implementation plan for GSWP-2 was presented in a GEWEX report that became the handbook for all participants (International GEWEX Project Office 2002).

A two-stage spinup process was used to minimize climate drift and to present a realistic state for the start of the 1986–95 core period. First, the models were started from an initial state on 1 July 1982 with global soil wetness set at 75% of saturation at all levels, soil temperatures set to the local mean June air temperature, and no snow cover. Each model was integrated over the first 12 months repeatedly until the modeler deemed that deep soil wetness and temperature had stabilized. Then each model was integrated forward through the remaining 2.5 years (July 1983–December 1985) to attain a realistic state of land surface anomalies at 1 January 1986. Meteorological forcing data for the entire 13.5-year period was generated and provided to all participants.

Great care was taken to construct hybrid meteorological forcing data that would be self consistent and as realistic as possible (Zhao and Dirmeyer 2003). The NCEP–Department of Energy (DOE) reanalysis (Kanamitsu et al. 2002) was combined with a number of in situ and remotely sensed datasets for precipitation (Rudolf et al. 1994; Huffman et al. 1997; New et al. 1999, 2000), humidity, temperature (New et al. 1999, 2000), and surface radiation (Stackhouse et al. 2000; Zhang et al. 2004). Undercatch of precipitation, especially of snow, was corrected based on a database of gauge aerodynamic properties from around the world (Motoya et al. 2002). Great care was also taken to ensure that during the re-gridding, meteorological data were not mapped from sea to land, since land–sea masks did not align perfectly among reanalysis and ISLSCP grids. The basic approach was the same as GSWP-1: 3-h reanalysis fields were scaled so their monthly means agreed with observationally based datasets. Additionally, temperature and surface pressure were adjusted to account for the differences in surface elevation between the 1° GSWP-2 grid and the reanalysis grid, and relative humidity was held constant during temperature adjustments. For sensitivity studies, regridged 40-yr ECMWF Re-Analysis (ERA-40) reanalyses were also used (Betts and Beljaars 2003). Vegetation distribution data came from Loveland et al. (2001) with an alternate dataset for sensitivity studies from DeFries et al. (1999), and vegetation radiative properties were derived from remote sensing (Los et al. 2000).

All input and output data were required to comply with the Assistance for Land Surface Modeling Activities (ALMA) data standard, which was developed for GEWEX by its Global Land–Atmosphere System Study panel (GLASS; Polcher et al. 2000). ALMA provided standards of variable names, units, and sign conventions, as well as Network Common Data Form (NetCDF) metadata, to aid the distribution and processing of datasets involved in land surface modeling. It has become the archetype for LDAS efforts as well as PILPS and other land model intercomparison endeavors.

An Intercomparison Center (ICC) was again set up at the University of Tokyo, where model output was submitted and basic quality control tests were performed. This is also the site where model output was shared. The forcing and parameter datasets were served from three sites—one in Japan, one in France, and one in the United States—using the Open-Source Project for a Network Data Access Protocol (OpenDAP) data sharing convention (Gallagher et al. 2007). The hope was that this would allow modeling groups to conduct their simulations without downloading the complete forcing dataset, but rather have it served directly to the executing model code via the Internet. However, bandwidth and network instabilities were found to remain limiting factors during the mid-2000s, and ultimately the modeling participants downloaded the datasets for local storage and execution.
Forcing data for a baseline simulation by all model groups was specified, but alternative datasets for many variables were prepared for use in sensitivity studies. A number of different precipitation datasets were constructed, as precipitation was deemed the most critical driving variable for soil moisture simulation. The default precipitation forcing included scaling for gauge undercatch and a two-stage renormalization to observations—one to the gauge-only based Global Precipitation and Climate Change (GPCC; Rudolf et al. 1994) dataset and a second to the blended gauge and satellite GPCP (Huffman et al. 1997) product where the rain gauge density from GPCC was below a critical threshold. Several of the sensitivity studies involved precipitation data where the GPCP scaling, gauge undercatch correction and GPCC scaling were sequentially removed. Also, a hybrid precipitation dataset based on ERA-40 precipitation was prepared. Alternative radiation datasets included versions where only the longwave or shortwave radiation were scaled to match observed monthly means, a version where reanalysis radiation was scaled to the International Satellite Cloud Climatology Project (ISCCP; Zhang et al. 2004) estimates instead of the Surface Radiation Budget (SRB; Stackhouse et al. 2000) data, and one where ERA-40 values were used. Also, datasets were prepared where no corrections were applied to any of the reanalysis data from either NCEP–DOE or ERA-40. Finally, several alternative sets of land surface parameters were prepared, including one with the University of Maryland land cover classes (DeFries et al. 1999) instead of the IGBP classifications prepared by the Earth Resources Observation and Science (EROS) Data Center (EDC; Loveland et al. 2001), and a dataset where the monthly varying vegetation parameters such as leaf area index, greenness, and vegetation cover fraction were replaced by a mean seasonal cycle. Some modeling groups also opted to conduct a LSM integration using their own default vegetation and soil grids. Ultimately, 15 models would be involved in the project. They are listed in Table 3.

