A heterogeneous land surface model initialization study

Roshan Shrestha[1] and Paul Houser[2]

Received 22 September 2009; revised 27 May 2010; accepted 17 June 2010; published 7 October 2010.

[1] Accurate energy and water storage initialization is critical for skillful land surface model prediction. A properly initialized model has equilibrium land surface states via spin-up runs. The present study investigated aspects of spin-up runs for the Noah Land Surface Model (LSM) at point stations using the 1 km grid of the Land Information System. The model is run repeatedly through a single year until a predefined equilibrium is achieved. Nine different model initialization methods were tested and compared in these spin-up runs at twelve southern Great Plains surface observation stations in the Midwestern United States. Soil moisture is used as the primary land surface state to evaluate the spin-up initializations. The model outputs are compared to evaluate the discrepancies over the annual cycle of repeated runs to learn how the model attains equilibrium land surface states. The results indicate that the climatological average state does not necessarily return the most efficient LSM initialization. Among the various tested methods, the spin-up runs initialized with the spatially heterogeneous states averaged over a short period are found to perform better than others. Heterogeneous land surface conditions are also found to play a vital role in the spin-up response. More stable land surface states are obtained through longer spin-up runs, which also produce more similar coefficient of variance, suggesting that the longer spin-up runs could yield similar heterogeneous fluxes irrespective of the initialization method. Comparing the results from different methods, a computationally economic technique for the single-year spin-up is proposed.

Citation: Shrestha, R., and P. Houser (2010), A heterogeneous land surface model initialization study, *J. Geophys. Res.*, *115*, D19111, doi:10.1029/2009JD013252.

1. Introduction

[2] The exchange of heat, moisture, and momentum between the land surface and the atmosphere is highly dependent on land surface processes. Variations in surface characteristics (e.g., topography, vegetation, soil, etc.) in a chosen modeling domain are often large, which make accurate land surface process modeling a challenging task. The philosophy of testing land surface models by forcing them with observed meteorological data, and then comparing the fluxes generated by the models with observations, has been practiced by many validation and intercomparison studies; for example, phase 2 of the Program for Intercomparison of Land surface Parameterization Schemes (PILPS) [*Henderson-Sellers et al.*, 1993, 1995]. However, the problems in the process of a model adjustment to forcing fields (i.e., model spin-up) may invalidate such a test [*Robock et al.*, 1995; *Koster et al.*, 2004]. Owing to the persistence and memory of a land surface model's state, it requires proper initialization to provide a skillful simulation of land surface fluxes and avoid erroneous interpretation of the model results. The spin-up problem can severely bias land surface simulations, and if not properly recognized, could potentially invite erroneous understanding of the land surface processes and degrade the value of initiation fields in coupled land-atmosphere modeling, therefore compromising associated weather and climate simulation skill. It has been shown that subsequent validation results are largely affected by the methods chosen to initialize a land surface model. The role of model initialization is critical in the prediction of land surface conditions that retain the memory of the land surface state for a considerably long period.

[3] There is neither a standard choice nor any scientifically established optimal technique for spinning up a land surface model. *Yang et al.* [1995] found that most land surface schemes require many years to come to thermal and hydrologic equilibrium with the forcing meteorology; the time needed depends on the total moisture-holding capacity and the moisture store initialization. *Cosgrove et al.* [2003] compared the behavior of several LSMs and found that the spin-up times showed a large spatial variation. The spin-up times were correlated most strongly with precipitation and temperature. *Rodell et al.* [2005] compared ten methods for initializing a land surface model and concluded that when multiple years of forcing data are not available, one of the best approaches is to use climatological average states

[1]Center for Research on Environment and Water, Institute of Global Environment and Society, Calverton, Maryland, USA.
[2]Department of Geography and Geoinformation Science, George Mason University, Fairfax, Virginia, USA.

0148-0227/10/2009JD013252

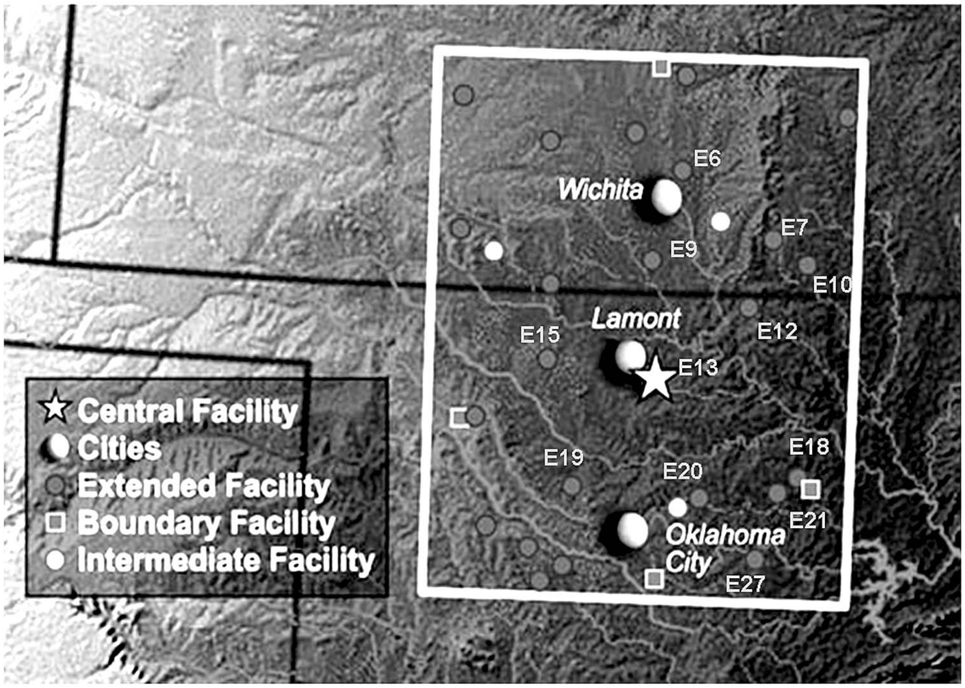

Figure 1. Map and location of stations.

derived from the same model for the time of year of initialization. *de Goncalves et al.* [2006] evaluated spin-up characteristics of the Simplified SiB (SSiB) model [*Xue et al.*, 1991, 1996] in the climate of South America and reported regional dependency of spin-up time on precipitation regime. Spin-up times in the study of *de Goncalves et al.* [2006] are noticeably less than those found by *Cosgrove et al.* [2003] which the authors have argued is related to the more abundant precipitation and denser vegetation cover.

