GEOPHYSICAL RESEARCH LETTERS, VOL. 34, L09402, doi:10.1029/2007GL029804, 2007

Improving land-surface model hydrology: Is an explicit aquifer model better than a deeper soil profile?

Lindsey E. Gulden,[1] Enrique Rosero,[1] Zong-Liang Yang,[1] Matthew Rodell,[2] Charles S. Jackson,[3] Guo-Yue Niu,[1] Pat J.-F. Yeh,[4] and James Famiglietti[4]

Received 24 February 2007; revised 3 April 2007; accepted 10 April 2007; published 5 May 2007.

[1] We use Monte Carlo analysis to show that explicit representation of an aquifer within a land-surface model (LSM) decreases the dependence of model performance on accurate selection of subsurface hydrologic parameters. Within the National Center for Atmospheric Research Community Land Model (CLM) we evaluate three parameterizations of vertical water flow: (1) a shallow soil profile that is characteristic of standard LSMs; (2) an extended soil profile that allows for greater variation in terrestrial water storage; and (3) a lumped, unconfined aquifer model coupled to the shallow soil profile. North American Land Data Assimilation System meteorological forcing data (1997–2005) drive the models as a single column representing Illinois, USA. The three versions of CLM are each run 22,500 times using a random sample of the parameter space for soil texture and key hydrologic parameters. Other parameters remain constant. Observation-based monthly changes in state-averaged terrestrial water storage (dTWS) are used to evaluate the model simulations. After single-criteria parameter exploration, the schemes are equivalently adept at simulating dTWS. However, explicit representation of groundwater considerably decreases the sensitivity of modeled dTWS to errant parameter choices. We show that approximate knowledge of parameter values is not sufficient to guarantee realistic model performance: because interaction among parameters is significant, they must be prescribed as a congruent set. **Citation:** Gulden, L. E., E. Rosero, Z.-L. Yang, M. Rodell, C. S. Jackson, G.-Y. Niu, P. J.-F. Yeh, and J. Famiglietti (2007), Improving land-surface model hydrology: Is an explicit aquifer model better than a deeper soil profile?, *Geophys. Res. Lett.*, *34*, L09402, doi:10.1029/2007GL029804.

1. Introduction

[2] With the growing recognition of groundwater–atmosphere interaction as a potentially significant influence on spatial and temporal climate variability, researchers in the field of terrestrial hydrometeorology have focused increasing attention on improving the process representations of subsurface hydrology within land-surface models (LSMs). Existing process representations fall within three broad classes: (1) multi-layered, relatively shallow soil columns in which groundwater storage is implicitly represented because the model conserves mass [e.g., *Oleson et al.*, 2004]; (2) many-layered, deep soil columns whose lower boundaries are beneath the climatological depth to the water table [*Koster et al.*, 2000; *Maxwell and Miller*, 2005]; and (3) multi-layered soil columns coupled to lumped, unconfined aquifer models [*York et al.*, 2002; *Liang et al.*, 2003; *Yeh and Eltahir*, 2005; *Fan et al.*, 2007; *Niu et al.*, 2007].

[3] Which of these methods best represents subsurface hydrology at a monthly time scale? We address this question for three different levels of parameter uncertainty: (1) when an optimal set of subsurface hydrologic parameters (e.g., percent sand, porosity, and specific yield) can be inferred from observations (the "ideal" case); (2) when no information about effective parameters can be obtained (the "worst" case); and (3) when only ranges for parameter values are known (the "real life" case).

[4] To ensure a fair comparison between methods, we isolate process representation as the primary source of uncertainty in model predictions. To limit input-data uncertainty, we employ the same meteorological forcing data and land-surface data for all runs. We use a Monte Carlo approach to explore the impact of parameter uncertainty. Unlike calibration studies, the underlying goal of this work is not to identify the optimal parameter set; instead our primary goal is to evaluate and compare the added value of process representations.

[5] Three questions frame our analysis: (1) When given a surrogate optimal parameter set, which of the ways to represent subsurface hydrology results in the most realistic simulation of monthly change in terrestrial water storage? (2) When no reliable information regarding effective subsurface hydrologic parameters exists, which process representation most consistently gives the best performance? (3) Does knowledge of approximate values for hydrologic parameters guarantee reasonably accurate simulation of monthly change in terrestrial water storage? Our results will inform LSM model development; more important, they characterize the level of confidence that can be placed in LSM-generated hydrologic predictions, especially when observations are scarce.

