

Improving parameter estimation and water table depth simulation in a land surface model using GRACE water storage and estimated base flow data

Min-Hui Lo,¹ James S. Famiglietti,¹ P. J.-F. Yeh,² and T. H. Syed³

Received 12 February 2009; revised 13 November 2009; accepted 16 December 2009; published 18 May 2010.

[1] Several previous studies have shown the significance of representing shallow groundwater in land surface model (LSM) simulations. However, optimal methods for parameter estimation in order to realistically simulate water table depth have received little attention. The recent availability of Gravity Recovery and Climate Experiment (GRACE) water storage data provides a unique opportunity to constrain LSM simulations of terrestrial hydrology. In this study, we incorporate both GRACE (storage) and estimated base flow (flux) data in the calibration of LSM parameters, and demonstrate the advantages gained from this approach using a Monte Carlo simulation framework. This approach improves parameter estimation and reduces the uncertainty of water table simulations in the LSM. Using the optimal parameter set identified from the multiobjective calibration, water table simulation can be improved due to close dependence of both base flow and total subsurface water storage on the water table depth. Moreover, it is shown that parameters calibrated from short-term (2003–2005) GRACE and base flow data can be validated using simulations for the periods of 1984–1998 and 2006–2007, which implies that the proposed multiobjective calibration strategy is robust. More important, this study has demonstrated the potential for the joint use of routinely available GRACE water storage data and streamflow records to constrain LSM simulations at the global scale.

Citation: Lo, M.-H., J. S. Famiglietti, P. J.-F. Yeh, and T. H. Syed (2010), Improving parameter estimation and water table depth simulation in a land surface model using GRACE water storage and estimated base flow data, *Water Resour. Res.*, 46, W05517, doi:10.1029/2009WR007855.

1. Introduction and Background

[2] Land surface hydrologic processes can record previous atmospheric forcing anomalies, and then manifest the effects of these anomalies in the following season or year. Enhanced knowledge about such land memory processes can improve weather forecasting and climate prediction at seasonal-to-interannual time scales [e.g., Dirmeyer, 2000; Koster *et al.*, 2000a; Koster and Suarez, 2001]. Several studies have shown that deeper soil layers have longer memory of precipitation anomalies than surface layers do [e.g., Liu and Avissar, 1999a, 1999b; Wu *et al.*, 2002; Wu and Dickinson, 2004; Amenu *et al.*, 2005].

[3] The representation of groundwater dynamics in land surface models (LSMs) has received considerable attention in recent years [e.g., Famiglietti and Wood, 1991, 1994; Koster *et al.*, 2000b; Liang *et al.*, 2003; Maxwell and Miller, 2005; Yeh and Eltahir, 2005a, 2005b; Fan *et al.*, 2007; Miguez-Macho *et al.*, 2007; Niu *et al.*, 2007]. These studies have shown the importance of representing shallow ground-

water and its interaction with soil moisture in land surface hydrologic simulations. For regions with shallow groundwater, the distribution of soil moisture in the vertical profile is highly dependent on water table dynamics. Gulden *et al.* [2007] also indicated that the high sensitivity of simulated terrestrial water storage variations to the selected parameter sets could be reduced by incorporating a groundwater representation into LSMs. While optimal methods for estimating groundwater parameter in LSMs have received little attention in the literature thus far, a recent study by Lo *et al.* [2008] demonstrated the advantage of incorporating estimated base flow in addition to streamflow measurements in LSM parameter calibration.

[4] Issues regarding how to best specify parameters in LSMs remain unclear. Several model intercomparison studies such as Project for Intercomparison of Land-surface Parameterization Schemes (PILPS [e.g., Shao and Henderson-Sellers, 1996; Chen *et al.*, 1997; Lohmann *et al.*, 1998; Bowling *et al.*, 2003]) and Model Parameter Estimation Experiment (MOPEX [Duan *et al.*, 2006]) have shown that with the same atmospheric forcing and common model parameters, the same amount of runoff with contrasting base flow and surface runoff compositions can be simulated. This deficiency is related to variations in the partitioning of water storage and runoff among LSMs, which in turn can be attributed to the inability to accurately determine model parameters related to these hydrologic processes. Despite this, the concept of distinguishing various water storages and

¹Department of Earth System Science, University of California, Irvine, California, USA.

²Institute of Industrial Science, University of Tokyo, Tokyo, Japan.

³Department of Applied Geology, Indian School of Mines University, Dhanbad, India.

runoff components has seldom been appreciated in LSMs. Global applications of LSMs with groundwater parameterizations generally assume spatially constant parameters across different climatic and hydrologic regions [Niu et al., 2007]. The disadvantage of this assumption, however, can be critical since no constraints are imposed on the water table depth and base flow simulations [Lo et al., 2008]. The primary reason for this simplification is the lack of appropriate observations and methodology to estimate the parameters on the global scale [Niu et al., 2007; Xie et al., 2007].

[5] New observations of terrestrial water storage (i.e., all of the snow, ice, surface water, soil moisture, and groundwater [Rodell and Famiglietti, 1999; Syed et al., 2008]) from the Gravity Recovery and Climate Experiment (GRACE) satellite mission [Tapley et al., 2004; Wahr et al., 2004] have provided an unprecedented opportunity to constrain LSM-simulated water storage variations. The GRACE mission now provides estimates of variations in terrestrial water storage for areas larger than $\sim 150,000 \text{ km}^2$ [Swenson et al., 2006] and at time scales ranging from 10 days [Rowlands et al., 2005] to monthly. Several previous studies have compared GRACE observations of terrestrial water storage to observations [Rodell et al., 2004b; Swenson et al., 2006; Syed et al., 2007, 2008] and models [Ramillien et al., 2004; Chen et al., 2005; Seo et al., 2006; Syed et al., 2008], with close agreement. Of importance to the present study, Rodell and Famiglietti [2002], Yeh et al. [2006], Rodell et al. [2007], Strassberg et al. [2007], and Swenson et al. [2008] have all demonstrated that GRACE observations, combined with in situ data or model simulations of the water mass above the water table, can be used to estimate groundwater storage variations. In addition, GRACE has also been applied to evaluate [Swenson and Milly, 2006; Niu and Yang, 2006; Niu et al., 2007] or constrain [Zaitchik et al., 2008; Werth et al., 2009] water storage simulations in LSMs. Ramillien et al. [2008] provide a thorough review of GRACE applications in hydrology.

[6] Our previous study [Lo et al., 2008] has indicated that base flow information estimated from observed streamflow data can help constrain water table depth simulations in LSMs. On the basis of this and the availability of GRACE data, here we explore whether parameter estimation and water table simulation can be enhanced by including both GRACE and base flow data in an optimization procedure. Several studies have shown that model parameter uncertainties can be greatly reduced through a multiobjective calibration framework [e.g., Gupta et al., 1998; Yapo et al., 1998; Houser et al., 2001; Crow et al., 2003]. In this study, GRACE total water storage anomalies (i.e., deviations from the long-term mean) and estimated base flow data will be utilized to estimate LSM parameters. In addition, it will be demonstrated that the use of multiobjective calibration can improve simulation of water table depth compared to using only GRACE or base flow data.

2. Model and Data

2.1. Community Land Model 3.0 With a Groundwater Parameterization

[7] The model used in this study is the Community Land Model 3.0 [Bonan et al., 2002; Oleson et al., 2004] coupled with an unconfined aquifer model originally developed by

Yeh and Eltahir [2005a, 2005b] as a flexible module to couple to any LSM to represent shallow water table dynamics. The coupled model is referred to as CLMGW by Lo et al. [2008], in which the water table is interactively linked to the soil moisture model through the exchange of groundwater recharge (i.e., soil drainage flux) and capillary rise at the bottom of the soil column. For detailed descriptions of the physics in the CLMGW, the reader is referred to Yeh and Eltahir [2005a, 2005b]. In the following, only the surface runoff and base flow generation schemes are briefly described for the purpose of this paper.

[8] The surface runoff generation scheme in the CLMGW is the same as the SIMTOP (simple TOPMODEL-based) surface runoff scheme developed by Niu et al. [2005]. It accounts for both saturation-excess and infiltration-excess runoff as described by the following equation [Oleson et al., 2004; Niu et al., 2005]:

$$R_s = F_{\max} e^{-cz} Q_{\text{in}} + (1 - F_{\max} e^{-cz}) \max(0, Q_{\text{in}} - I_{\max}), \quad (1)$$

where R_s [L/T] is the surface runoff, F_{\max} [] is the maximum saturated fraction for a grid cell, c [] is a coefficient for fitting an exponential function to the cumulative distribution function of the topographic index ($c = 0.5$ as suggested by Niu et al. [2005], and will be used throughout this study), z [L] is the water table depth, f [1/L] is the soil decay factor (i.e., the length scale for the exponential decrease in the saturated hydraulic conductivity with depth), Q_{in} [L/T] is the effective precipitation, and I_{\max} [L/T] is the soil infiltration capacity.

[9] On the basis of observed data in eight catchments in Illinois, base flow (Q_{gw}) at the local scale was formulated using the following nonlinear threshold relation by Yeh and Eltahir [2005b]:

$$\begin{aligned} Q_{\text{gw}} &= Q_0(d_0 - d_{\text{gw}}) & \text{if } 0 < d_{\text{gw}} < d_0 \\ Q_{\text{gw}} &= 0 & \text{if } d_{\text{gw}} > d_0 \end{aligned}, \quad (2)$$

where Q_0 [1/T] is the outflow constant inversely proportional to the aquifer residence time, d_0 [L] is the threshold depth at which groundwater runoff is initialized, and d_{gw} [L] is the water table depth. When applying equation (2) to a grid cell in an LSM, the grid-scale groundwater runoff (Q_{gw}) cannot be determined solely from the grid-mean water table depth (d_{gw}) because of the spatial variability of d_{gw} and the nonlinear relationship in equation (2). Yeh and Eltahir [2005b] proposed a statistical-dynamical approach to account for the influence of the subgrid heterogeneity of water table depth on the grid scale Q_{gw} , which can be derived by integrating equation (2) with respect to an assumed statistical distribution of d_{gw} . The resulting grid-scale groundwater runoff Q_{gw} can be written as [Yeh and Eltahir 2005b, equation (5)]

$$\begin{aligned} Q_{\text{gw}} &= \frac{Q_0 \lambda^\alpha}{\Gamma(\alpha)} \left\{ d_0 \left[\frac{(\alpha - 1)!}{\lambda^\alpha} - e^{-\lambda d_0} \sum_{k=0}^{\alpha-1} \frac{(\alpha - 1)!}{k!} \frac{d_0^k}{\lambda^{\alpha-k}} \right] \right. \\ &\quad \left. - \left[\frac{\alpha!}{\lambda^{\alpha+1}} - e^{-\lambda d_0} \sum_{k=0}^{\alpha} \frac{\alpha!}{k!} \frac{d_0^k}{\lambda^{\alpha-k+1}} \right] \right\}, \quad (3) \end{aligned}$$

where $\Gamma(\alpha)$ is the gamma function and α and λ are the shape and scale parameters of the Gamma distribution of water table depth, respectively. The two statistical parameters (α and λ)

are related to each other through the grid-mean water table depth [Yeh and Eltahir, 2005b, equation (4)].