### Table 3. Land surface models participating in GSWP-2.

<table>
<thead>
<tr>
<th>Modeling group</th>
<th>Model(s)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Texas</td>
<td>CLM2-TOP</td>
<td>The Community Land Model (CLM) with TOPMODEL hydrology</td>
</tr>
<tr>
<td></td>
<td>VISA</td>
<td>CLM with the Versatile Integrator of Surface and Atmospheric processes</td>
</tr>
<tr>
<td>COLA</td>
<td>SSiB</td>
<td>Simple SiB (COLA version)</td>
</tr>
<tr>
<td>NASA GSFC Climate and Radiation Branch</td>
<td>HY-SSiB</td>
<td>SSiB with improved hydrology</td>
</tr>
<tr>
<td>Kyoto University</td>
<td>SiBUC</td>
<td>SiB with explicit representation of urban landscapes</td>
</tr>
<tr>
<td>CNRM and Météo-France</td>
<td>ISBA</td>
<td>Interaction Sol–Biosphère–Atmosphère</td>
</tr>
<tr>
<td>NASA GSFC Hydrology Branch</td>
<td>Mosaic</td>
<td></td>
</tr>
<tr>
<td>NASA GSFC Global Modeling and Assimilation Office (GMAO)</td>
<td>Catchment</td>
<td></td>
</tr>
<tr>
<td>United Kingdom Meteorology Office (UKMO)</td>
<td>MOSES2</td>
<td>Meteorological Office Surface Exchange Scheme version 2</td>
</tr>
<tr>
<td>NCEP–EMC</td>
<td>Noah</td>
<td></td>
</tr>
<tr>
<td>Russian Academy of Sciences</td>
<td>SWAP</td>
<td>Soil–Water–Atmosphere–Plants</td>
</tr>
<tr>
<td>NOAA Geophysical Fluid Dynamics Laboratory (GFDL)</td>
<td>LaD</td>
<td>Land Dynamics</td>
</tr>
<tr>
<td>IPSL</td>
<td>ORCHIDEE</td>
<td>Organizing Carbon and Hydrology in Dynamic Ecosystems</td>
</tr>
<tr>
<td>University of Maryland</td>
<td>SLand</td>
<td>Simple Land model</td>
</tr>
<tr>
<td>University of Tokyo Institute of Industrial Science (IIS)</td>
<td>Bucket</td>
<td></td>
</tr>
</tbody>
</table>

### b. Data products

The main data product of GSWP-2 was a multimodel analysis (MMA) consisting of a simple arithmetic mean of 13 of the 15 participating LSMs [those that simulate complete closed surface energy and water budgets—simple land (SLand) and bucket were excluded]. The MMA was documented by Dirmeyer et al. (2006) and distributed on DVD and via the Internet. Gao and Dirmeyer (2006) had investigated strategies for creating a land surface MMA, and found that in the absence of in situ validation data at all but about 100 of the 15 325 grid boxes in the GSWP-2 dataset, more sophisticated regression-based methods of combining model output failed to produce better results than a simple mean. Nevertheless, the simple mean used in the GSWP-2 MMA was found to be superior to any individual model and better than other soil wetness estimates, including those from reanalyses, in reproducing temporal variations of observed soil moisture (Guo et al. 2007). Nevertheless, great care was taken to check each variable of every model for quality and consistency, and a number of minor problems, as well as few major problems, were found in individual LSMs. These were largely rectified while producing the MMA.
By simulating across a decade with more than a dozen LSMs, robust statistics on the global energy and water budgets could be estimated using the MMA (Dirmeyer et al. 2006). Figure 4 shows estimates of terms of these budgets, as well as snowpack water balance and components of evapotranspiration. The range of the annual cycle and interannual variability is also shown. Inter-model variability was used as a measure of uncertainty caused by the variations among parameterizations in LSMs. Dirmeyer et al. (2006) also presented a table listing every variable in the MMA and a quantification of the model-based uncertainties associated with them.