2. Methods

[6] We use the National Center for Atmospheric Research's Community Land Model (CLM) [*Bonan et al.*, 2002; *Oleson et al.*, 2004; *Niu et al.*, 2005] as the host model in which to test three methods for representing vertical water flow within the LSM soil column. The

[1]Department of Geological Sciences, John A. and Katherine G. Jackson School of Geosciences, University of Texas at Austin, Austin, Texas, USA.
[2]Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[3]Institute for Geophysics, John A. and Katherine G. Jackson School of Geosciences, University of Texas at Austin, Austin, Texas, USA.
[4]Department of Earth System Science, University of California, Irvine, California, USA.

0094-8276/07/2007GL029804$05.00

Table 1. Ranges and Distributions of Randomly Sampled Subsurface Hydrologic Parameters

Parameter	Range	Distribution
Sand[a]	5 to 90%	uniform
Clay[a]	5 to (100 − [% sand])%	semi-uniform
Porosity	0.01–0.50 $m^3 m^{-3}$	uniform
e-folding depth of saturated hydraulic conductivity	0.1–100 m	uniform
Maximum rate of baseflow (rsbmax)	1×10^{-11}–1×10^{-3} m s^{-1}	log uniform
Specific yield[b]	0.01–0.25	uniform

[a]CLM calculates hydraulic conductivity and matric potential as a function of percent sand and percent clay according to the methods of *Clapp and Hornberger* [1978] and *Cosby et al.* [1984]. Percent silt is 100−(% sand + % clay).

[b]Specific yield is used only by the AQUIFER runs, not by the SHALLOW or DEEP runs.

versions of CLM calculate surface and subsurface runoff (i.e., baseflow) as a function of topographic characteristics [*Niu et al.*, 2005] and are identical except for the method that they use to represent vertical water transfer in the soil column.

[7] The first version of CLM (hereafter "SSOIL") uses the standard 10-layer, relatively shallow 3.43-m soil profile with topography-based runoff parameterizations [*Niu et al.*, 2005]. Because it conserves mass, the model implicitly represents groundwater dynamics; however, the true depth to the water table often exceeds the depth of the model's lower boundary. The second model (hereafter "DEEP") is identical to SSOIL except that it uses a 30-layer, 11.2-m soil profile, thereby extending the depth of the model soil profile to encompass a wider range of groundwater fluctuations. The third version (hereafter "AQUIFER") couples a lumped unconfined aquifer model to the standard 10-layer soil profile [*Niu et al.*, 2007]; it allows two-directional vertical water transfer between the unsaturated zone and the aquifer down a hydraulic gradient.

[8] We run each version of the model as a single column representing the state of Illinois, USA. Illinois covers ∼146,000 km^2. Crops and grass dominate the landscape. The climate is temperate and continental, and the topographic relief is relatively low. (See *Changnon et al.* [1988] and *Yeh et al.* [1998] for detailed descriptions of regional climate and hydrogeology.)

[9] Meteorological forcing and land-surface input data are the area-weighted arithmetic averages of high-resolution datasets over the state of Illinois. The forcing is provided by the North American Land Data Assimilation System [*Cosgrove et al.*, 2003]. A CLM-compatible land-cover dataset derived from Advanced Very High Resolution Radiometer and Moderate Resolution Imaging Spectroradiometer data [*Lawrence and Chase*, 2007] provides vegetation type distributions, biomass densities, and soil colors.

[10] A Monte Carlo approach allows us to extensively explore the range of model responses across parameter space. We run SSOIL, DEEP, and AQUIFER 22,500 times each. A unique set of subsurface hydrologic parameters is used for each run. We randomly sample uniform or semi-uniform distributions that span physically reasonable ranges of values for soil texture parameters and other hydrologic parameters (Table 1). Each Monte Carlo run is initialized with a spun-up dataset created by running the model three times through the period 1997–2005 using default parameters. To allow for additional spin-up, the first year of each run is omitted from the analysis.

[11] We assess the accuracy of model output using the statewide-average change in total column terrestrial water storage (dTWS), which we constructed from soil moisture and groundwater observations obtained by the Illinois State Water Survey (ISWS) [*Hollinger and Isard*, 1994; *Robock et al.*, 2000] following the methods of *Rodell and Famiglietti* [2001]. dTWS is a suitable constraint because it integrates the hydrologic behavior of the landscape; it is directly observable everywhere on Earth using Gravity Recovery and Climate Experiment (GRACE) measurements [*Chen et al.*, 2006]; and it properly represents the land storage term of the coupled atmospheric-terrestrial water budget.