2.2. Water Table Depth and Streamflow Data

[10] In this study, the model simulation domain is in the state of Illinois, where a unique hydrologic data set covering most of the hydrologic variables exists. In the following, the data used in this study for model simulation, calibration, and validation are briefly summarized. Further details of this extensive data set, including the time spans, sampling locations and quality control, are given by Yeh *et al.* [1998], Yeh and Famiglietti [2009], and Lo *et al.* [2008].

[11] Water table depth data consist of monthly observations from 19 shallow groundwater wells scattered throughout Illinois. These wells, maintained by the Illinois State Water Survey, have been used to monitor the response of unconfined aquifers to climatic fluctuations. The unconfined aquifers in Illinois are relatively shallow, with the state-averaged water table depth about 3 m and the spatial distribution among the 19 wells ranging between 1 and 10 m [Eltahir and Yeh, 1999]. The major soil texture around these wells is silt loam [Yeh *et al.*, 1998].

[12] Streamflow data collected by the U.S. Geological Survey consist of daily discharge records at the outlets of the three largest river basins in Illinois: the Illinois River, the Rock River, and the Kaskaskia River. The total drainage area of the three rivers covers more than two thirds of the total area of Illinois. The locations of groundwater well and streamflow networks are shown in Figure 1.

[13] In this study, the 24 year (1984–2007) daily discharge records from the three major basins are weighted by their drainage areas to yield an estimate of average streamflow in Illinois. The 10 wells with complete records for the 1984–2007 period are used for model calibration and validation. To be consistent with the time span of GRACE data, the 2003–2005 spatially averaged monthly water table depth and streamflow data are used for model calibration, while the data from 1984 to 1998 and 2006–2007 are used for model validation.

2.3. GRACE Total Water Storage Anomalies

[14] The most recently released GRACE data (RL04), produced by the Center for Space Research (CSR) at the University of Texas at Austin [Chambers, 2006], were used in this study. The data extend over more than 5 years from August 2002 to January 2008 (excluding June 2003 and January 2004). Spatial smoothing of GRACE data is required to decrease the influence of noisy short-wavelength Stokes coefficients in the water storage estimates. We have tested the sensitivity of the smoothing kernel (300 km and 500 km), and found that the magnitude of GRACE-derived land water storage estimates is rather insensitive to the half width of the kernel used in our study domain (Illinois). As another independent check, global data sets of simulated total water storage from the WaterGAP Global Hydrology Model (WGHM) [Döll *et al.*, 2003; Güntner *et al.*, 2007] and the Global Land Data Assimilation System (GLDAS) [Rodell *et al.*, 2004a] are also used to test the Gaussian filtering bias. Results (not shown here) indicate that both the phase and amplitude of total water storage anomalies match well between the cases of 300 km smoothing and the one-degree model simulations (i.e., no filtering adopted) averaged over

Illinois. Moreover, we have also tested total water storage variations in the WGHM (no filtering adopted) for the regions surrounding Illinois with areas 2 and 4 times greater than that of the state. We found that the storage variations averaged over the larger surrounding regions are close to those within Illinois. This indicates that the 300 km smoothing radius is suitable for use in this study domain. Therefore, throughout this study, a Gaussian averaging kernel with a half width of 300 km is adopted. Spatially averaged total water storage anomalies over Illinois from 2003 to 2005 are used for the calibration of CLMGW parameters, while the data from 2006 to 2007 are used for validation.

2.4. Atmospheric Forcing Data

[15] To drive the CLMGW model in an off-line simulation, seven input atmospheric forcings are required: precipitation, downwelling solar radiation, downwelling longwave radiation, near-surface air temperature, air humidity, air pressure, and wind speed. Precipitation and temperature were taken from the National Climate Data Center (NCDC; <http://www.ncdc.noaa.gov/oa/ncdc.html>) Integrated Surface Hourly data set. Seventeen NCDC stations uniformly distributed in Illinois were used to derive state mean values by simple averaging. Air humidity, pressure, wind speed, and solar radiation are taken from National Centers for Environmental Prediction/Department of Energy (NCEP/DOE) 6 hourly reanalysis data [Kanamitsu *et al.*, 2002] and linearly interpolated to the 3 hourly resolution. A state-wide average (87.5°W–90.5°W, 37°N–42.5°N) of the NCEP/DOE Reanalysis Data was retrieved to derive the spatially averaged forcing in Illinois. In addition, solar radiation is bias corrected by adjusting the monthly mean to be consistent with the NASA Surface Radiation Budget data set (http://eosweb.larc.nasa.gov/PRODOCS/srb/table_srb.html). The above forcings were used for the 1984–2005 simulations. For the 2006–2007 simulations, the seven input atmospheric forcings were taken from the North American Land Data Assimilation System (NLDAS) [Cosgrove *et al.*, 2003]. Illinois state-averaged forcing data are used to drive the CLMGW as a single-point simulation for the Monte Carlo simulations presented in section 3.

2.5. Base Flow Estimation

[16] Following Lo *et al.* [2008], a digital recursive filter technique was adopted here to separate base flow from daily streamflow records in Illinois. For the equations used in the digital filter technique, see Lo *et al.* [2008] and the references cited below. The digital recursive filter technique has gained popularity in the recent hydrologic literature [e.g., Nathan and McMahon, 1990; Chapman, 1991; Arnold *et al.*, 1995; Mugo and Sharma, 1999; Eckhardt, 2005]. These studies have indicated that the digital recursive filter technique is efficient, reproducible, and objective. The performance of the digital recursive filter technique has been considered as satisfactory as other traditional hydrograph separation approaches [Arnold *et al.*, 1995, 2000; Mau and Winter, 1997] and physically based simulations of base flow [Szilagyi, 2004]. Although the digital filter technique lacks a physical basis, it is more objective and easier to implement than traditional graphical separation techniques. Considering the potential global implementation of the CLMGW model, the digital filter approach is perhaps the only method

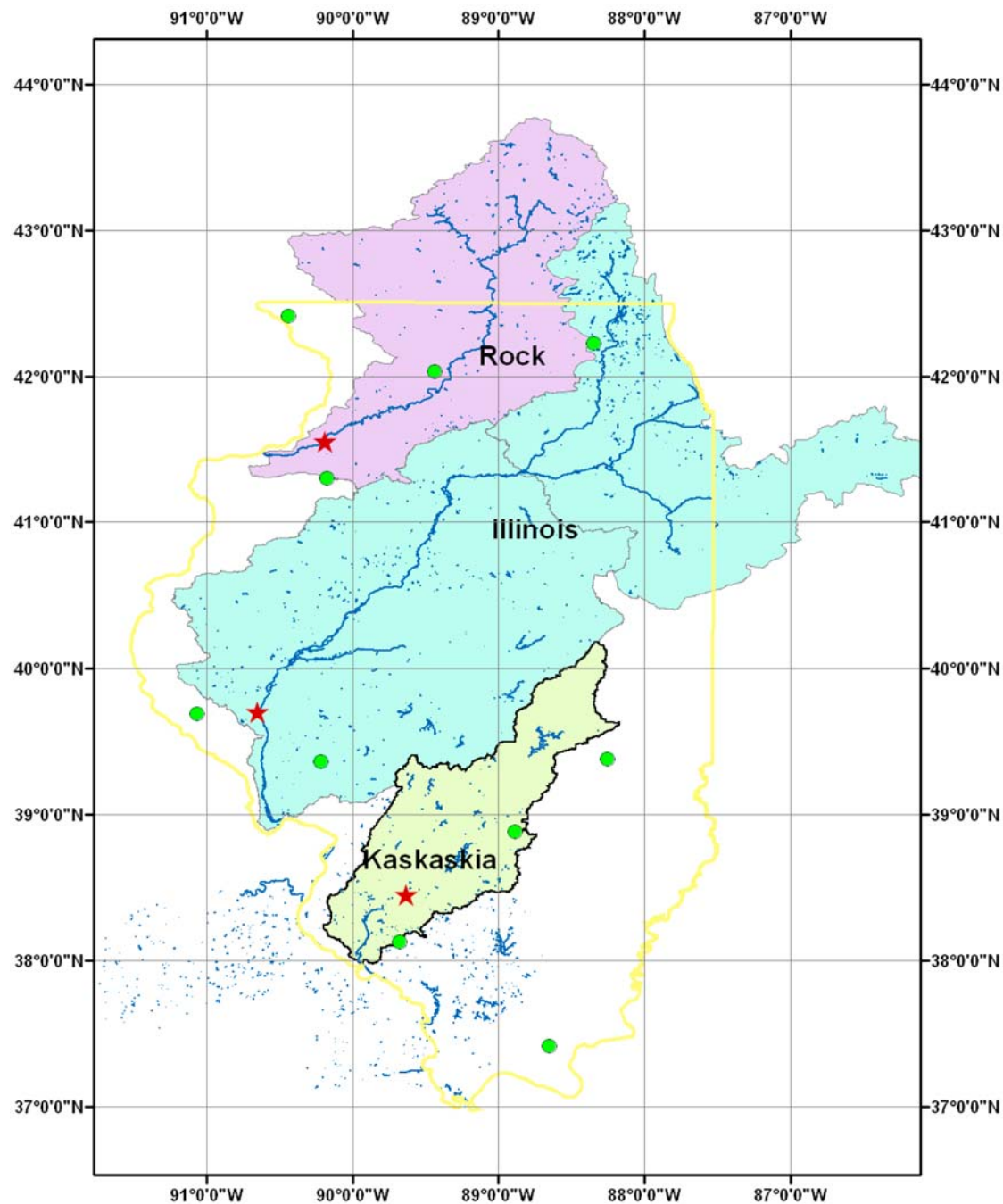


Figure 1. Locations of observational networks of groundwater well (green circles) and streamflow gauges (red stars), and the three largest river basins in Illinois.

currently suitable for this purpose. It yields a first-order estimate of base flow without being limited by data availability, which is extremely important for global-scale land surface modeling.