A study of the sensitivity of GSWP-2 LSMs to various meteorological forcings and surface parameter sets was conducted by Guo et al. (2006). Not every model performed every sensitivity test, so some conclusions are tentative. Using the GSMDB field measurements for validation, it was found that the hybridization of the forcing data greatly improved soil wetness simulations by LSMs. The impact on skill was greatest for variations among the precipitation forcings. Radiation had the second largest impact, and the choice of vegetation distribution was third. Most remarkably, the variations for a single LSM among all choices of forcing data were as large as that across all LSMs with the baseline forcing, suggesting that the magnitude of model uncertainty and forcing uncertainty are on par with one another. Wei et al. (2008a) examined sensitivity in general, as opposed to impact on skill, and found precipitation uncertainties easily dominate all other sources in terms of effect on soil wetness, with a particularly strong impact in semiarid regions, as shown in Fig. 5. Radiation uncertainties were found to have a large seasonal cycle in their impact, being more important during summer. Guo and Dirmeyer (2006) compared the participating LSMs and found a wide range of skill across models, with geographical variations as well.

Several authors have assessed GSWP-2 analyses. Because reliable estimates of evapotranspiration are particularly difficult to come by, data from GSWP-2 have been compared to other sources (Jiménez et al. 2009, 2011) applied in global water budget evaluations (Schlosser and Gao 2010) and regional assessments (Feng and Houser 2008). Decharme and Douville (2006) revisited the role of uncertainties in precipitation forcing on river discharge, much as Oki et al. (1999) had done in GSWP-1, and found the wind corrections for gauge undercatch in GSWP-2 were too severe. Materia et al. (2010) used
runoff from GSWP-2 models to compare precipitation uncertainty to intramodel uncertainty as a source of error for river discharge modeling.

c. Applications of GSWP-2

The GSWP-2 modeling methodology has been applied to problems in water resources (Islam et al. 2007; Hanasaki et al. 2008; Kim et al. 2009). Yang et al. (2011) found that a multiparameterization ensemble of a single LSM had the same superior performance compared to individual configurations as found in GSWP-2 with the MMA versus individual LSMS. Oki and Kanae (2004) and Hanasaki et al. (2010) have applied GSWP-2 data to estimate the worldwide trade of virtual water (water utilized in one nation for food or goods that are exported to another). Wei et al. (2008b) found evidence of a negative feedback between soil moisture variations and precipitation in certain situations using GSWP-2 MMA data. Dirmeyer et al. (2009) used GSWP-2 MMA estimates of soil wetness to determine soil moisture memory and correlations between soil wetness and evapotranspiration from GSWP-2 to establish regions of potential land state control of surface fluxes. Combined with independent estimates of atmospheric water recycling (Dirmeyer and Brubaker 2007), a synthesis of the global climatology of land–atmosphere interactions was derived. Koster et al. (2009) used GSWP-2 soil wetness states as initial conditions for GCM simulations to demonstrate the pitfalls of haphazardly applying soil wetness states from one LSM in another LSM without making accommodations for the differences in their climatologies. This stems from the problem that soil moisture or soil wetness is neither defined consistently among models nor across various methods of measurement (Dirmeyer 2004). Jeong et al. (2008) used the Community Land Model (CLM) simulations to initialize coupled CLM–Community Atmospheric Model (CAM) simulations and assess their effect. Teuling et al. (2009) used GSWP-2...
data to extrapolate relationships between surface radiation, evapotranspiration, and climate change impacts.

The forcing data from GSWP-2 have been used by a large number of researchers for a variety of applications, including model development (Gibelin et al. 2006; Decharme 2007; Decharme and Douville 2007; Balsamo et al. 2009; Dutra et al. 2010), hydrologic modeling (Oki and Kanae 2004; Islam et al. 2007; Decharme et al. 2008; d’Orgeval et al. 2008; Hanasaki et al. 2008; Biancamaria et al. 2009), model comparison (Essery et al. 2009; Boone et al. 2010), remote sensing retrieval testing (Gao et al. 2004; Balsamo et al. 2006), surface heat flux simulation (Alton et al. 2009), ecological (Yuan et al. 2010), and urban modeling (Menon et al. 2010). Land surface parameter data from GSWP-2 have also been used in other applications (e.g., Serreze et al. 2006; Gao et al. 2008; Kim et al. 2009).