[12] Looking only at data from 1998–2005, we score parameter sets with the following metric:

$$F = \mathrm{RMSE} \times (1 - r) \tag{1}$$

where RMSE is the root mean square error between modeled and observed dTWS:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(o_i - m_i)^2} \qquad \begin{aligned} n &= \text{length of time series} \\ o_i &= \text{observed dTWS at time } i \\ m_i &= \text{modeled dTWS at time } i \end{aligned} \tag{1a}$$

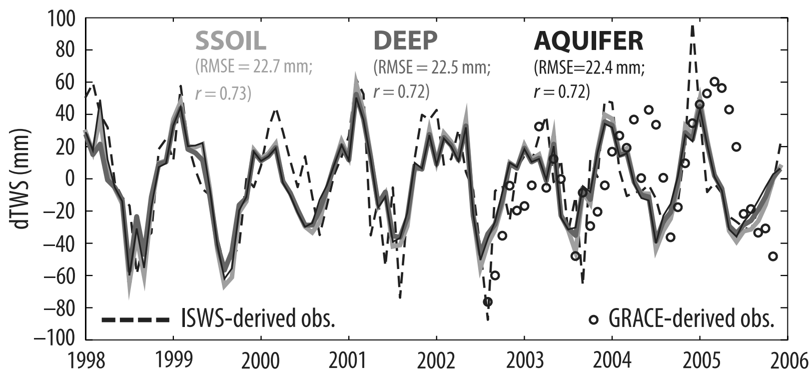

Figure 1. Observed monthly dTWS compared with that simulated by each model with its optimal parameter set. GRACE-derived data [*Chen et al.*, 2006] are shown only for reference; they were not used to score model output.

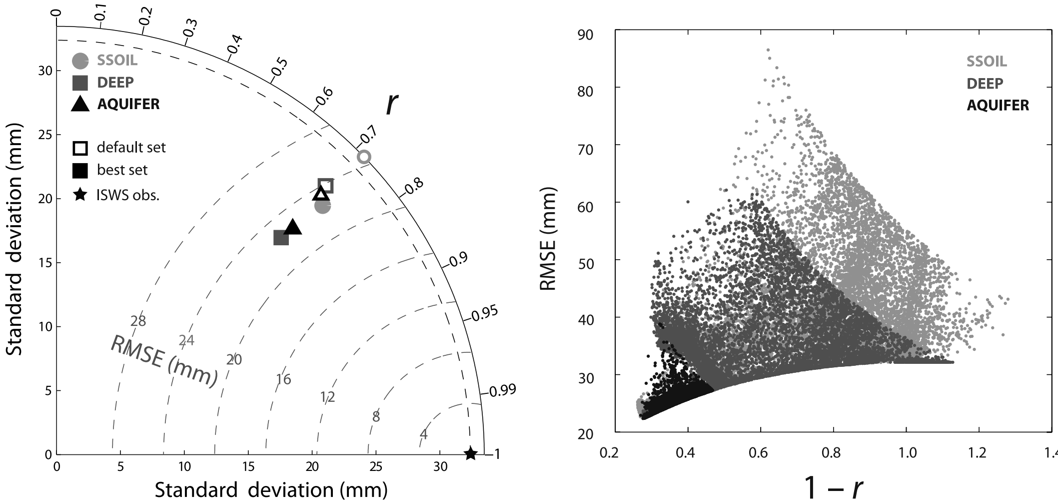

Figure 2. Model performance. A Taylor diagram including (left) scores from runs using default and best parameter sets and (right) scores for the top-scoring 50% of runs. Figure 2 (right) contains an equal number (11,250) of black, dark gray, and light gray points.

and r is the correlation coefficient, defined as:

$$r = \frac{\sum_{i=1}^{n} (o_i - \bar{o})(m_i - \overline{m})}{\left(\sum_{i=1}^{n} (o_i - \bar{o})^2 \sum_{i=1}^{n} (m_i - \bar{m})^2\right)^{1/2}} \quad \begin{array}{l} n, o_i, m_i \text{ defined in(1a)} \\ \bar{o} = \text{mean observed dTWS} \\ \overline{m} = \text{mean modeled dWTS} \end{array} \tag{1b}$$

We use F because it allows us to select parameter sets for which both the timing and amplitude of the modeled seasonal cycle match observations. We define the best parameter set as that which minimizes F; we use it as a surrogate optimum. We perform the exhaustive parameter exploration, which mimics a single-criteria manual calibration.