3. Multiobjective Calibration Using GRACE Data and Base Flow Estimates

[17] Swenson *et al.* [2006, Figure 4] showed a close agreement between the total water storage anomalies derived from GRACE and the combined in situ soil moisture and groundwater measurements in Illinois, although

GRACE underestimated the peaks in observed total water storage. Yeh *et al.* [1998] and Rodell and Famiglietti [2001] have shown that groundwater storage change in Illinois is equal in magnitude to soil moisture change, and both are typically the largest components of monthly terrestrial water storage variations in Illinois relative to snow and surface water.

[18] Figure 2 presents scatterplots between the estimated base flow and GRACE total water storage anomalies against the observed water table depth averaged over the Illinois for the period of 2003–2005. The 22 year (1984–2005) estimated base flow and water table depth is also plotted in

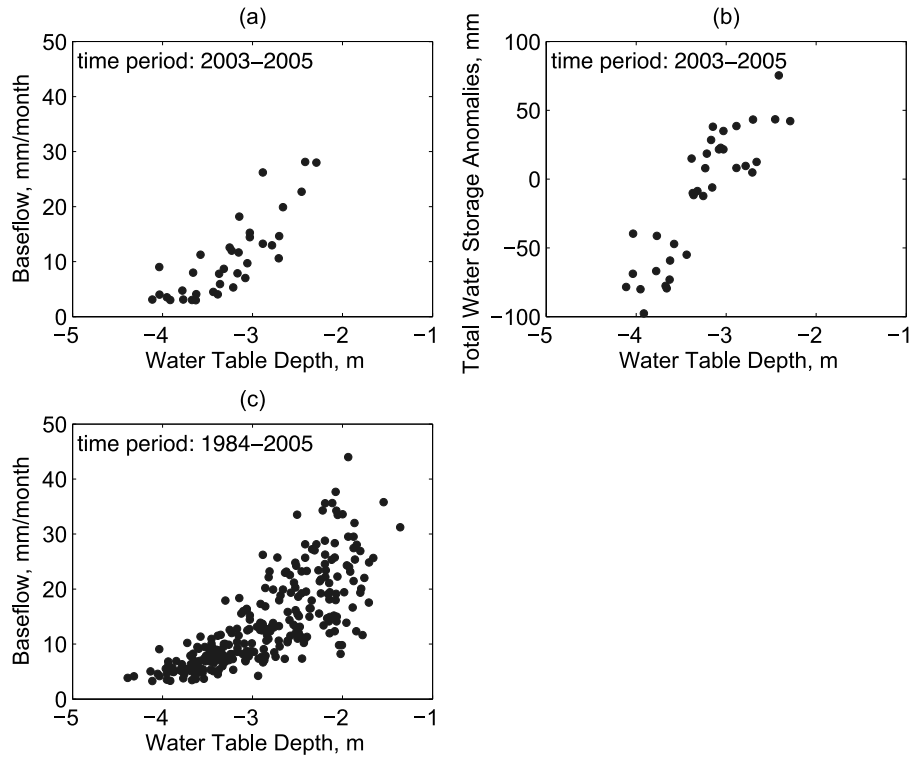


Figure 2. Scatterplots of observed water table depth versus (a) estimated base flow from spatially average daily streamflow data in Illinois and (b) Gravity Recovery and Climate Experiment (GRACE) total water storage anomalies over the Illinois for the period of 2003–2005. (c) Scatterplots of water table depth and estimated base flow for the 22 year (1984–2005) period.

Figure 2c for comparison. As shown, a high correlation exists between these three variables: The correlation coefficient is 0.82 (p-value < 1%) between base flow and water table depth and 0.87 (p-value < 1%) between total water storage anomalies and water table depth, both for the period of 2003–2005, and 0.76 (p-value < 1%) between base flow and water table depth for the 22 year (1984–2005) period.

[19] Motivated by the strong correlation shown in Figure 2, in this study we utilize the routinely measured GRACE data and estimated base flow simultaneously in model parameter calibration to better constrain water table depth simulations in the CLMGW. In the following, the feasibility of the proposed approach will be tested within a Monte Carlo simulation framework, whereby the calibration parameters are randomly sampled from their physically reasonable ranges. The additional advantage gained from this approach relative to the single-objective calibration strategy using only base flow or GRACE water storage will be demonstrated based on the evaluation of water table depth simulations.

3.1. Calibration Parameters

[20] Parameters related to water storage and runoff generation in the CLMGW can be categorized into the following groups: (1) groundwater parameters, which are base flow parameters d_0 and Q_0 (equations (2) and (3)), the parameter characterizing the statistical distribution of water table depth, α (equation (3)), and the specific yield of the aquifer, S_y ; and (2) soil parameters, which are surface runoff- and infiltration-related parameters c , f , and F_{\max} (equation (1)),

and the soil pore size index B from the water-retention equation of *Clapp and Hornberger* [1978], i.e.,

$$\begin{aligned} K &= K_{\text{sat}} \left(\frac{\theta}{\theta_{\text{sat}}} \right)^{2B+3} \\ \psi &= \psi_{\text{sat}} \left(\frac{\theta}{\theta_{\text{sat}}} \right)^{-B}, \end{aligned} \quad (4)$$

where K [L/T] is the hydraulic conductivity, K_{sat} [L/T] is the saturated hydraulic conductivity, θ [%] is the soil water content (%), θ_{sat} [%] is the soil water content at saturation (i.e., porosity), B [%] is the empirical soil pore size index, ψ [L] is the soil water potential, and ψ_{sat} [L] is the saturated soil water potential.

[21] Consistent with previous work [Lo *et al.*, 2008], the parameters with the largest influence on runoff generation and water table dynamics were f , B , d_0 , and Q_0 . Only these parameters are treated as calibration parameters in this study in order to keep the number of total simulations economic. Table 1 summarizes the reasonable physical ranges of the calibration parameters. These parameters have combined direct and indirect influences upon the partitioning of runoff components as well as the simulation of other hydrologic variables (infiltration, surface runoff, groundwater recharge, soil moisture, and groundwater storages). The general sensitivities of model simulations of land surface hydrologic variables can be summarized as follows: (1) Parameter f critically controls the infiltration rate and hence determines surface runoff and the available water for groundwater recharge, which balances base flow over the long-term; (2) B

Table 1. Feasible Ranges of Four Calibration Parameters in the CLMGW Model Used in the Monte Carlo Simulation Framework

Parameter	Unit	Ranges
Decay factor, f	per meter	0.5–3.0
Clapp and Hornberger, B		5–11
d_0	m	0.5–3.5
Q_0	per month	10–280

affects the temporal distribution of groundwater recharge more so than its amount (determined by f); (3) d_0 significantly controls both the amount and temporal variation of base flow; it also determines the long-term average water table depth, which further indirectly feeds back to affect the amount of surface runoff via equation (1); and (4) Q_0 affects the temporal variation of base flow rather than the amount (determined by d_0). It should be mentioned that due to water balance constraints, the calibration parameters d_0 and Q_0 exhibit a strong compensating effect. Given a specific amount of average base flow (determined by the residual of precipitation, evapotranspiration, and surface runoff), a larger (smaller) d_0 would facilitate (inhibit) base flow generation, and thus a higher tendency for a smaller (larger) Q_0 to be optimized. Figure 3 schematically illustrates these direct and indirect feedbacks.

[22] It is well known that various model parameter combinations can reproduce equally well simulations of total runoff due to parameter interactions, but the flow partitioning, the simulations of water table depth, and other hydrologic variables can be rather different, i.e., the equifinality problem from which most hydrologic models have suffered [e.g., Sorooshian and Gupta, 1983; Beven and Binley, 1992; Beven and Freer, 2001]. In order to reduce the equifinality difficulty, it is necessary to calibrate model parameters simultaneously with respect to multiple objectives in order to have realistic partitioning among various water flux and storage components. In section 3.2, we will demonstrate the improvement in water table depth simulation obtained from utilizing a multiobjective calibration framework, i.e., the combination of GRACE water storage and estimated base flow data into the parameter estimation procedures of the CLMGW.

3.2. Monte Carlo Simulations

[23] A Monte Carlo simulation framework is adopted here to demonstrate whether a multiobjective calibration strategy can better constrain water table simulations in the CLMGW. We have conducted 10,000 simulations by a Monte Carlo search over the feasible parameter space. Table 1 summarizes the physically reasonable ranges of four calibration parameters (f , B , d_0 , and Q_0). Each parameter value was randomly (independently) sampled from uniform distributions that span the feasible parameter space (Table 1). All 10,000 simulations were driven by the 10 year (1996–2005) 3 hourly atmospheric forcing data for Illinois described in section 2.4. In order to remove the effect due to uncertain initial conditions, the 1996–2002 period was treated as spin-up, and only the 3 year (2003–2005) simulations were used in the analysis. Each of the 10,000 simulations was scored by a cost function (F) defined to be the weighted sum of the

normalized root mean square error ($NRMSE$) of base flow and total water storage:

$$F = R_b * NRMSE_b + (1 - R_b) * NRMSE_t$$

$$NRMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (o_i - p_i)^2} \frac{1}{\sigma(o)} \quad (5)$$

where the subscript b denotes base flow, t denotes the total water storage anomalies, n is the number of sample points, o is the observation, p is the model estimate, and $\sigma(o)$ is the standard deviation of observations. The weighting ratio R_b is uniformly distributed from 0 to 1 with intervals of 0.1. If R_b is taken as zero or one, F becomes single-objective. The optimal R_b will be determined as the value that minimizes F in equation (5).