Specific fields from the GSWP-2 MMA have been used widely for model comparison and validation (e.g., Teuling et al. 2006; Oleson et al. 2007; Biancamaria et al. 2008; Boé and Terray 2008; d’Orgeval et al. 2008; Wang and Liang 2008; Cao et al. 2009; Corti et al. 2009; Essery et al. 2009). Soil wetness fields from GSWP-2 have been used in GCM predictability studies to compare their impacts to the effect of sea surface temperatures (Conil et al. 2007) and snowpack (Douville 2009). The characteristics of the surface hydrologic cycle in arid regions were studied by Shen and Chen (2010) with the help of GSWP-2 analyses. Using GSWP-2 data, Zribi et al. (2010) derived a lagged relationship between soil wetness and a vegetation index observable from space that could potentially be used to predict vegetation states in semiarid regions using real-time remotely sensed soil wetness.

A follow up to the original GEWEX Global Land–Atmosphere Coupling Experiment (GLACE; Koster et al. 2004) was conducted to examine the impact of “realistic” land surface state initialization on retrospective forecasts in a multimodel framework. GLACE-2 (Koster et al. 2010) used GSWP-2 integrations of LSMs to supply initial states for spring and summer of 1986–95 to the GCMs participating in the experiment. It was found that realist initial soil wetness does increase the forecast skill of near-surface temperature and precipitation over random land surface initialization in certain areas of the globe out to 45–60 days. This skill increase is greatest when the magnitudes of the initial soil wetness anomalies are largest. There are, however, different responses to wet versus dry anomalies, and considerable variability among GCMs (Koster et al. 2011). Regional differences in the response exist; impacts are less over Europe than North America because of spatial variations in the potential predictability gain from soil wetness (van den Hurk et al. 2011).

6. Summary and future

The Global Soil Wetness Project has provided a framework for integrating, validating, and comparing LSMs and a tool to aid model development. Datasets from GSWP have been applied to a wide range of research applications, and the land surface analyses have filled a need that cannot be met by current observing systems. GSWP demonstrated the feasibility of producing global gridded analyses of the land surface using state-of-the-art land surface models. GSWP was able to show the primary sources of error and uncertainty in estimates of land surface state variables and fluxes and where they have their greatest impact. Precipitation measurements are the most crucial for forcing LSMs. GSWP also showed that an ensemble of different models could be used to reduce the errors caused by inadequacies in the LSM parameterizations themselves. This approach is now gaining wide acceptance.

Figure 6 is a demonstration of our progress in understanding the land surface’s role in climate, as well as an indication of what is still to be done. The idealized curve from the empirical relationship of Miyakoda et al. (1979), discussed in the introduction, has been plotted here. On top is shown a scatter of the gridbox values of mean July soil wetness as a function of antecedent mean precipitation from the GSWP-2 MMA. The idealized estimate captured the most salient feature—increasing soil wetness with increasing precipitation. Miyakoda et al. (1979) recognized that their idealization failed in energy-limited domains, such as the high latitudes, which is evident in Fig. 6 by the cluster of points in the upper-left corner of the plot. The scatter seen is symptomatic of the fact that many factors control the distribution and variation of soil wetness, and current models, despite many sophisticated parameterizations of relevant physical processes, actually remain more deterministic than observations suggest the real world to be. Clearly, LSMs do not include all of the salient processes involved. Newer schemes that explicitly model the relationship between the carbon and water cycles, which meet primarily in the biological processes of plants and bacteria, are more realistic than most of the LSMs that participated in GSWP.

Today there are regional and global land data assimilation system (LDAS) efforts that perform much the same data production function as GSWP, but in a context that is closer to operational than research. These apply multiple LSMs in a framework of common grids and forcing datasets, using operational observing networks and meteorological analyses, to update the land surface data to within days or hours of real time. Software systems have been developed to facilitate running multiple models in this parallel mode; for example, the
Land Information System (LIS; Peters-Lidard et al. 2007). Thus, the modeling paradigm of GSWP is transitioning from research to operations. Going forward, there have been calls for a much longer dataset than the 10 years of GSWP-2, and for the data to be generated at a higher spatial resolution—0.5° or finer. The usefulness of longer datasets is obvious—more climate variations and case studies can be examined, and statistical confidence in the simulated climatology is increased. Furthermore, there exists the potential to capture the climate change signal that would emerge ever more clearly as the period is extended. There exist a number of single-model datasets of this type (e.g., Fan and van den Dool 2004). However, GSWP demonstrated the benefits to having a multimodel land surface analysis approach.