3. Results and Discussion

3.1. When Given a Pseudo-Optimum Parameter Set, Which Process Representation Is Better?

[13] When given their best parameter sets, SSOIL, DEEP, and AQUIFER are equivalently adept at simulating monthly dTWS in Illinois. For all three models, 22.4 mm $\leq$ RMSE $\leq$ 22.7 mm and $r \geq 0.72$ (Figure 1). In the ideal case where observations exist and calibration identifies the optimal parameter set, the most computationally efficient model (either SSOIL or AQUIFER) should be used. Single-criteria analysis does not provide sufficient information with which to distinguish the overall performance of the models. Future work will use automatic multi-criteria parameter estimation to further explore the variation in model skill.

3.2. When Little Is Known About Parameter Values, Which Process Representation Is Best?

[14] In the absence of specific information, modelers often use default parameter sets recommended by model developers. The Taylor diagram [*Taylor*, 2000] in Figure 2 shows the performance of both the default and best sets for each of the three models. For all three models, dTWS simulated using the best set has lower variance than observed dTWS, and the improvement over the default set is marginal. The good performance of the default set is not surprising: as one of the few extensive hydrologic datasets in the world, ISWS observations regularly inform LSM development and default parameter estimations.

[15] For most locations, little reliable information about subsurface hydrologic parameters exists, and we have no way to know whether the default parameter set adequately represents effective parameters. Figure 2 (right) shows the

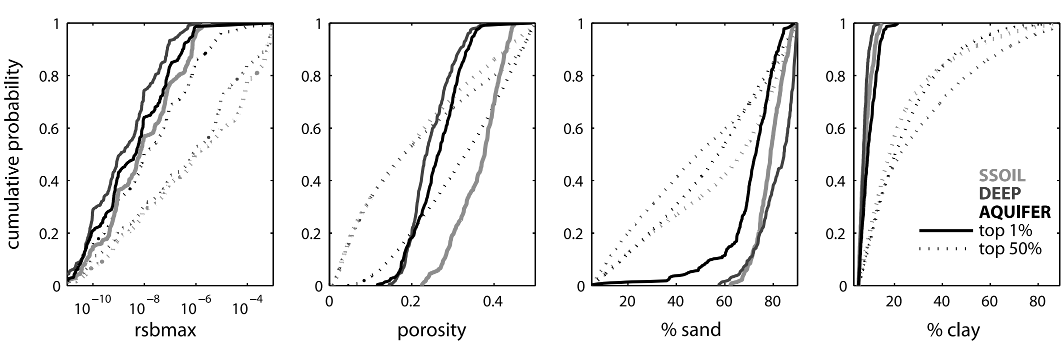

Figure 3. Empirical cumulative distribution functions for parameters used in top-scoring 1% and 50% of runs. Distributions for specific yield and e-folding depth of hydraulic conductivity are not shown.

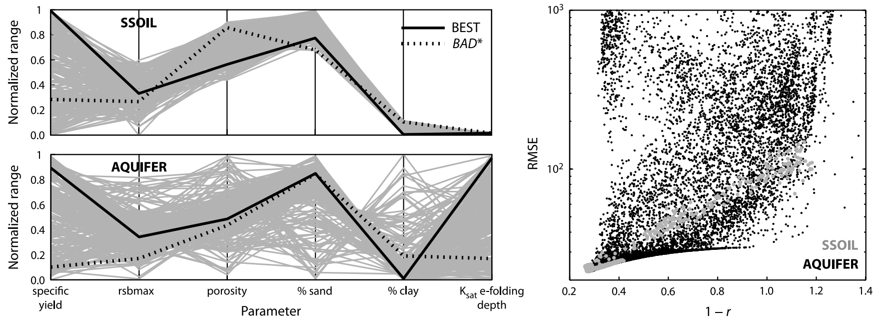

Figure 4. (left) Parameter sets corresponding to the top-scoring 1% of runs in normalized parameter space. (right) Scores of all the sets that fall within the envelope created by the top 1%. *BAD* is not a member of the top-scoring 1%; it is the worst-scoring parameter set in the envelope.

variations in the consistency of model performance between SSOIL, DEEP, and AQUIFER. Each point represents a single model run in which a unique set of hydraulic parameters was used. Model skill improves as points near the origin. Of the three models, AQUIFER is the least sensitive to choice of parameters. Its robustness likely results from the buffering capacity of the augmented subsurface reservoir. When this buffering mechanism is absent, adequate simulation of dTWS (i.e., realistic regulation of the flux of water into the soil column) depends entirely on accurate assignation of effective parameters. DEEP is slightly less sensitive to faulty parameter values than is SSOIL.