[24] Figure 4 shows the mean and standard deviation of the $NRMSE$ for the water table depth simulations for the top-scoring 0.5% (50) of the total 10,000 runs for different weighting ratio R_b . The single-objective calibrations are at the edges of Figure 4: $R_b = 1$ corresponds to base flow calibration, whereas $R_b = 0$ to the total water storage calibration. As seen in Figure 4, the minimum water table depth $NRMSE$ is found to be at $R_b = 0.7$; therefore the combination of 70% of estimated base flow information and 30% of GRACE total water storage information results in the optimal calibration. Compared to the single-objective calibration using base flow (total water storage) only, the multiobjective calibration improves the $NRMSE$ by about 41% (80%), which suggests that water table depth simulation can be further improved by combining information from these two sources. Crow *et al.* [2003] indicated that the choice of the best threshold value (0.5%) from the Monte Carlo simulations is arbitrary. We therefore have also tested the best 0.25%, 1%, 1.5%, 2%, and 2.5% thresholds. Results were found to be insensitive to these changes.

[25] Figure 5 shows the 3 year (2003–2005) monthly time series of simulated water table depth, total water storage anomalies (i.e., deviations from the 3 year mean), and base flow of the best 0.5% (50) runs for the cases of $R_b = 0$ (total water storage calibration alone), $R_b = 1$ (base flow calibration alone), and $R_b = 0.7$ (identified optimal weighting factor). For comparison, observed monthly water table depth, total water storage, base flow, and streamflow are also plotted in Figure 5. It can be seen from Figure 5a that

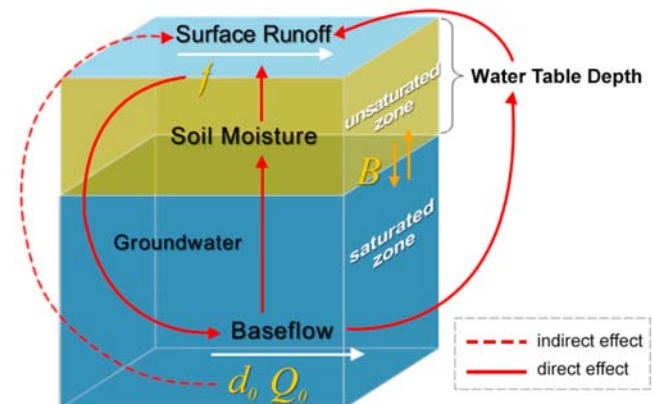


Figure 3. Schematic of the direct and indirect feedbacks for the four parameters (f , B , d_0 , and Q_0).

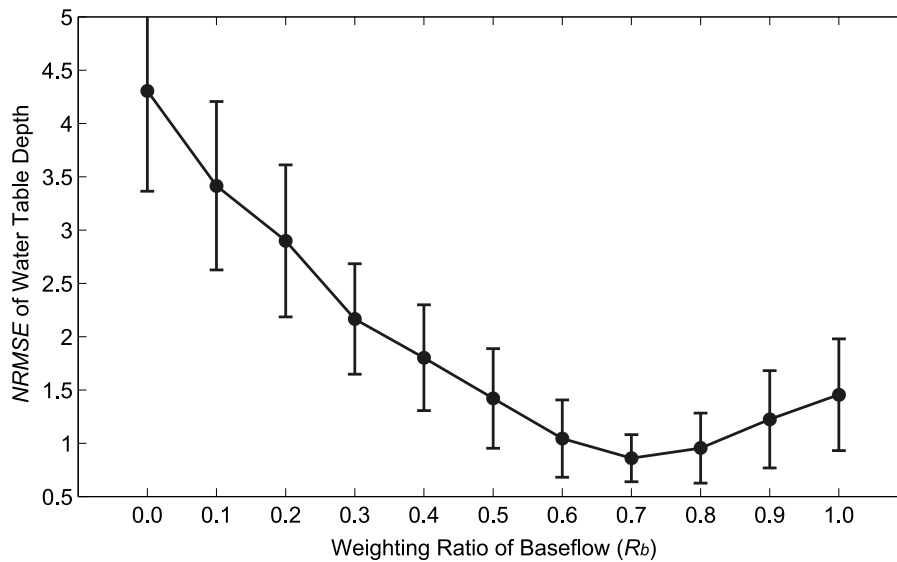


Figure 4. Mean and standard deviation of water table depth $NRMSE$ for the top-scoring 0.5% of the cost function (F) from the 10,000 runs with a specific R_b choice. The analysis is for 2003–2005.

when using multiobjective calibration, water table depth simulations converge to the observation and exhibit less spread compared to the two single-objective calibrations. In addition, when calibrating with GRACE data alone, the performance of total water storage anomaly simulations is better than the other two cases as clearly shown in Figure 5b. On the other hand, the base flow calibration alone gives better base flow simulations (Figure 5c).

[26] Simulations of total water storage anomalies for the case of $R_b = 0.7$ are better constrained than that of $R_b = 1$, which can be attributed to incorporating 30% of the constraint by GRACE data (Figure 5b). Figure 5 also shows a tendency that when using multiobjective calibration, simulations are closer to that by using base flow calibration. The reason is that the optimal weighting factor is the combination of 70% of base flow and 30% of GRACE data, i.e., the base flow provides more information in the multiobjective framework in this case. In addition, Figure 5 shows that the water table depth simulations have wider spreads compared to the total water storage anomalies and base flow simulations for different calibration cases. This also indicates that the regional water table depth simulation is relatively more difficult to be constrained in LSMs than total water storage simulation.

[27] Figure 5b shows that the simulated total water storage variations are larger than the GRACE signal, particularly after 2004. *Lo et al.* [2008] have shown that the seasonal cycle of water table depth becomes progressively flatter as the simulated mean water table depth increases. Correspondingly, the simulated total water storage also shares the same characteristic, i.e., the storage variations tends to be smaller when mean water storage decreases. Therefore, if only GRACE data are used in the calibration, in order to yield better storage simulation, the model tends to simulate lower water storage resulting in a lower water table depth, as shown in Figure 5a. Moreover, notice that the simulated base flow sometimes even exceeds observational total runoff (Figure 5c) if the model is only calibrated using GRACE,

although it does improve total water storage simulations as seen most notably in the second half of 2005 (Figure 5b).

[28] In general, GRACE provides monthly variations of water storage rather than the absolute value so the monthly variations of water table depth can be better captured by incorporating GRACE data, but not the location of mean water table depth [*Lettenmaier and Famiglietti, 2006*]. Therefore it is necessary to incorporate base flow information, which gives a first-order constraint on mean water table depth [*Lo et al., 2008*], while using GRACE to constrain the groundwater storage behavior.

[29] Figure 6 plots the distribution of parameter values for the best 0.5% (50) runs for the cases of $R_b = 0$, $R_b = 1$, and $R_b = 0.7$, respectively. The values of four calibration parameters are normalized by the corresponding maximum values within the feasible ranges as listed in Table 1 (with 1 denoting the maximum). From the comparison of two single-objective calibration cases (Figure 6), it is clearly seen that the parameters f and B (d_0 and Q_0) exhibit higher identifiability for the case of total water storage (base flow) calibration. Figure 6a shows that the optimal value of f is small for the best 50 runs for the GRACE calibration case; namely, using any combination of another three parameters with a low f from those plotted in Figure 6a, the water table variations can be simulated well. However, when f is low, surface runoff is high (because of higher saturated fractional area; see equation (1)) and less infiltration into the soil, resulting in a deeper water table depth [*Lo et al., 2008*], which explains why GRACE-only calibration tends to adjust to a lower mean water table depth as shown in Figure 5b.

[30] Note that parameters d_0 and Q_0 show a strong compensating effect in their optimal values, i.e., a large d_0 with a small Q_0 , and vice versa. Their compensating effect on the water table depth can be clearly seen from Figure 7, where three groups of monthly groundwater rating curves of the best 0.5% runs for the cases of $R_b = 0$, 0.7, and 1 are compared to the observed (state-averaged) groundwater rating curve. The groundwater rating curve, as in equation (2)