[4] One of the most commonly employed methods for LSM initialization is to start from an initial guess or observation. Use of synthetic data is another widely employed method, which is often obtained from a long-term simulation of the same or different model. The primary objective of the synthetic run is to obtain bias-free land surface states suitable for model initialization. However, it increases the dependency of the spin-up experiments on the choice of synthetic model set up. Other widely used methods for initializing a LSM are generally based on blind guesses. In this study, we use the conventional blind guess and in situ observation for model initialization assuming that the observed data should provide better land surface initialization than synthetically prepared data from a predefined set of model runs. The motivating factor of the study is to investigate how the spin-up results are affected when a LSM is run at a point station using a set of well-observed high-resolution in situ data. Considering that the land surface spatial heterogeneity plays a vital role in the LSM initialization, this study also investigates how the model initialization responds to the land surface heterogeneity.

[5] This study evaluates the spin-up characteristics of the Noah land surface model at point stations using the 1 km grid of the LIS. The stations are located at the Atmospheric Radiation Measurement (ARM) site in the southern Great Plains (SGP) of the Midwestern United States (see Figure 1). The simulations are confined to running over a single year repeatedly. Looping over a single year or a fixed period is the most commonly employed method to spin-up a model, particularly if the available data is limited (as in this study), which is collected during an enhanced observation period (details on the data are in section 3).

2. LIS Noah Model

[6] The LIS provides a comprehensive software framework to integrate a various community of land surface models, observations and the necessary computing infrastructure [*Kumar et al.*, 2006, 2008]. The LIS architecture is designed to allow interoperability of land surface models, meteorological inputs, land surface parameters and observational data. It can plug-in various land surface models such as the Noah LSM, Community Land Model (CLM), Variable Infiltration Capacity (VIC) and Catchment LSM. The LIS enables running models on points, regions or the globe at a range of spatial resolutions (2.5 degree to 1 km). The LIS is also compliant with the Earth System Modeling Framework (ESMF; see *Hill et al.*, 2004] and Assistance for Land Surface Modeling Activities (ALMA) (see "The ALMA data exchange convention," available at http://web.lmd.jussieu.fr/~polcher/ALMA/) standards. With this flexibility, the LIS allows for testing of a variety of configurations with multiple land surface models and compares the impacts of the choice of models and data sets [e.g., *Kato et al.*, 2007]. With the opportunity of testing land surface models in various ways, the LIS helps improve our understanding on the land surface processes.

Table 1. Specifications of the LIS-Noah Model

Parameter	Value
Spatial resolution	0.01° × 0.01°
Model time step	15 min
Model output time step	30 min
Time period	2003–2004
Forcing source	SGP-ARM in situ observation
Soil layers	10, 30, 60, and 100 mm

[7] The Noah LSM [*Chen et al.*, 1996; *Koren et al.*, 1999] evolved in a series of modeling experiments conducted to simulate land surface process since the early 1990s. Recent versions of the Noah LSM is the result of broad partnership among the Office of Hydrological Development (OHD) of the National Weather Service, National Environmental Satellite Data and Information Service (NESDIS), National Aeronautics and Space Administration (NASA), National Center for Atmospheric Research (NCAR), the U.S. Air Force, Oregon State University (OSU) and other universities. The Noah LSM has been widely involved in studies coupling LSM with atmospheric models and many other experimental and operational simulations. For example, the National Oceanic and Atmospheric Administration-National Center for Environmental Prediction (NOAA-NCEP) executes the Noah LSM as the land component in their climate models. *Ek et al.* [2003] gives a brief review of the evolution of Noah LSM and its broad partnership among various participating organizations.

[8] The Noah LSM is a 1-D column model that can be used in either coupled or uncoupled mode. This model uses longwave radiation, shortwave radiation, precipitation, surface wind, and humidity as the forcing input to simulate soil moisture (both liquid and frozen), soil temperature, skin temperature, snow depth, snow water equivalent, canopy water content and land surface energy and water fluxes. The governing equations of the physical processes of the soil-vegetation-snowpack medium in the Noah LSM are integrated by the finite difference spatial discretization methods and a Crank-Nicholson time-integration scheme. This model continues to benefit from a steady progression of improvements in performance, both in an offline mode (that is, atmospheric-forced LSM-only runs for specific sites or in two-dimensional horizontal land surface domains), as well as coupled in fully three-dimensional operational mesoscale analysis and forecast systems [*Betts et al.*, 1996; *Ek et al.*, 2003]. (See Table 1.)

3. SGP ARM Data

[9] Data used for forcing the model and comparison with the spin-up states were obtained from the SGP ARM surface observation stations. The data observation stations in the SGP ARM site are equipped with a large existing network of weather and climate research and instrumentation. The instruments are configured for automatic data collection and the data is archived on the site data system. The SGP ARM site is located in relatively homogeneous geography but has a wide variability of climate and surface flux properties with large seasonal variation in temperature and specific humidity. This makes the site data sets ideal for testing land surface models and sufficient to resolve land surface interaction.

[10] The SGP ARM observation stations are included in the Coordinated Enhanced Observation Period (CEOP) reference data set. CEOP requires each reference site to undertake a quality check to ensure high-quality control. The data collected during CEOP Enhanced Observation Period (EOP)-3 (1 October 2002 through 30 September 2003) and CEOP EOP-4 (1 October 2003 through 30 December 2004) are available from CEOP Data Archive (CDA) in the form of 30 min resolution composite data that are organized into four different components (1) surface meteorological and radiation, (2) fluxes, (3) soil temperature and soil moisture, and (4) meteorological observations. The data sets also contain quality-control information, which can be used to mask out any suspicious or bad observations. The forcing inputs of the SGP ARM region (e.g., precipitation, surface air temperature, downward solar and longwave radiation, surface wind and relative humidity, etc.) are used to resolve the land surface interaction at the stations.

[11] The CDA does not implement or provide a gap-filling procedure. The gaps in SGP ARM data are very few, however, they might interrupt running the LSM smoothly. In order to avoid possible interruptions owing to missing data, the gaps in the forcing input are filled by interpolation.

4. Experiment Design and Configuration

[12] Several research questions motivate the experimental design. The first is to investigate whether the LSM spin-up period could be universal regardless of the variables we consider. Defining homogeneous initial states may be a good way to start, but there may be better LSM initialization methods. Is there a way to incorporate the spatial and temporal heterogeneity into the initialization? Are mean states for the given time of year appropriate for initialization or does the LSM respond indifferently throughout a year? Can we use a coarser resolution data, averaged over a larger domain for efficient initialization or do we need to have data from each station separately?

[13] Soil moisture is an essential initial condition for the Noah LSM. Other variables used to initialize the model are soil temperature, canopy water content, snow depth and water-equivalent snow depth. Among these initialization variables, the soil moisture fluctuation is relatively slower and exhibits consistent response to the dynamics of water and energy balance than others. Also, the subsurface soil layers holding the moisture respond slower to the changes of other land surface states causing a longer spin-up period. So, the soil moisture state is used as a surrogate variable for LSM initialization to test the spin-up performance.