As duration, resolution, and the number of models increase, the appetite for computing and networking resources grows accordingly. A centralized approach to computing and data storage seems advantageous as potential participants begin to encounter resource limitations at their own institutes. Resolution and duration of integrations may increase faster than storage and network bandwidth, the latter being a particular problem for international projects. This would require a fundamental change to the execution of such a project. No longer would forcing data be distributed to individual modelers, who then send their results to a collection point after integrating their LSMs. Rather, modelers would submit their model code to a central computing center where all data reside, and computational cycles would be dedicated to the project. There are obviously benefits to the existing distributed approach as well. The development of universal couplers (e.g., Polcher et al. 1998; Best et al. 2004) allow modelers to set up and easily link their LSMs to forcing and output data streams for testing and execution. Centralizing the infrastructure could hinder the learning and benchmarking process. In any such arrangement, modelers should have first access to the results from their own models.

To make any new phase of such an experiment more than an engineering exercise, there need to be new and interesting research topics that can be addressed. One potential direction is to include the carbon cycle in the experiment and shift emphasis to those LSMs with an explicit treatment of carbon and other nutrients. Land surface modeling has evolved greatly since Manabe’s bucket. It has progressed through so-called second-generation models with explicit closed energy and water budgets to third-generation models that include carbon and nutrient cycles (Sellers et al. 1997). Biomass, vegetation phenology, plant respiration, carbon stores, biogeochemistry, and land use change could be added to the scientific topics explored—topics that have so far been dominated by surface hydrology. The study of linked responses between the carbon and hydrologic cycles in a multimodel framework would be an important new direction for research. Such an approach would also make any new project more relevant to the study of climate change. In fact, there could be interesting frontiers of research in terrestrial climate by using output from the GCMs used in future climate projections for the Intergovernmental Panel on Climate Change (IPCC).
assessment reports. Time-slice experiments (several decades from the late twenty-first century compared to late twentieth century) in the GSWP framework could be used to study terrestrial climate change impacts, the effect of the range of uncertainty among projections, downscaling of results, and a number of other land surface studies that remain too computationally expensive to execute in the coupled ocean–atmosphere–land modeling context of IPCC. Whatever direction future efforts take, it is important that science remain the prime motivation.

Acknowledgments. Bart van den Hurk and an anonymous reviewer made many helpful comments during the review process. Many people have contributed to the evolution and success of GSWP since its inception in the early 1990s. I would like to acknowledge several individuals who played crucial roles. First and foremost I thank Piers Sellers, with whom it all started within ISLSCP. Nobuo Sato and Han Dolman were seasoned mentors to me as a greenhorn recent Ph.D. in whose lap this international project landed. Thanks to the GEWEX IGPO in the persons of Rick Lawford, Paul Try, Sam Benedict, and through it all Dawn Ehrlich. Pavel Kabat provided much support from the IGBP side, and later as head of ISLSCP after Piers Sellers left to follow his aspirations. Taikan Oki became my copilot as GSWP-1 advanced into GSWP-2. Aaron Boone, GSWP-1 participant and leader of the Rhône-AGG experiment, has always been a strong supporter. Jared Entin wrote his Ph.D. thesis on soil moisture analysis and GSWP-1, and later as a program manager at NASA helped to support GSWP-2. Among many of my colleagues at COLA, Xiang Gao was invaluable for bringing GSWP-2 to fruition and making the MMA such a good product. Others who made substantial contributions to GSWP while at COLA include Fanrong Zeng, C. Adam Schlosser, Mei Zhao, Zhichang Guo, and Jianfeng Wei. Jan Polcher, as first chair of GLASS, gave GSWP an organizational and scientific home, and the GLASS Panel has been supportive and encouraging throughout. Finally, thanks to J. Shukla and Jim Kinter for being always supportive of GSWP and providing any COLA resources the project needed.

REFERENCES


DeFries, R. S., J. R. G. Townshend, and M. C. Hansen, 1999: Continuous fields of vegetation characteristics at the global scale at 1-km resolution. *J. Geophys. Res.*, **104**, 16 911–16 923.


simulations of the boreal summer climate. *J. Hydrometeor.*, 9, 1231–1248.


Motoya, K., K. Masuda, K. Takata, and T. Oki, 2002: Sensitivity of precipitation gauge correction for the estimation of the global