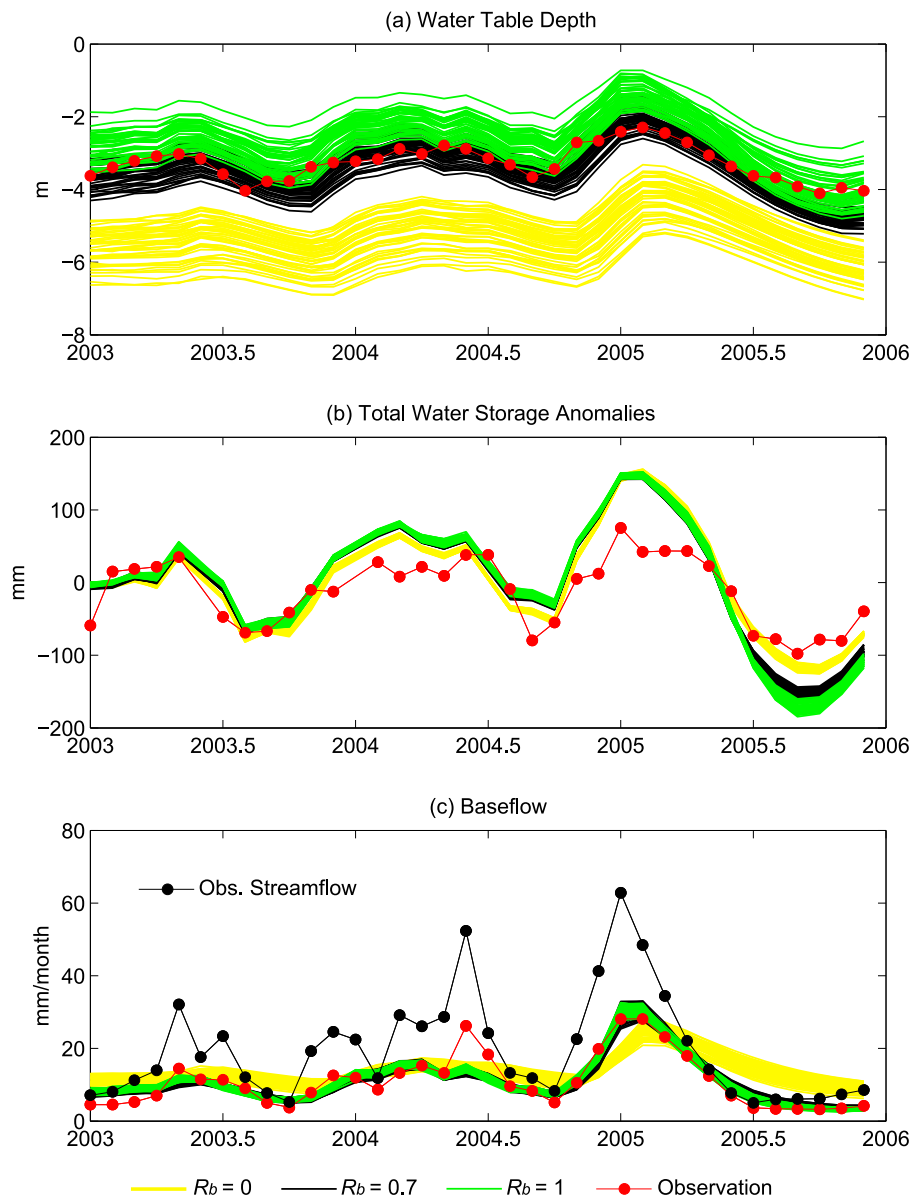


Figure 5. Three-year (2003–2005) monthly time series of simulated (a) water table depth, (b) total water storage anomalies (i.e., deviations from the 3 year mean), and (c) base flow of the best 0.5% (50) runs for the cases of $R_b = 0$ (total water storage calibration alone), $R_b = 1$ (base flow calibration alone), and $R_b = 0.7$ (identified optimal weighting factor).

for the local scale and equation (3) for the grid scale, defines a one-to-one relationship between base flow (Q_{gw}) and water table depth (d_{gw}) [Yeh and Eltahir, 2005b]. The group of green rating curves is plotted from equation (3) with various combinations of d_0 and Q_0 calibrated using base flow; the red curves calibrated using both GRACE and base flow; and the yellow curves calibrated using GRACE. Black dots are the “observed” base flow estimated from the digital filter technique, which are plotted versus observed water table depth. Three distinct groups of calibrated rating curves can be discerned: The red curves cross the observed data, whereas the green and yellow form the lower and upper bounds, respectively, that envelope the observed rating curve. Therefore the water table depth is shallower (deeper)

than observed, as shown in Figure 5a, when using only base flow (GRACE) in calibration. Moreover, the yellow and green rating curves exhibit significantly wider spreads compared to the red curves, which is also reflected in the simulated water table depth of single-objective calibration, and increases the uncertainty of model simulations of single-objective calibration as can be seen from the standard deviation of *NRMSE* plotted in Figure 4.

[31] Taken together, the above implies that after the incorporation of both base flow and GRACE information, the identifiability of calibration parameters (f , B , d_0 , and Q_0) is enhanced, and hence a more accurate groundwater rating curve can be derived with less uncertainty. Moreover, some of the calibration parameters adjust to converge toward their

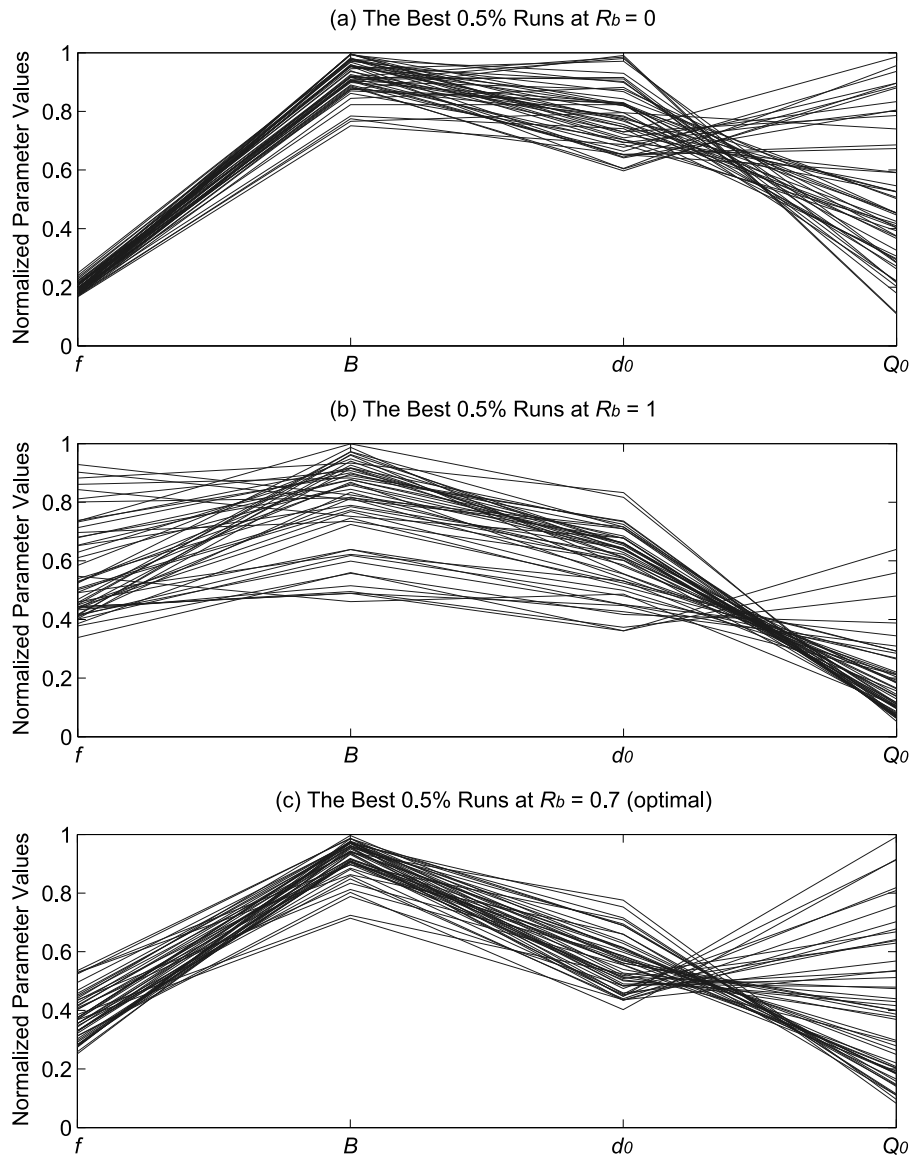


Figure 6. Normalized parameter values for the four parameters of the best 0.5% (50) runs for the cases of (a) $R_b = 0$ (total water storage calibration alone), (b) $R_b = 1$ (base flow calibration alone), and (c) $R_b = 0.7$ (identified optimal weighting factor). The values of four calibration parameters have been normalized by the maximum in their feasible ranges as listed in Table 1. The analysis is for 2003–2005.

optimal ranges when the additional objective is incorporated. This not only improves model simulations but greatly enhances parameter identifiability and alleviates the problem of equifinality.

4. Validation

[32] It is important to validate whether the optimal parameter set identified from 2003 to 2005 simulation can reproduce the observations of water table depth, streamflow, and total water storage in another preferably longer time period. The split-sample approach is therefore utilized here for this purpose. Since GRACE data are available only after August 2002, two additional simulations (1984–1998 and 2006–2007) were conducted. The first simulation was in fact from 1975 to 1998, with 1975–1983 treated as the spin-up

period, while the second was from 1998 to 2007, with 1998–2005 treated as spin-up.

[33] The ratio (R_b) of the cost function (F) obtained from the Monte Carlo simulation framework is 0.7 so the cost function (F) can be rewritten as

$$F = 0.7 * NRMSE_b + 0.3 * NRMSE_t. \quad (6)$$

We further use the Shuffled Complex Evolution (SCE) algorithm [Duan *et al.*, 1992] to determine the single optimal parameter set for the validation test. The simulation period for the SCE test is 2003–2005. The major advantages of SCE are the competitive evolution and complex shuffling, which can avoid the tendency of falling into local minima [Duan *et al.*, 1993]. While using the SCE to minimize F in equation (6), the sampled parameter values span over the ranges listed in Table 1. Consequently, the following unique parameter set

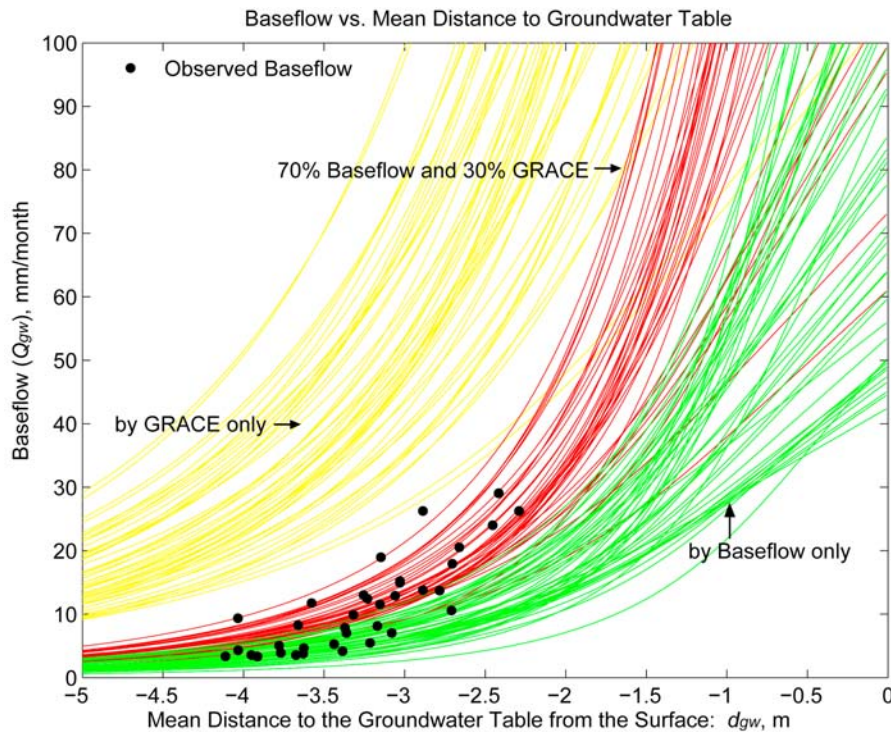


Figure 7. Three groups of groundwater rating curves (i.e., average water table depth d_{gw} versus base flow Q_{gw}) obtained from the top-scoring 0.5% runs of the base flow-only calibration, 30% GRACE and 70% base flow calibration, and GRACE-only calibration, respectively, compared to the observed groundwater rating curve in Illinois. Green rating curves are plotted from equation (3) with various combinations of d_0 and Q_0 calibrated from estimated base flow data; the red curves are calibrated from both GRACE and base flow data; the yellow curves are calibrated from GRACE data. Black dots are the scatterplot of observed water table depth versus estimated base flow in Illinois. The analysis is for 2003–2005.

was identified: $f = 1.07 \text{ m}^{-1}$, $B = 8.87$, $d_0 = 1.59 \text{ m}$, and $Q_0 = 155 \text{ month}^{-1}$. Because of the optimal searching process in SCE, the use of this algorithm will tend toward obtaining more base flow information in the parameter space, and thus leads to better water table simulations compared to the best run from the 10,000 Monte Carlo simulations (i.e., the root mean square error of water table depth simulation improves from 0.35 m to 0.28 m).