[14] The experiment consisted of 108 simulations in 12 different stations, which were performed with identical specifications except the initialization method. Each model spin-up simulation is set by forcing the model with a year of data and a default set of initialized states. The model is repeatedly run for 7 years by restarting the states simulated in the previous year. This choice is preferable when long-term forcing data are not available for spin-up as it allows having spin-up runs for several years via looping a single year until the desired equilibrium level is achieved.

[15] The heat flux and soil moisture values simulated from the land surface model are used to compare the results. The results are analyzed by comparing the land surface fluxes

Table 2. Specifications of Experimental Simulations Key Initialization Techniques

Simulation	Specification	Technique
A	dry: soil at 10% saturation	blind guess
B	wet: soil at 70% saturation	blind guess
C	average: soil at 30% saturation	blind guess
D	direct insertion of observation	in situ
E	direct insertion of station averaged observation	in situ
F	direct insertion of annual average at a station	in situ
G	direct insertion of annual average at all stations	in situ
H	direct insertion of monthly average	in situ
I	monthly shift of spin-up start time	in situ

obtained in repeated simulations looped on an annual basis. The results are analyzed in two ways to understand (1) the sensitivity of different land surface fluxes to the spin-up test, and (2) the sensitivity of different initialization methods. By doing this, the analysis attempts to understand the controlling factors in land surface model spin-up. (See Table 2.)

[16] The different initialization methods are based on how the soil water contents are varied in the initialization. In all spin-up simulations, the soil temperatures were set to 290 K, and all other state fields (i.e., canopy water content, snow depth and water-equivalent snow depth) were initialized to zero. These settings limit the study to the sensitivity of initial soil moisture state and keeping other initial land variables constant. While involving more than one land variables in the spin-up evaluation is an option to investigate whether sensitivity of multiple land variables may provide a further enhanced spin-up performance, the option is not carried out in this study. Since the soil moisture acts as a surrogate to other land variables and the other initialization variables have shorter memory compared to the soil moisture, it is likely that the improvement involving sensitivity of more land variable is insignificant. All together nine different methods of initialization are used in the Noah land surface model, which is described below.

[17] Method A, dry start: Soil water content in all four layers was set equal to 10% of saturation. This technique would be suitable if the main factor controlling spin-up time was drying of the soil to equilibrium conditions.

[18] Method B, wet start: Soil water content in all four layers was set equal to 70% of saturation. This technique would be suitable if the main factor controlling spin-up time was wetting of the soil to equilibrium conditions.

[19] Method C, average moisture start: Soil water content in all four layers was set equal to 30% of saturation (an average value). This is a commonly used method to initialize an LSM when other options are not feasible.

[20] Method D, direct observation insertion: Soil water content in all four layers was set equal to the observed saturation at the time of initialization. This technique would be suitable if the main factor controlling spin-up time was specifying the actual observed soil moisture to obtain equilibrium conditions.

[21] Method E, direct station averaged observation insertion: Soil water content in all four layers was set equal to the spatially averaged saturation at the time of initialization. This technique would be suitable if the main factor controlling spin-up time was the average soil moisture state of a region. This technique would also be suitable to control the spin-up effect of observational error in nearby stations.

[22] Method F, direct annual average insertion: Soil water content in all four layers was set equal to annual average saturation at the respective station. This technique would be suitable if the main factor controlling spin-up time was the climatology or long-term average data.

[23] Method G, direct annual station average insertion: Soil water content in all four layers was set equal to annual average saturation at all the stations considered in the experiment. This technique would be suitable if the main factor controlling spin-up time was the climatology or long-term average data of the region. This technique would also be suitable to control the spin-up effect coming from the possible observational error in a few stations.

[24] Method H, direct monthly average insertion: Soil water content in all four layers was set equal to monthly averaged saturation at the month of initialization. This technique would be suitable if the main factor controlling spin-up time was the average monthly soil moisture state of a region.

[25] Method I, monthly shift of spin-up start time: In this experiment, we repeat experiment H but shifting the start of simulation to one month apart from the previous run. This technique would be suitable to show the sensitivity of spin-up start time and its control on the spin-up time. This technique would also be suitable to highlight the seasonal dependency of the spin-up experiment.

5. Results and Discussions

[26] The LIS-Noah model is set up to run offline spin-up tests using SGP ARM forcing inputs for 2003 at each station separately. The first set of spin-up runs is in cold-start mode, running from 1 January 2003 to 31 December 2003. Then the model is run in restart mode using the land surface states of 31 December 2003 as the initial condition in the next run. In order to investigate how land surface variables perform in the repeated run, the land surface fluxes between the two runs are compared. This comparison provides a mean to judge whether the model is spun up or not. The land surface variables should attempt to achieve an equilibrium state from the repeated runs. Additional model runs are repeated to compare the state of the land surface variables between the successive runs and observe the discrepancy in the fluxes, which is given by

$$d_Y(x,t,c) = \mathrm{Y}(x,t,c) - \mathrm{Y}(x,t,c+1) \quad (1)$$

Here d_Y (x,t,c) is the discrepancy of land surface flux Y at station x in spin-up cycle c; Y is a land surface flux being analyzed such as sensible heat flux (Qh), latent heat flux (Qle), ground heat flux (Qg) or soil moisture (Sm); t is the time. The spin cycle c denotes the discrepancy between current and the next set of model run. If the spin-up cycle c = 1 (or 2), it compares the outcomes of simulation set 1 (or 2) and set 2 (or 3). The mean discrepancy is given by