[16] Figure 3 compares the empirical cumulative distribution functions of parameter sets of the top-scoring 1% of runs with the distributions of the top-scoring 50%. Because the distributions of the top 1% differ from those of the top half, the parameters shown in Figure 3 are sensitive and merit calibration [*Bastidas et al.*, 1999]. More important, the gentler curvature of AQUIFER's parameter distributions indicates a decrease in sensitivity to percent sand, percent clay, porosity, and maximum rate of subsurface runoff with respect to the other two models, whose cumulative distribution curves are sharper and steeper. Application of a two-sided Kolmogorov–Smirnov test confirms that the difference in sensitivity is statistically significant.

[17] In the foreseeable future, for large model domains, the scientific community is unlikely to be able to confidently assign subsurface hydrologic parameters either by direct observation or by calibration against subsurface hydrologic observations. Decreasing the sensitivity of model output to faulty parameter choices is therefore of utmost practical importance for improving model prediction capability. However, if soil texture properties are the dominant control on regional subsurface hydrologic variation in nature, then AQUIFER's lower sensitivity to parameter values is likely problematic, and a significant increase in data collection and subsequent parameter estimation is warranted.

3.3. Does Knowledge of Parameter Ranges Guarantee Reasonable Model Output?

[18] Figure 4 (left) presents the top-scoring 1% of parameter sets for SSOIL and AQUIFER. Within the envelopes created by the top 1%, the best parameter combination is highlighted in black. Figure 4 (right) presents the scores of all parameter sets for which all values fall within the ranges defined by the envelope created by the top 1%. Note that Figure 4 (right) does not only show the scores of the "good" runs, which are clustered close to the origin; it also presents the scores of the runs that used parameter sets that are near those that resulted in the top-scoring 1% of simulations. For instance, "*BAD*" (Figure 4, left, dashed line) is a parameter set that, despite of having values within the envelope, performs very poorly (e.g., for SSOIL, RMSE $\approx$ 0.2 m; for AQUIFER, RMSE $\approx$ 10 m). Data for DEEP is not shown but is qualitatively similar to that shown for SSOIL. For most parameter sets, AQUIFER performs well. However, we show that there exist parameter sets that are adjacent to top-scoring sets but that result in extremely unrealistic model output. Because of parameter interaction, knowledge of approximate parameter values is insufficient to guarantee realistic simulation of dTWS.

4. Conclusions and Implications

[19] When a surrogate optimal parameter set is used, the model with the 3.43-m, 10-layer soil profile; that with the 30-layer, 11.2-m soil profile; and that in which a lumped unconfined aquifer is coupled to the shallow soil profile are equivalently adept at simulating monthly dTWS over the state of Illinois. When knowledge of subsurface hydraulic parameter values is limited, the coupled aquifer model makes CLM significantly less sensitive to errant parameter values; that is, the explicit aquifer representation is the most robust of the three parameterizations. However, knowledge of ranges for individual parameters is insufficient to guarantee realistic simulation of monthly dTWS.

[20] **Acknowledgments.** We are grateful to Yasir Kaheil, MinHui Lo, and two anonymous reviewers for their insightful, constructive comments. Thanks to Bob Scott at the ISWS for the soil moisture data and to Peter Lawrence for the land-surface dataset. Rosero and Gulden thank the NASA GEST Graduate Student Summer Program. Texas Advanced Computing Center supplied computing resources. NASA grants NAG5-10209 and NAG5-12577 and NOAA grant NAO3OAR4310079 provided funding.

References

Bastidas, L. A., H. V. Gupta, S. Sorooshian, W. J. Shuttleworth, and Z. L. Yang (1999), Sensitivity analysis of a land surface scheme using multicriteria methods, *J. Geophys. Res.*, *104*, 19,481–19,490.

Bonan, G. B., et al. (2002), The land surface climatology of the community land model coupled to the NCAR community climate model, *J. Clim.*, *15*, 3123–3149.