[34] Figures 8a, 8b, and 8c compare simulated monthly water table depths, base flow, and total runoff to observations for both periods of 1984–1998 and 2006–2007. Figure 8d shows simulated total water storage anomalies for the 2006–2007 period compared to GRACE data. In general, simulations using the optimal parameter set reproduce the observed monthly variability of base flow and water table depth reasonably well, in particular the anomalously low (high) base flow and water table depth in the extreme 1988 drought (1993 flood) conditions in the U.S. Midwest. The root mean square error and the correlation coefficient for base flow and water table depth simulations are 5.21 mm/month (0.83, p-value < 1%) and 0.48 m (0.82, p-value < 1%), respectively, for the entire 17 year simulation. However, obvious underestimation in the total runoff simulation is apparent in Figure 8c, most likely due to the underestimation of surface runoff. *Lo et al.* [2008] explain that when average atmospheric data are used as model input forcing, part of the temporal variability in precipitation are smoothed out; as a result surface runoff is underestimated. Although

the optimal parameter set based on base flow and total water storage calibration is used in Figure 8, it may not improve the surface runoff simulation due to the different flow generation mechanisms and controlling parameters. Figure 8d also shows that monthly variations of the total water storage anomalies are well captured, except for the first half of 2007.

[35] The longer-term (1984–1998) validation test has shown that the identified optimal parameter set from the multiobjective calibration can be applied beyond the calibration period. This also implies that the multiobjective calibration strategy demonstrated here has the potential to enhance the robustness of model parameter estimation even when the period of calibration data is relatively short, as the use of GRACE data in this study.

5. Evaluation of Soil Moisture Simulation

[36] Figure 9a shows the simulated soil moisture profile averaged from 1984 to 2003 for the cases of $R_b = 0, 0.7,$ and 1.0 from the best 50 runs of the Monte Carlo simulations compared to the observed spatially averaged soil moisture profile. The default CLM3.5 (with groundwater model developed by *Niu et al.* [2007]) simulation is also plotted for comparison. As seen in Figure 9a, none of these four simulations can reproduce the observed vertical soil moisture profile. *Zeng and Decker* [2009] also have shown that the NCAR CLM has limitations in reproducing the vertical soil moisture profile and have proposed a modified Richards

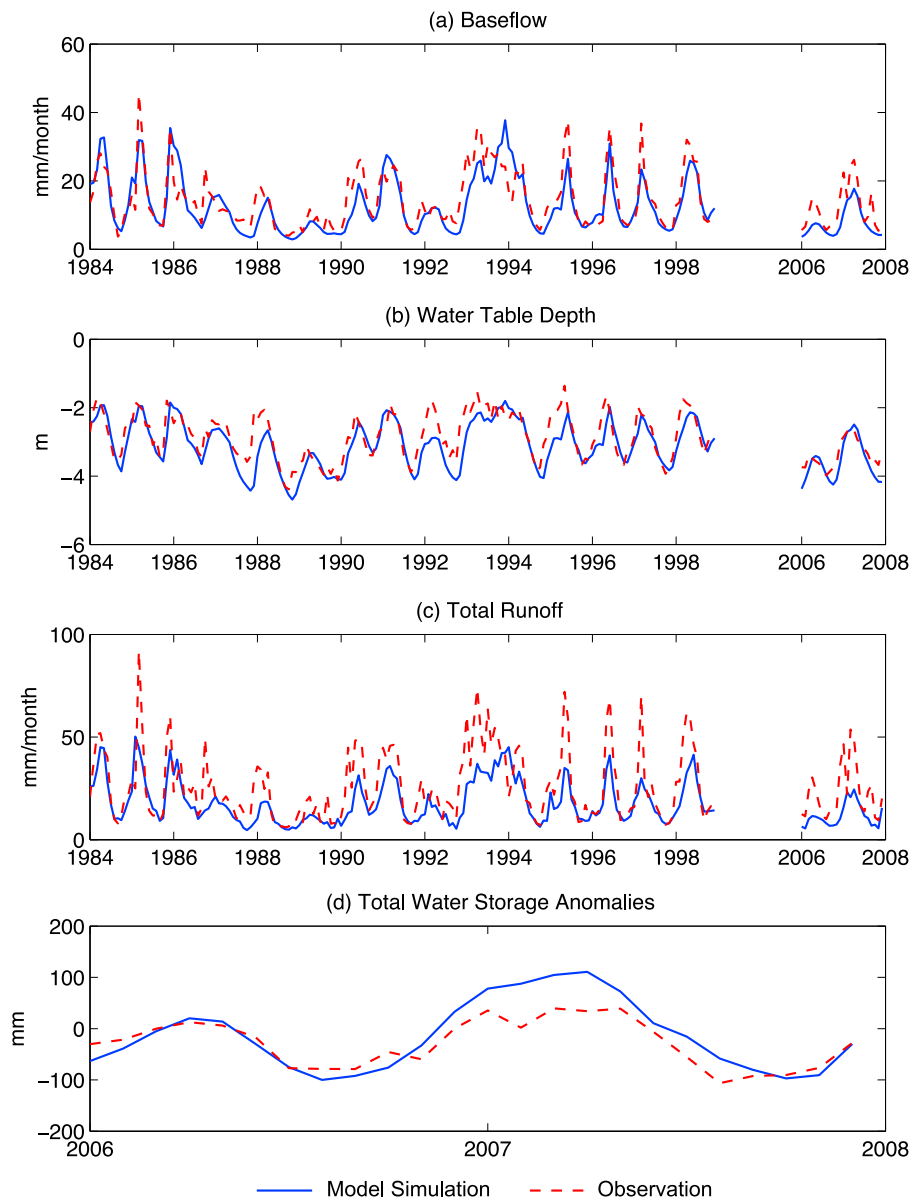


Figure 8. The 17 year (1984–1998 and 2006–2007) monthly time series of simulated (a) base flow, (b) water table depth, and (c) total runoff using the optimal parameter set identified from the multiobjective calibration in comparison with observations. (d) Simulated total water storage anomalies for the 2006–2007 period compared to the GRACE data.

equation as a potential solution. Since our focus in this study is on the improvement of water table depth simulations, soil moisture is not calibrated. However, the parameters in the modified Richards equation could be considered as additional calibration targets if the improvement of soil moisture simulation is a major concern.

[37] Moreover, the results from three calibration cases are consistent with those in water table depth simulations (Figure 5a). Since the water table depth is shallower (deeper) than the observation while using base flow (GRACE) alone in calibration, the resulting soil moisture is wetter (drier).

[38] Figures 9b and 9c show the seasonal cycle comparisons of the observations, CLM3.5, and the average from the best 50 runs of the Monte Carlo simulations using CLMGW. As shown, in all simulations the seasonal amplitude of surface layer soil moisture (0–0.5 m) is too small to capture the

observed variability. For deep layers (0.5–1.6 m), however, the case of $R_b = 0.7$ well reproduces the observed seasonal variations in term of the phase and amplitude. For the default CLM3.5, the seasonal amplitude is too small for the surface layers and too large for the deep layers.

6. Discussion and Conclusions

[39] Although several recent land surface modeling studies have demonstrated the importance of water table dynamics and various groundwater parameterizations have been developed, the problem of how to best specify groundwater parameters for realistic simulations of water table depth dynamics has received little attention. Here, we use GRACE total water storage data combined with estimated base flow data in the model calibration. This approach