$$D_Y(x,c) = \frac{1}{n}\sum_{t=1}^{n}\sqrt{d_Y(x,t,c)^2} \quad (2)$$

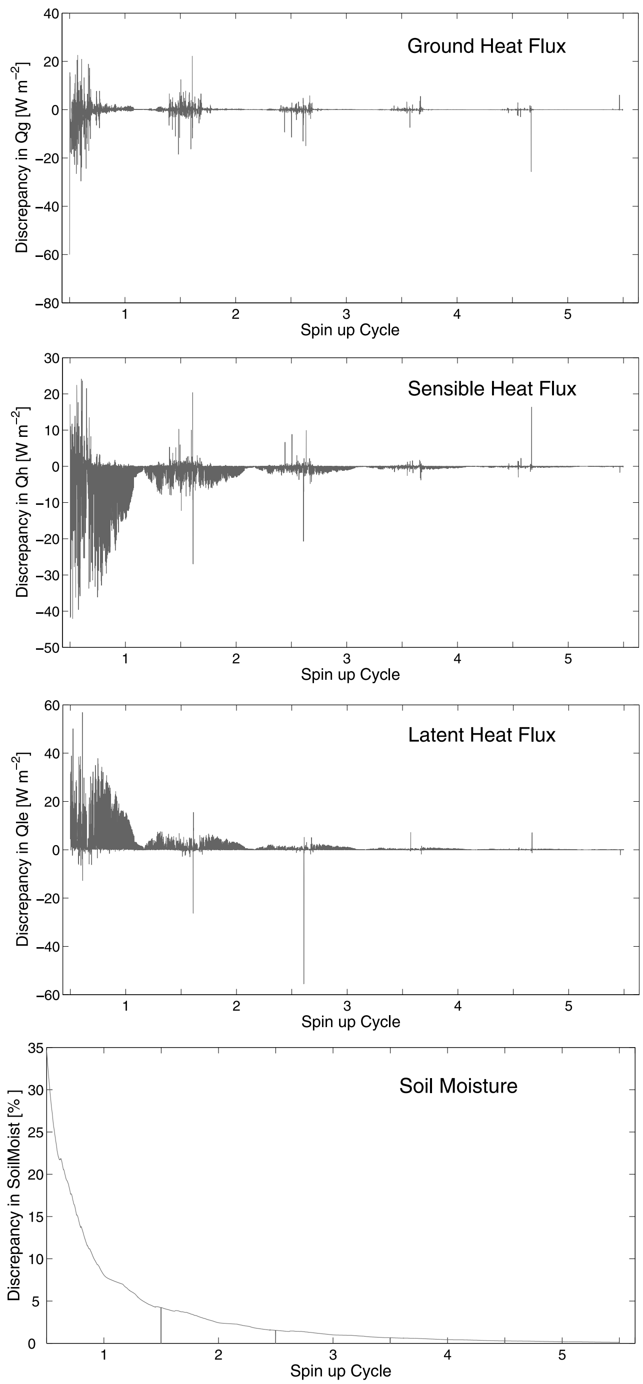

Figure 2. Incremental discrepancy in the land surface states decreases as the number of spin-up cycle increases. The discrepancy represents differences between two consecutive annual spin-up runs.

Here D_Y (x,c) is the mean discrepancy, which resembles the root mean square error formulation. It is notable that the D_Y (x,c) should reach zero in an idealized spin-up that would assure an equilibrium state of the land surface variable.

5.1. Sensitivity of Land Surface Fluxes to Spin-Up

[27] Figure 2 shows that the discrepancy d_Y (x,t,c) looks like noise in early stages of spin-up runs. As the model spin-up progresses forward, the noise is reduced gradually for any x. We have noted that the d_Y (x,t,c) tends to stabilize after some time. The trend of stabilization is also different for other land surface variables. For example, the d_Y (x,t,c) of *Qh* and *Qle* and *Qg* fluctuates and tends to stabilize from $c = 3$ to 4 (see Figure 2), where as $d_{soilmoisture}$ (x,t,c) does not fluctuate and tends to stabilize at $c = 2$. This indicates the model took several years to reach the equilibrium state and also hints at how long a model should be allowed to run before getting close to the equilibrium state. The preferable spin-up period may vary depending upon the chosen land surface flux. In this result, it required more than a 4 year run time for *Qh*, *Qg* and *Qle*, but soil moisture spin-up was nearly stabilized after a 3 year run.

[28] The D_Y (x,c) plots (Figure 3) show how the spin-up runs vary over time and space for different land surface fluxes. Each trace on the plot (Figure 3) represents the D_Y (x,c) calculated for each station from a run using the same initial condition. The decreasing trend of D_Y (x,c) indicates how well the spin-up runs are performing to bring the land surface state close to equilibrium. There is large spatial variation in the D_Y (x,c) in early cycles (i.e., $c = 1$). The higher variance corresponds to the heterogeneity of land surface conditions among stations. It suggests that using the same initial condition for different spatial locations is an inappropriate way to initialize the model. However, the variance in the D_Y (x,c) tends to disappear quickly in few additional cycles of spin-up runs. This shows the spin-up runs are adjusting land surface states rapidly with the local land surface conditions such that the spatial variation in the D_Y (x,c) does not appear in higher c runs despite having a lot of spatial heterogeneity among the model stations. This kind of rapid adjustment is expected to occur in the cyclic spin-up test where the forcing input patterns are the same for each cycle. It might be hard to observe such a rapid adjustment in a continuous spin-up test that may receive varying forcing input patterns.

[29] The variance of D_Y (x,c) for ground heat flux (*Qg*); that is, D_{Qg}, is found low at $c = 1$. This indicates that the *Qg* is less sensitive to the heterogeneity of initial land surface conditions and therefore displaying less divergence at lower c. Also, the variance of D_{Qg} (x,c) does not decrease much at higher c compared to the variance of *Qh*, *Qle* and soil moisture. The decreasing variance of D_Y (x,c) points out that land surface states are increasingly moving toward the equilibrium. This is an indicator of collective spin-up performance. However, the variance of D_Y (x,c) may actually increase, but as long as the overall trend of the discrepancy stays toward a lower value of D_Y (x,c), one could still find that the model was moving toward the equilibrium, up to a certain point at least. A higher variance of $D_Y(x,c)$, in such condition, indicates that the model failed to perform collectively well in moving toward the equilibrium. Therefore, the high divergence of D_{Qg} (x,c) at later spin-up cycles (higher c) indicates that the ground heat flux does not adjust to the land surface condition heterogeneity as quickly as other land surface fluxes such as *Qh*, *Qle* and soil moisture. This is understood as the long memory effect of ground heat flux, which needs a longer time to adjust in more heterogeneous land surface conditions.