Changnon, S., et al. (1988), Relations between precipitation and shallow groundwater in Illinois, *J. Clim.*, *1*, 1239–1250.
Chen, J. L., C. R. Wilson, and K.-W. Seo (2006), Optimized smoothing of Gravity Recovery and Climate Experiment (GRACE) time-variable gravity observations, *J. Geophys. Res.*, *111*, B06408, doi:10.1029/2005JB004064.
Clapp, R. B., and G. M. Hornberger (1978), Empirical equations for some soil hydraulic-properties, *Water Resour. Res.*, *14*, 601–604.
Cosby, B. J., et al. (1984), A statistical exploration of the relationships of soil-moisture characteristics to the physical-properties of soils, *Water Resour. Res.*, *20*, 682–690.
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.
Fan, Y., G. Miguez-Macho, C. Weaver, R. Walko, and A. Robock (2007), Incorporating water table dynamics in climate modeling. 1: Water table observations and the equilibrium water table, *J. Geophys. Res.*, doi:10.1029/2006JD008111, in press.
Hollinger, S. E., and S. A. Isard (1994), A soil moisture climatology of Illinois, *J. Clim.*, *7*, 822–833.
Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar (2000), A catchment-based approach to modeling land surface processes in a general circulation model. 1: Model structure, *J. Geophys. Res.*, *105*, 24,809–24,822.
Lawrence, P. J., and T. N. Chase (2007), Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0), *J. Geophys. Res.*, *112*, G01023, doi:10.1029/2006JG000168.
Liang, X., Z. Xie, and M. Huang (2003), A new parameterization for surface and groundwater interactions and its impact on water budgets with the variable infiltration capacity (VIC) land surface model, *J. Geophys. Res.*, *108*(D16), 8613, doi:10.1029/2002JD003090.
Maxwell, R. M., and N. L. Miller (2005), Development of a coupled land surface and groundwater model, *J. Hydrometerol.*, *6*(3), 233–247.
Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, and L. E. Gulden (2005), A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, *J. Geophys. Res.*, *110*, D21106, doi:10.1029/2005JD006111.
Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, L. E. Gulden, and H. Su (2007), Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data, *J. Geophys. Res.*, *112*, D07103, doi:10.1029/2006JD007522.
Oleson, K. W., et al. (2004), Technical description of the Community Land Model (CLM), *Tech. Note NCAR/TN-461+STR*, 174 pp., Natl. Cent. for Atmos. Res., Boulder, Colo., http://www.cgd.ucar.edu/tss/clm/distribution/clm3.0/TechNote/CLM_Tech_Note.pdf.
Robock, A., et al. (2000), The Global Soil Moisture Data Bank, *Bull. Am. Meteorol. Soc.*, *81*, 1281–1299.
Rodell, M., and J. S. Famiglietti (2001), An analysis of terrestrial water storage variations in Illinois with implications for the Gravity Recovery and Climate Experiment (GRACE), *Water Resour. Res.*, *37*, 1327–1339.
Taylor, K. E. (2000), Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, *106*, 7183–7192.
Yeh, P. J.-F., M. Irizarry, and E. A. B. Eltahir (1998), Hydroclimatology of Illinois: A comparison of monthly evaporation estimates based on atmospheric water balance and soil water balance, *J. Geophys. Res.*, *103*, 19,823–19,837.
Yeh, P. J.-F., and E. A. B. Eltahir (2005), Representation of water table dynamics in a land surface scheme. part I: Model development, *J. Clim.*, *18*, 1861–1880.
York, J. P., M. Person, W. J. Gutowski, and T. C. Winter (2002), Putting aquifers into atmospheric simulation models: An example from the Mill Creek Watershed, northeastern Kansas, *Adv. Water Resour.*, *25*, 221–238.

J. Famiglietti and P. J.-F. Yeh, Department of Earth System Science, University of California, Irvine, 230 Rowland Hall, Irvine, CA 92696-3100, USA.

L. E. Gulden, G.-Y. Niu, E. Rosero, and Z.-L. Yang, Department of Geological Sciences, John A. and Katherine G. Jackson School of Geosciences, University of Texas at Austin, 1 University Station #C1100, Austin, TX 78712-0254, USA. (liang@mail.utexas.edu)

C. S. Jackson, Institute for Geophysics, John A. and Katherine G. Jackson School of Geosciences, University of Texas at Austin, J. J. Pickle Research Campus, Building 196, 10100 Burnet Road, (R2200), Austin, TX 78758-4445, USA.

M. Rodell, Hydrological Sciences Branch, NASA Goddard Space Flight Center, Code 614.3, Greenbelt, MD 20771, USA.