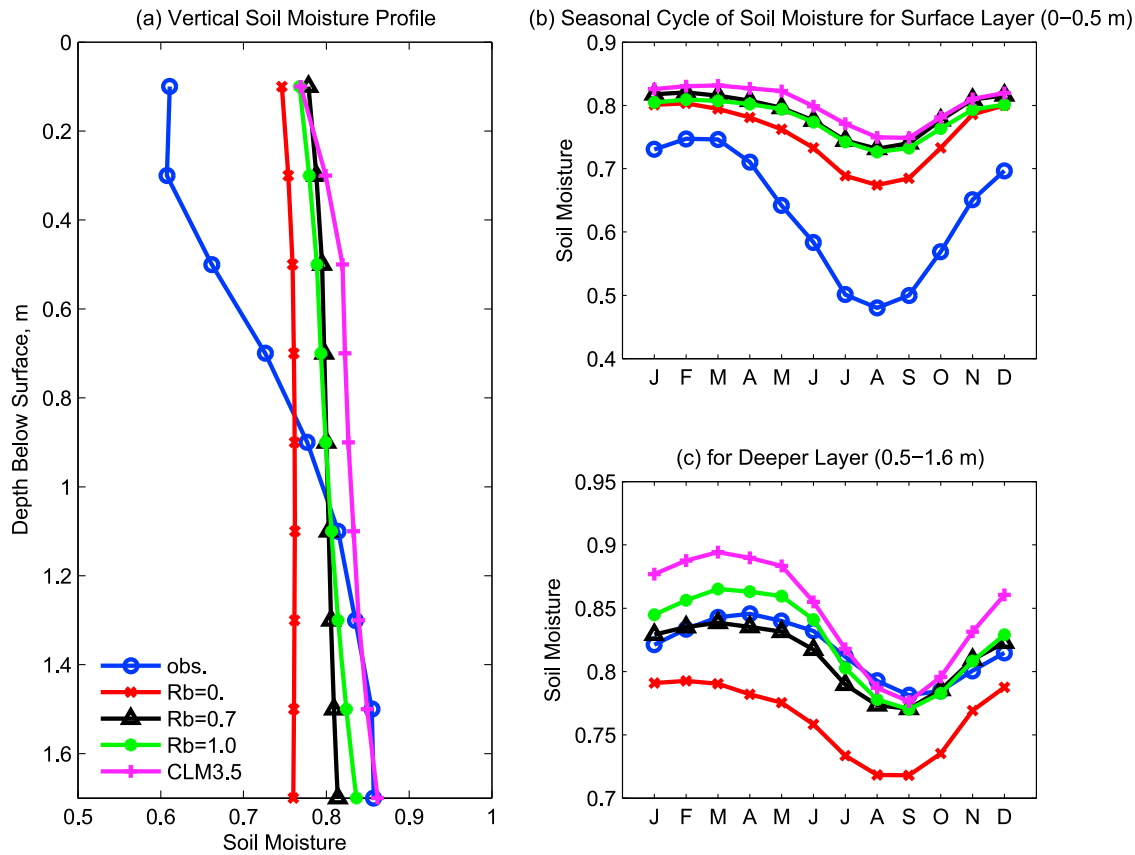


Figure 9. (a) Simulated soil moisture profile averaged from 1984 to 2003 for the cases of $R_b = 0, 0.7$ and 1.0 , from the best 50 runs of the Monte Carlo simulations compared to the observed spatially averaged soil moisture profile. The default CLM3.5 simulation is also plotted for comparison. (b) Seasonal cycle comparisons of the observations, CLM3.5, and the average from the best 50 runs of the Monte Carlo simulations using CLMGW for the surface layer (0–0.5 m) and (c) for the deeper layer (0.5–1.6 m).

improves parameter estimation, i.e., enhances parameter identifiability, and reduces the uncertainty of water table simulations in an LSM. Using the optimal parameter set identified from the multiobjective (base flow and total water storage in this study) calibration, water table simulation can be improved due to the close dependence of base flow and total water storage on the water table depth. It has also been shown that parameters calibrated from the short-term (2003–2005) GRACE and base flow data can be validated for other time periods (1984–1998 and 2006–2007), which implies that the proposed multiobjective calibration strategy is robust. Importantly, this study demonstrates the potential for the joint use of routinely available GRACE and streamflow measurements to constrain LSM simulations at the global scale.

[40] This study shows that total water storage and base flow have two distinct sensitivities among calibrated parameters. The soil parameters f and B have more influence on the total water storage simulation, whereas the groundwater parameters d_0 and Q_0 have more influence on the water table simulations. Hence, while using the dual objectives in the CLMGW calibration, optimal parameter sets can be identified from their respective optimal ranges from which it lead to improved water table simulations.

[41] Crow *et al.* [2003] pointed out that it is necessary to incorporate at least one surface energy flux and one state variable to calibrate LSMs in order to obtain correct simu-

lations. In this study, we use estimated base flow, a flux, and GRACE total water storage, a state, to calibrate CLMGW model. In general, the data on state variables (e.g., soil moisture and water table depth) are extremely difficult to obtain at the regional and global scale, but streamflow records are available in many locations of the world and hence are most commonly used in calibration. Nevertheless, GRACE data can provide unique information on the variations of water storages at the global scale for LSM calibration that is unlikely to be achieved by using steamflow alone. Moreover, it should be noted that since the generation mechanisms of surface runoff and base flow are quite different, calibrations using base flow and water storage information are not necessary to improve surface runoff simulation. If streamflow data are used as an additional calibration target, surface runoff simulation can be better constrained.

[42] In this study, we have demonstrated a convincing strategy of using GRACE and estimated base flow data to constrain LSM simulations. For regions lacking water table observations, the information derived from GRACE and base flow can be used to constrain water table simulation in LSMs, while improving the partitioning between fluxes (surface runoff and base flow) and storage (soil moisture and groundwater) components. Therefore it can be concluded that groundwater parameters in LSMs can be better calibrated by the combined use of GRACE and estimated

base flow data. In this case study using Illinois data, the weighting ratio (R_b) is found to be 0.7, indicating that the contribution to water table simulation from base flow information is larger than that from GRACE. However, this result is not transferable to other regions. In deep groundwater areas where base flow only contributes a small percentage of streamflow, GRACE data may have a more dominant weight. When applying the proposed multiobjective calibration approach at large scales, we suggest that R_b and the resultant calibration parameter sets should be derived regionally in order to account for the spatial heterogeneity of land hydrologic processes. Similarly, when applying the proposed approach where other storage components (soil moisture, snow, lakes or wetlands) than groundwater dominate terrestrial water storage variations, the correlation between water table depth and the GRACE signal may be weaker. Therefore the contribution of GRACE information to groundwater parameter calibration could be smaller (the R_b closer to 1) than the case of Illinois.

[43] The purpose of the multiobjective calibration framework proposed in this study is to constrain water table depth simulations in an LSM. Hence, if the GRACE total water storage signal can be successfully decomposed into different components, the proposed approach would be more robust for constraining groundwater storage variations. However, GRACE signal decomposition is still an open research question. *Frappart et al.* [2008] estimated surface water storages in Negro River basin by using the TOPEX/POSEIDON (T/P) altimetry satellite. Thus the contributions of soil moisture and groundwater to the total storage changes can be isolated from GRACE. In fact, a new NASA satellite mission, Surface Water and Ocean Topography (SWOT), has been proposed to provide high-resolution measurements of surface water bodies such as lakes, wetlands, and rivers. This together with other ongoing satellite missions (e.g., AMSR-E for snow and soil moisture retrieval) hold great potential to help decompose GRACE total water storage signal, and hence are essential for applying the proposed multiobjective calibration framework globally.

[44] It should be noticed that GRACE data provide monthly variations of water storage rather than the absolute values of storages. Therefore, whereas the monthly variations of water table depth can be improved by the calibration with GRACE data, the position of mean water table depth may not be accurately located. Therefore it is necessary to incorporate base flow information to constrain mean water table depth, since a strong nonlinear relationship exists between water table depth and base flow [*Eltahir and Yeh*, 1999]. For regions without streamflow data, GRACE can still provide a first-order constraint on the water storage variations. Recently, *Syed et al.* [2009] used GRACE and reanalysis data to estimate terrestrial freshwater discharge, which can be used in model calibration for those regions without streamflow observations. Moreover, for those regions where the GRACE signals are relatively small, calibration using GRACE data alone will not be effective. Therefore when using GRACE alone to constrain global LSM simulations, it is important to consider the amplitude of total water storage variations.

[45] Finally, the downwelling longwave radiation used in this study was estimated from an empirical formula as a function of near-surface atmospheric vapor pressure and temperature [*Oleson et al.*, 2004]. *Yang et al.* [1997] found a

large uncertainty when estimating longwave radiation using this formula. We compared the estimated longwave radiation to the GLDAS data set and found that the difference in the annual mean (1980–2005) is negligible (322.9 W/m^2 from the empirical formula versus 322.7 W/m^2 from the GLDAS data). However, the estimated longwave radiation is higher (lower) than that of GLDAS data during summer (winter), which results in the higher snow depth during the winter and higher evapotranspiration during the summer. *Yang et al.* [1997] have shown that in order to more accurately estimate longwave radiation, it is necessary to consider cloud type and cloud amount. Phase II of NLDAS, which provides hourly data from January 1979 to present at $1/8^\circ$ resolution over the contiguous United States, is available online. For future research, this data set can be used to provide more accurate atmospheric forcing for offline LSM simulations.

[46] **Acknowledgments.** This research was sponsored by NOAA CPPA grant NA05OAR4310013, NASA Earth and Space Science Fellowship NNX08AV06H, and the U. C. Irvine Institute of Geophysics and Planetary Physics. We express our gratitude to Andreas Güntner, Susanna Werth, and Caroline de Linage for valuable communication and to the Illinois State Water Survey for providing the hydrologic data used here. GRACE data were processed by D. P. Chambers, supported by the NASA Earth Science REASON GRACE Project, and are available at <http://grace.jpl.nasa.gov>. The radiation data were obtained from the NASA Langley Research Center Atmospheric Science Data Center. Computations were supported by Earth System Modeling Facility NSF ATM-0321380.