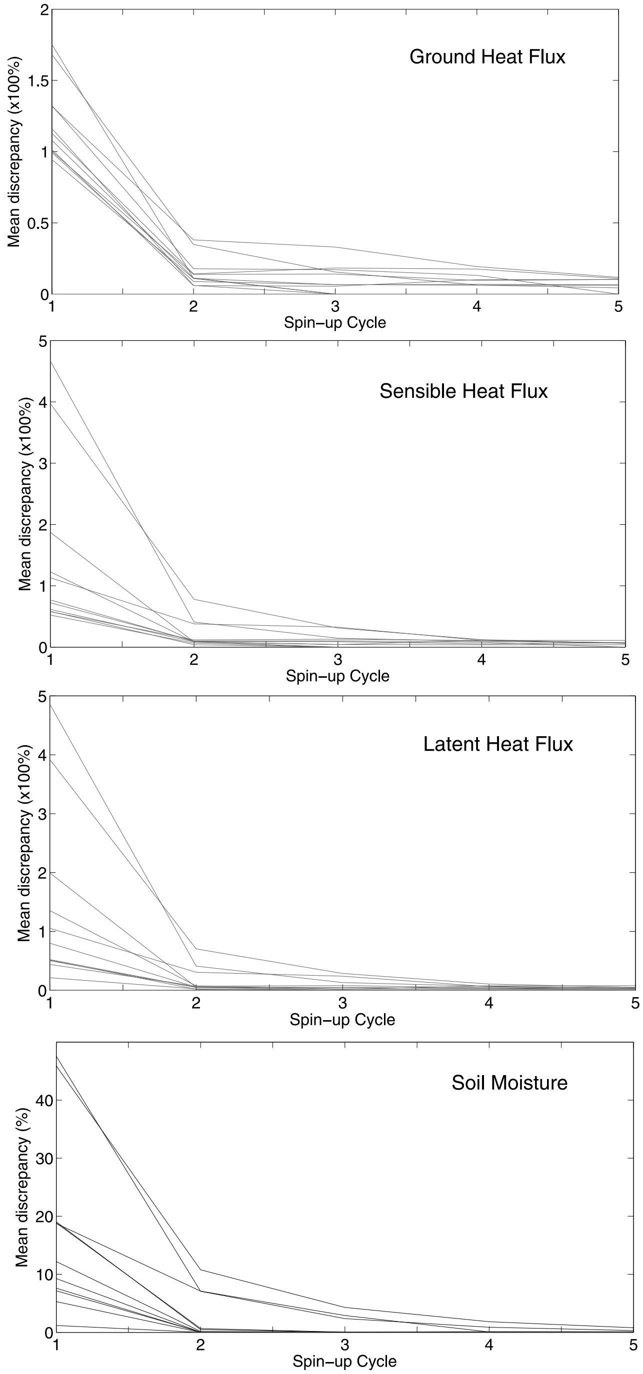

Figure 3. Ensemble plot of decreasing discrepancy in land surface states with increasing repetition cycle obtained from method C.

5.2. Sensitivity of Different Initializing Methods

[30] Depending upon the methods chosen to initialize the spin-up runs, the model test runs produce different outcomes. We report analysis showing how the soil moisture states are affected from the methods chosen to initialize the test runs. The analysis is based on the discrepancy mean and variance $D_Y\,(x,c)$ using the total column soil moisture field outputs from the test runs.

[31] The results in Figure 4 show that a dry start (method A) appears to be the worst, as the discrepancy is higher than 5% after completion of four cycles. The wet start (method B) is found to be an improper choice but slightly better than the dry initialization. The average moisture start (method C) is found to be much better than the dry and wet start. *Rodell et al.* [2005] report similar results. Hence, the hypothesis that an average value is preferable to a wet or dry value seems to be correct in a range of LSM experiment scales.

[32] Other initialization methods such as the direct observation insertion (method D), the direct station averaged observation insertion (method E), the direct annual station average insertion (method F) have performed poorly than the average moisture start (method C) in early cycles of the test runs conducted. These three direct insertion-based methods are found better than both dry and wet start methods but they are not found as good as the average moisture start method. The direction annual station average insertion (method G) and the direct monthly average insertion (method H) are found to perform better than the average moisture start (method C) in the comparison after the first cycle. Method G is found slightly worse after the second cycle but overall performance of the methods C, G and H in the spin-up test runs is better than others.

[33] The results indicate that the direct observation insertion (method D) could be risky as it may cause poorer spin-up performance if there is flaw in the data observation system. This method brings the observational uncertainty into the model initialization. There is also a chance that the observed state does not resemble the equilibrium model state because there are numerous other factors and parametric controls in the model that would significantly differ from the real interaction of land surface fluxes at the particular time of the start of spin-up. The possible inappropriate value inserted from the direct observation may cause the model take longer to spin-up. One way of reducing the observation uncertainty is to average the data over multiple stations. Model initialization using the averaged observation (method E) is found to perform consistently better than method D. This method inserts the mean state field for the time of the start of spin-up and does not consider the soil moisture states over a longer period. The time-averaged state, which is obtained from the observed soil moisture at a

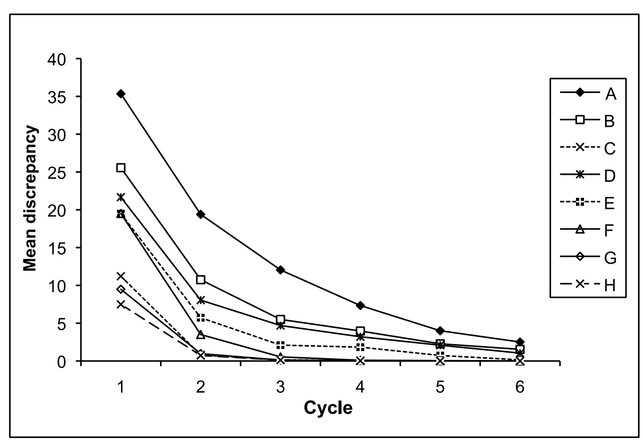

Figure 4. Comparison of mean discrepancies in soil moisture field from different initialization methods.

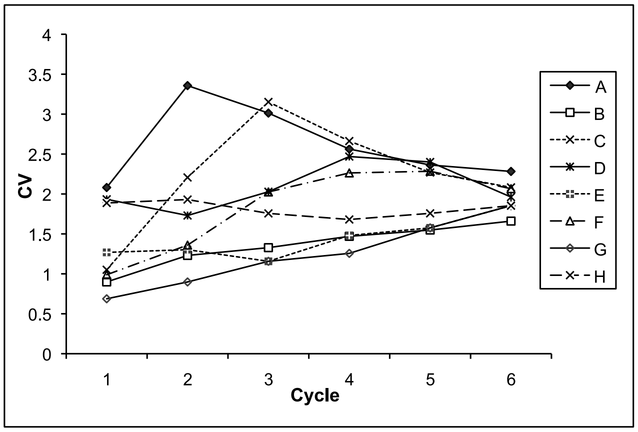

Figure 5. Comparison of CV of discrepancies in soil moisture field from different initialization methods.

station for a longer time period (method F), is found to be better for initializing the model than method E. Method F uses the time averaged mean state and therefore sets the model initialization close to the climatologically representative state for the particular station.

[34] By assigning the different initialization conditions for each station taken for the experiment (method F) assures more heterogeneous initialization than the space-time mean state obtained from all the stations (method G). However, method G is found to perform better than method F. The discrepancy in the results drops from ~20% (method F) to 10% (method G) in the first cycle and gets close to equilibrium rapidly in the successive cycles. Method G is found to be better than the average start, which was performing better until this test. The better performance of method G clearly indicates that model initialization may be improved by using a spatially and temporally averaged observation record.

[35] There is large soil moisture space-time variation over a year and across the spatial domain. The better results obtained from method G, which ignores the significance of spatial and temporal variability, could be revealed as a constraint in the model spin-up test that we might understand as if the space-time variation has an insignificant role in model initialization. By releasing the constraint and using a shorter averaging window separately over the stations (method H), the model initialization performance has further improved. The improvement in model spin-up while using the monthly average observation (method H) indicates that the use of long-term averages may not be the most efficient for model spin-up. The improved result of method G over method F could be due to a compensating effect of spatial averaging which has partially corrected the inappropriateness of the long-term average state used in the method F.

5.3. Spatial Heterogeneity of Spin-Up Results

[36] The discrepancies observed in the land surface fluxes during the spin-up runs vary from one station to another. The mean results obtained from several point stations (Figure 4) display the trend showing how the LSM advances to equilibrium during the spin-up runs. The mean result, however, is derived from several nonuniform results, which are the outcomes of spatially distributed model runs at several stations. These spatially distributed model runs exhibit an interesting phenomenon relating to how the different initialization methods work out the spatial heterogeneity in the spin-up outcomes. It is found that the model spin-up tends to build a certain level of heterogeneity as the spin-up runs get closer to the equilibrium state and the initialization methods are sensitive to the process of acquiring heterogeneous equilibrium state.

[37] Figure 5 compares the coefficient of variation (CV) of soil moisture discrepancies from different initialization methods. The coefficient of variation is a dimensionless simple statistical index based on standard deviation and mean statistics which measures dispersion of probability distribution. A lower CV indicates uniform discrepancies among stations and vice versa. The CV may increase when the variability of discrepancy increases or the mean discrepancy decreases. Previous analysis has shown the mean discrepancy decreases continuously in cyclic spin-up test. If the decrease were uniform among the stations, the CV would remain the same. Finding a change in CV is the consequence of nonuniform decrease of discrepancy resulting from the spatial heterogeneity in the land surface states.

[38] An example of the dry start (method A) can be taken to illustrate the interplay of the CV discrepancies and the spatial heterogeneity, which has yielded CV = 2.1 at cycle 1. This CV represents a standard deviation that is nearly twice as large as the mean value of the discrepancies between the first and second spin-up year. The CV is increased nearly to 3.5 at cycle 2 while the LSM has attempted to reach closer to its equilibrium state in the next cycle. There is a decrease in the mean discrepancy at the next spin-up cycle, and at the same time, the LSM has attempted to fit the local land surface condition, which is not uniform owing to the station spatial variability. As a result, the standard deviation and mean values of discrepancies are nonuniform. The resulting higher values of CV implies that the standard deviation is not decreasing proportionately to the mean discrepancies, but trying to fit with the proper pattern of spatial heterogeneity as the spin-up cycles progress. The process of attempting to fit with the spatial heterogeneity as well as yielding the equilibrium land surface state continues and after five cycles of spin-up runs, the dry start (method A) spin-up test yields CV close to 2.2.

[39] In the subsequent use of other methods to initialize the model, the CV fluctuates differently. Figure 5 shows that method A has comparatively higher CV since the early spin-up runs. This is because of the dry start condition, which causes high soil moisture variation among the stations. However, method B has a wet start (i.e., saturated soil condition), which causes a lower variation of soil moisture in the initial runs. As the spin-up cycle progresses further, the LSM gradually adapts to the natural pattern of spatial heterogeneity and hence the CV increases. The average moisture start (method C) has a low CV almost equal to the wet start (method B) in the beginning. This shows that the LSM did a little adjustment in the moisture level at the initial spin-up runs. The CV in method C increased rapidly until cycle three and then it started to decrease adjusting the CV close to what other methods had yielded.

[40] The CV of method D is high all the time showing the role played by the direct insertion of the observation in the

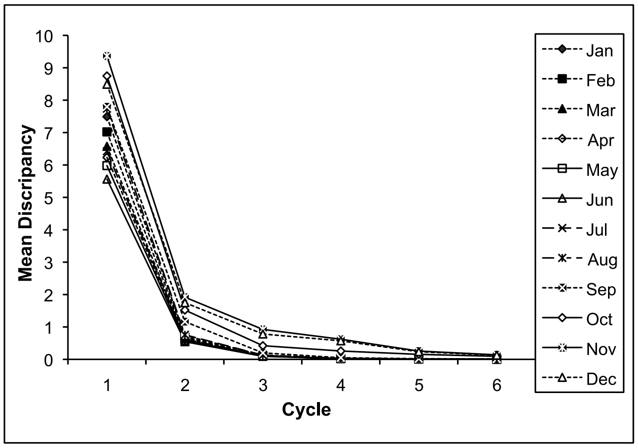

Figure 6. Comparison of initialization begun at different months.

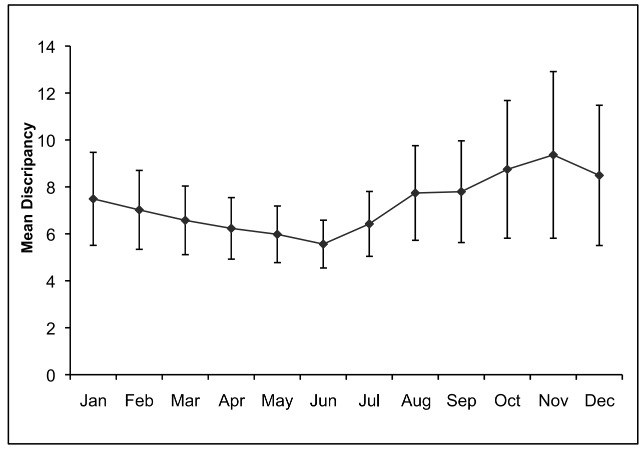

Figure 7. Comparison of mean discrepancy at initial spin-up cycle in different months.

spin-up runs. There is slight decrease, then an increase and then again a decrease in the CV as the spin-up cycle progresses, which shows how the spin-up runs have struggled to align with the natural spatial heterogeneity. Method E has lower CV in the initial cycles as the result of averaged value used for direct insertion. The CV is found almost constant up to cycle three and then has gradually increased as it did with the method B. Method F also has a lower CV similar to method C. The CV increased slowly and steadily until cycle three, and then remained nearly constant. Method G has the lowest CV in the initial spin-up cycle, which increased consistently until the last cycle. Unlike other methods, method H has an almost constant CV throughout the test showing that it has been close to the needed level of spatial heterogeneity since the initial cycle.

[41] From the heterogeneity viewpoint, it is clear that all methods have struggled to attain the heterogeneity equivalent to the CV of about 2. Most methods took five to six cycles to achieve the CV of ~2 except methods D, F, and H. These three methods took only three cycles to yield the CV of ~2. These methods have inserted spatially variable states for the model initialization, which plays an important role in the model spin-up. The methods A, D and H achieve the CV of ~2 in the first cycle but methods D and H only achieve CV of ~2 in the second cycle. The superiority of method H in maintaining proper degree of heterogeneity is evident from the results.

5.4. Starting the Spin-Up Run in a Different Month

[42] The earlier analysis has shown that the model spin-up performance is better when monthly average data are used for model initialization (method H). Those results were obtained from the January average, which is the starting month for all the simulations. Although January is the starting month of the year, the nature or other physical processes may not have any particular starting preference. It is interesting to investigate how the spin-up response would differ if the model initializations start at a different time, say, to start the next month. This set of experiments has investigated the dependence of spin-up results starting the model a month apart. This experiment not only tests the sensitivity of the monthly data, but it also helps to understand the land surface dynamics on a monthly or seasonal basis.

[43] The discrepancy plots combined show that the spin-up results have a great degree of variation based on the choice of spin-up start (see Figure 6). The trend of decreasing discrepancy is unaffected but the mean values of discrepancies are not exactly the same. These variations in the mean discrepancies are solely due to the monthly soil moisture variation pattern.

[44] Figure 7 shows the monthly fluctuation of the mean discrepancies and the standard deviation in the first spin-up cycle. The mean discrepancies decrease from January to June and then increase afterward. Summer months are found to yield lower discrepancies than winter months. The standard deviation varies likewise.

[45] Figure 8 shows the variation of CV in the different spin-up cycles begun at different months. The CV plot shows that the results are not consistent throughout the spin-up runs. The January spin-up runs do not show the best result in both mean discrepancies and CV plots. June has a lower discrepancy and lower CV, showing the higher degree of uniformity in the discrepancies at the initial spin-up runs.

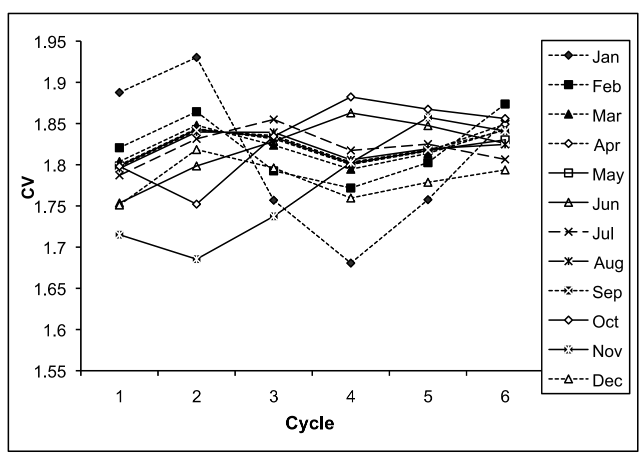

Figure 8. Comparison of CV of discrepancy at various spin-up cycles started in different months.

The July and September months have quite stable CVs, however September has the higher value of mean discrepancy. Viewing the uniform nature of CV plot and lesser discrepancies, the month of July seems to perform better than other months to begin the spin-up runs in this kind of test. The monthly input used to start the cycles is from the year studied but not from the monthly climatological averages. It might make a difference if one had an unusual monthly season and used that month as the spin-up initialization data. Since the studied year has no report of unusual monthly records, it is assumed that the reported results and analysis should represent the land surface conditions appropriately.

6. Summary and Conclusion

[46] Accurate initialization of land surface model energy and water stores is critical for making skillful land surface predictions. The present study investigated aspects of spin-up runs for LIS-Noah LSM run at point stations using a 1 km grid setting. The model outputs are compared to evaluate the discrepancies over the annual cycle of repeated runs to find out how the model attains an equilibrium state in the land surface fluxes.

[47] The spin-up results are closely linked with non-meteorological controls, metrological controls and the chosen initialization methods. Primary nonmetrological controls include the type and structure of model and parameters such as soil depth, soil type, hydrological conductivity, root depth, degree of wetness, vegetation parameters and additional model specific parameters. Primary metrological controls include precipitation, temperature, wind, sunshine hours and net radiation. All of these variables regulate the dynamics of the water and energy balance, moisture storage on layers of soil column and transfer of radiative energy and are likely to control the spin-up results significantly.

[48] The experiments conducted here are in a small region where the spatial regimes of nonmetrological and metrological controls do not differ greatly. Yet, the land variables at those experimental stations and their spin-up response are highly heterogeneous. This manifests the land surface process complexities even if a small region is studied with a high-resolution model and high-resolution observations. The model results are found to accommodate the variability when sufficient spin-up is allowed. The study might have benefited from the repeated use of annual forcing data over multiple years. If the forcing data is affected by a systematic multiannual trend (increasing, decreasing, or fluctuating) it may affect the model and may need a longer spin-up run.

[49] It is noted that the appropriate spin-up period can vary depending upon the chosen land surface states. The spin-up response is also found dependent on the land surface spatial heterogeneity. Early termination of spin-up runs could fail to adapt with the large land surface spatial variations. The model adjusts land surface states rapidly with the local land surface conditions but it needs multiple spin-up cycles. The fluxes (such as sensible heat or latent heat) are adjusted faster than the ground heat flux, which retains memory of the land surface state for considerably longer periods.

[50] Nine different methods of initializations are evaluated to investigate how the soil moisture states are affected from the chosen initialization methods. The results demonstrate that certain initialization methods are superior. All the spin-up methods have used the same approach of looping the simulation on a single annual forcing cycle. This technique has a drawback of overlooking annual anomalies in the meteorological forcing that may accumulate artificial anomalies in longer runs. However, this is one of the most commonly used approaches and is most suitable for data poor spin-up runs. Moreover, using the same approach is preferable for sound comparison of the results obtained from the different initialization methods.

[51] Among the methods chosen to initialize the LSM, the average moisture start method is found to be surprisingly better than other blind guess methods such as dry and wet start. Therefore, it could be a good choice to use the average initial states and loop through available years of forcing until a desired equilibrium level is attained. However, this may be just a coincidence in that particular region where the expected best initialization is close to the guessed average value. It is found that inserting the in situ data directly into the model initialization may not yield better spin-up results. In some cases, the direct insertion methods are found counter-productive compared to blind guess. The spin-up performance can be improved by using the mean state fields. Use of temporal mean state field is better than the spatial mean state field for the precise spin-up start time. However, a longer temporal mean may not be better for initializing the model. Among the tested methods, the monthly averaged mean state performed better than others studied here.

[52] The land surface model has to respond appropriately with the land surface process spatial heterogeneity. It is found that the spin-up process builds up a certain level of heterogeneity along the process of getting into the equilibrium state. The results suggest that using the same initial condition for different spatial locations is an inappropriate way to initialize the model. One of the ways to improve the spin-up efficiency could be to recognize the heterogeneity from the model initialization start such that the spin-up runs would achieve the desired level of heterogeneity and equilibrium faster. Use of spatially uniform states to initialize the model is found inferior despite yielding a lower discrepancy in the mean results.

[53] As stated above, the monthly averaged mean state performs better for the model initialization in both measures: faster to reach the desired level of equilibrium, and faster to adapt with the desired level of heterogeneity. However, there is a seasonal effect that could guide the preferred choice of the month to start the model spin-up. The land surface spin-up runs started in the summer months are found to have a lesser degree of uncertainty than the winter months both in the measure of mean and variance of discrepancies. Among the tested results, June and July are found to be the best months of the year to begin spin-up runs. The results and findings presented here provide important understanding about the factors controlling the spin-up results and are likely to benefit in the spin-up runs of similar land surface model for the regions having similar hydrometeorological regime. The effectiveness of the described initialization methods may be quite different in other regions of different hydrometeorological regime and in other land surface models.

References

Betts, A. K., J. H. Ball, A. C. M. Beljaars, M. J. Miller, and P. A. Viterbo (1996), The land surface-atmosphere interaction: A review based on observational and global modeling perspectives, *J. Geophys. Res.*, *101*, 7209–7225, doi:10.1029/95JD02135.

Chen, F., K. Mitchell, J. Schaake, Y. Xue, H.-L. Pan, V. Koren, Q. Y. Duan, M. Ek, and A. Betts (1996), Modeling of land surface evaporation by four schemes and comparison with FIFE observations, *J. Geophys. Res.*, *101*, 7251–7268, doi:10.1029/95JD02165.

Cosgrove, B. A., et al. (2003), Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project, *J. Geophys. Res.*, *108*(D22), 8842, doi:10.1029/2002JD003118.

de Goncalves, L. G. G., W. J. Shuttleworth, E. J. Burke, P. Houser, D. L. Toll, M. Rodell, and K. Arsenault (2006), Toward a South America Land Data Assimilation System: Aspects of land surface model spin-up using the Simplified Simple Biosphere, *J. Geophys. Res.*, *111*, D17110, doi:10.1029/2005JD006297.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley (2003), Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, *J. Geophys. Res.*, *108*(D22), 8851, doi:10.1029/2002JD003296.

Henderson-Sellers, A., Z. L. Yang, and R. E. Dickinson (1993), The project for intercomparison of land-surface parameterization schemes, *Bull. Am. Meteorol. Soc.*, *74*, 1335–1349, doi:10.1175/1520-0477(1993)074<1335:TPFIOL>2.0.CO;2.

Henderson-Sellers, A., A. J. Pitman, P. K. Love, P. Irannejad, and T. H. Chen (1995), The project for intercomparison of land surface parameterization schemes (PILPS): Phase 2 and 3, *Bull. Am. Meteorol. Soc.*, *76*, 489–503, doi:10.1175/1520-0477(1995)076<0489:TPFIOL>2.0.CO;2.

Hill, C., C. Deluca, X. Balaji, M. Suarez, and A. Da Silva (2004), The architecture of the Earth System Modeling Framework, *Comput. Sci. Eng.*, *6*, 18–28, doi:10.1109/MCISE.2004.1255817.

Kato, H., M. Rodell, F. Beyrich, H. Cleugh, E. Gorsel, H. Liu, and T. P. Meyers (2007), Sensitivity of land surface simulations to model physics, land characteristics, and forcings, at four CEOP sites, *J. Meteorol. Soc. Jpn.*, *85A*, 187–204, doi:10.2151/jmsj.85A.187.

Koren, V., J. Schaake, K. Mitchell, Q.-Y. Duan, F. Chen, and J. M. Baker (1999), A parameterization of snowpack and frozen ground intended for NCEP weather and climate models, *J. Geophys. Res.*, *104*, 19,569–19,585, doi:10.1029/1999JD900232.

Koster, R. D., M. J. Suarez, P. Liu, U. Jambor, A. Berg, M. Kistler, R. Reichle, M. Rodell, and J. Famiglietti (2004), Realistic initialization of land surface states: Impacts on subseasonal forecast skill, *J. Hydrometeorol.*, *5*, 1049–1063.

Kumar, S. V., et al. (2006), Land information system: An interoperable framework for high resolution land surface modeling, *Environ. Modell. Software*, *21*, 1402–1415, doi:10.1016/j.envsoft.2005.07.004.

Kumar, S. V., C. D. Peters-Lidard, J. L. Eastman, and W.-K. Tao (2008), An integrated high-resolution hydrometeorological modeling testbed using LIS and WRF, *Environ. Modell. Software*, *23*, 169–181, doi:10.1016/j.envsoft.2007.05.012.

Robock, A., K. Y. Vinnikov, C. A. Schlosser, N. A. Speranskaya, and Y. Xue (1995), Use of midlatitude soil moisture and meteorological observations to validate soil moisture simulations with biosphere and bucket models, *J. Clim.*, *8*, 15–35, doi:10.1175/1520-0442(1995)008<0015:UOMSMA>2.0.CO;2.

Rodell, M., P. R. Houser, A. A. Berg, and J. S. Famiglietti (2005), Evaluation of 10 methods for initializing a land surface model, *J. Hydrometeorol.*, *6*, 146–155.

Xue, Y., P. J. Sellers, J. L. Kinter II, and J. Shukla (1991), A simplified biosphere model for global climate studies, *J. Clim.*, *4*, 345–364, doi:10.1175/1520-0442(1991)004<0345:ASBMFG>2.0.CO;2.

Xue, Y., F. J. Zeng, and C. A. Schlosser (1996), SSiB and its sensitivity to soil properties: A case study using HAPEX-Mobilhy data, *Global Planet. Change*, *13*, 183–194, doi:10.1016/0921-8181(95)00045-3.

Yang, Z.-L., R. E. Dickinson, A. Henderson-Sellers, and A. J. Pitman (1995), Preliminary study of spin-up processes in land surface models with the first stage data of Project for Intercomparison of Land Surface Parameterization Schemes Phase 1(a), *J. Geophys. Res.*, *100*, 16,553–16,578, doi:10.1029/95JD01076.

P. Houser, Department of Geography and Geoinformation Science, George Mason University, Fairfax, VA 22030, USA.

R. Shrestha, Center for Research on Environment and Water, Institute of Global Environment and Society, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705, USA. (roshan@iges.org)