References

- Amenu, G. G., P. Kumar, and X. Z. Liang (2005), Interannual variability of deep-layer hydrologic memory and mechanisms of its influence on surface energy fluxes, *J. Clim.*, *18*, 5024–5045, doi:10.1175/JCLI3590.1.
- Arnold, J. G., P. M. Allen, R. Muttiyah, and G. Bernhardt (1995), Automated baseflow separation and recession analysis techniques, *Ground Water*, *33*, 1010–1018, doi:10.1111/j.1745-6584.1995.tb00046.x.
- Arnold, J. G., R. S. Muttiyah, R. Srinivasan, and P. M. Allen (2000), Regional estimation of baseflow and groundwater recharge in the upper Mississippi River basin, *J. Hydrol.*, *227*, 21–40, doi:10.1016/S0022-1694(99)00139-0.
- Beven, K. J., and A. M. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Processes*, *6*, 279–298, doi:10.1002/hyp.3360060305.
- Beven, K. J., and J. Freer (2001), Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems, *J. Hydrol.*, *249*, 11–29, doi:10.1016/S0022-1694(01)00421-8.
- Bonan, G. B., K. W. Oleson, M. Vertenstein, S. Levis, X. Zeng, Y. Dai, R. E. Dickinson, and Z.-L. Yang (2002), The land surface climatology of the Community Land Model coupled to the NCAR Community Climate Model, *J. Clim.*, *15*, 3123–3149, doi:10.1175/1520-0442(2002)015<3123:TLSCOT>2.0.CO;2.
- Bowling, L., et al. (2003), Simulation of high-latitude hydrological processes in the Torne-Kalix basin: PILPS Phase 2(e): 1. Experiment description and summary intercomparisons, *Global Planet. Change*, *38*(1–2), 1–30, doi:10.1016/S0921-8181(03)00003-1.
- Chambers, D. P. (2006), Evaluation of new GRACE time-variable gravity data over the ocean, *Geophys. Res. Lett.*, *33*, L17603, doi:10.1029/2006GL027296.
- Chapman, T. (1991), Comment on “Evaluation of automated techniques for baseflow and recession analyses” by R. J. Nathan and T. A. McMahon, *Water Resour. Res.*, *27*, 1783–1784, doi:10.1029/91WR01007.
- Chen, J. L., C. R. Wilson, J. S. Famiglietti, and M. Rodell (2005), Spatial sensitivity of the Gravity Recovery and Climate Experiment (GRACE) time-variable gravity observations, *J. Geophys. Res.*, *110*, B08408, doi:10.1029/2004JB003536.
- Chen, T., et al. (1997), Cabauw experimental results from the Project for Intercomparison of Land-Surface Parameterization Schemes (PILPS), *J. Clim.*, *10*, 1194–1215, doi:10.1175/1520-0442(1997)010<1194:CERFTP>2.0.CO;2.

- Clapp, R. B., and G. M. Hornberger (1978), Empirical equations for some soil hydraulic properties, *Water Resour. Res.*, *14*, 601–604, doi:10.1029/WR014i004p00601.
- 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.
- Crow, W. T., E. F. Wood, and M. Pan (2003), Multiobjective calibration of land surface model evapotranspiration predictions using streamflow observations and spaceborne surface radiometric temperature retrievals, *J. Geophys. Res.*, *108*(D16), 4725, doi:10.1029/2002JD003292.
- Dirmeyer, P. A. (2000), Using a global soil wetness dataset to improve seasonal climate simulation, *J. Clim.*, *13*, 2900–2922, doi:10.1175/1520-0442(2000)013<2900:UAGSWD>2.0.CO;2.
- Döll, P., F. Kaspar, and B. Lehner (2003), A global hydrological model for deriving water availability indicators: Model tuning and validation, *J. Hydrol.*, *270*, 105–134, doi:10.1016/S0022-1694(02)00283-4.
- Duan, Q. Y., S. Sorooshian, and V. K. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, *28*, 1015–1031, doi:10.1029/91WR02985.
- Duan, Q. Y., V. K. Gupta, and S. Sorooshian (1993), Shuffled complex evolution approach for effective and efficient global minimization, *J. Optim. Theory Appl.*, *76*(3), 501–521, doi:10.1007/BF00939380.
- Duan, Q. Y., et al. (2006), An overview of science strategy and major results from the second and third workshops, *J. Hydrol.*, *320*(1–2), 3–17, doi:10.1016/j.jhydrol.2005.07.031.
- Eckhardt, K. (2005), How to construct recursive digital filters for baseflow separation, *Hydrol. Processes*, *19*(2), 507–515, doi:10.1002/hyp.5675.
- Eltahir, E. A. B., and P. J.-F. Yeh (1999), On the asymmetric response of aquifer water level to floods and droughts in Illinois, *Water Resour. Res.*, *35*, 1199–1217, doi:10.1029/1998WR900071.
- Famiglietti, J. S., and E. F. Wood (1991), Evapotranspiration and runoff from large land areas—Land surface hydrology for atmospheric general circulation models, *Surv. Geophys.*, *12*, 179–204, doi:10.1007/BF01903418.
- Famiglietti, J. S., and E. F. Wood (1994), Multiscale modeling of spatially variable water and energy balance processes, *Water Resour. Res.*, *30*, 3061–3078, doi:10.1029/94WR01498.
- Fan, Y., G. Miguez-Macho, C. P. 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.*, *112*, D10125, doi:10.1029/2006JD008111.
- Frappart, F., F. Papa, J. S. Famiglietti, C. Prigent, W. B. Rossow, and F. Seyler (2008), Interannual variations of river water storage from a multiple satellite approach: A case study for the Rio Negro River basin, *J. Geophys. Res.*, *113*, D21104, doi:10.1029/2007JD009438.
- 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.
- Güntner, A., J. Stuck, S. Werth, P. Döll, K. Verzano, and B. Merz (2007), A global analysis of temporal and spatial variations in continental water storage, *Water Resour. Res.*, *43*, W05416, doi:10.1029/2006WR005247.
- Gupta, H. V., S. Sorooshian, and P. O. Yapo (1998), Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resour. Res.*, *34*, 751–763, doi:10.1029/97WR03495.
- Houser, P. R., H. V. Gupta, W. J. Shuttleworth, and J. S. Famiglietti (2001), Multiobjective calibration and sensitivity of a distributed land surface water and energy balance model, *J. Geophys. Res.*, *106*, 33,421–33,433, doi:10.1029/2000JD900803.
- Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter (2002), NCEP-DOE AMIP-II Reanalysis (R-2), *Bull. Am. Meteorol. Soc.*, *83*, 1631–1643, doi:10.1175/BAMS-83-11-1631(2002)083<1631:NAR>2.3.CO;2.
- Koster, R. D., and M. J. Suarez (2001), Soil moisture memory in climate models, *J. Hydrometeorol.*, *2*, 558–570, doi:10.1175/1525-7541(2001)002<0558:SMMICM>2.0.CO;2.
- Koster, R. D., M. J. Suarez, and M. Heiser (2000a), Variance and predictability of precipitation at seasonal-to-interannual timescales, *J. Hydrometeorol.*, *1*, 26–46, doi:10.1175/1525-7541(2000)001<0026:VAPOPA>2.0.CO;2.
- Koster, R. D., M. J. Suarez, A. Ducharme, M. Stieglitz, and P. Kumar (2000b), A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure, *J. Geophys. Res.*, *105*(D20), 24,809–24,822, doi:10.1029/2000JD900327.
- Lettenmaier, D. P., and J. S. Famiglietti (2006), Water from on high, *Nature*, *444*, 562–563, doi:10.1038/444562a.
- 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.
- Liu, Y., and R. Avissar (1999a), A study of persistence in the land-atmosphere system using a general circulation model and observations, *J. Clim.*, *12*, 2139–2153, doi:10.1175/1520-0442(1999)012<2139:ASOPIT>2.0.CO;2.
- Liu, Y., and R. Avissar (1999b), A study of persistence in the land-atmosphere system with a fourth order analytical model, *J. Clim.*, *12*, 2154–2168, doi:10.1175/1520-0442(1999)012<2154:ASOPIT>2.0.CO;2.
- Lo, M. H., P. J.-F. Yeh, and J. S. Famiglietti (2008), Using baseflow to constrain water table depth simulations in the NCAR Community Land Model (CLM), *Adv. Water Resour.*, *31*(12), 1552–1564, doi:10.1016/j.advwatres.2008.06.007.
- Lohmann, D., et al. (1998), The Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) phase 2(c) Red-Arkansas River basin experiment: 3. Spatial and temporal analysis of water fluxes, *Global Planet. Change*, *19*(1–4), 161–179, doi:10.1016/S0921-8181(98)00046-0.
- Mau, D. P., and T. C. Winter (1997), Estimating ground-water recharge and baseflow from streamflow hydrographs for a small watershed in New Hampshire, *Ground Water*, *35*, 291–304, doi:10.1111/j.1745-6584.1997.tb00086.x.
- Maxwell, R. M., and N. L. Miller (2005), Development of a coupled land surface and groundwater model, *J. Hydrometeorol.*, *6*, 233–247, doi:10.1175/JHM422.1.
- Miguez-Macho, G., Y. Fan, C. P. Weaver, R. Walko, and A. Robock (2007), Incorporating water table dynamics in climate modeling: 2. Formulation, validation, and soil moisture simulation, *J. Geophys. Res.*, *112*, D13108, doi:10.1029/2006JD008112.
- Mugo, J. M., and T. C. Sharma (1999), Application of a conceptual method for separating runoff components in daily hydrographs in Kimikia Forest catchments, Kenya, *Hydrol. Processes*, *13*, 2931–2939, doi:10.1002/(SICI)1099-1085(19991215)13:17<2931::AID-HYP838>3.0.CO;2-N.
- Nathan, R. J., and T. A. McMahon (1990), Evaluation of automated techniques for baseflow and recession analysis, *Water Resour. Res.*, *26*, 1465–1473, doi:10.1029/WR026i007p01465.
- Niu, G.-Y., and Z.-L. Yang (2006), Accessing a land surface model's improvements with GRACE estimates, *Geophys. Res. Lett.*, *33*, L07401, doi:10.1029/2005GL025555.
- 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.
- Ramillien, G., A. Cazenave, and O. Brunau (2004), Global time variations of hydrological signals from GRACE satellite gravimetry, *Geophys. J. Int.*, *158*, 813–826, doi:10.1111/j.1365-246X.2004.02328.x.
- Ramillien, G., J. S. Famiglietti, and J. Wahr (2008), Detection of continental hydrology and glaciology signals from GRACE: A review, *Surv. Geophys.*, *29*(4–5), 361–374, doi:10.1007/s10712-008-9048-9.
- Rodell, M., and J. S. Famiglietti (1999), Detectability of variations in continental water storage from satellite observations of the time-dependent gravity field, *Water Resour. Res.*, *35*(9), 2705–2723, doi:10.1029/1999WR900141.
- Rodell, M., and J. S. Famiglietti (2001), An analysis of terrestrial water storage variations in Illinois with implications for Gravity Recovery and Climate Experiment (GRACE), *Water Resour. Res.*, *37*(5), 1327–1339, doi:10.1029/2000WR900306.
- Rodell, M., and J. S. Famiglietti (2002), The potential for satellite-based monitoring of groundwater storage changes using GRACE: The High Plains aquifer, central U. S., *J. Hydrol.*, *263*, 245–256, doi:10.1016/S0022-1694(02)00060-4.
- Rodell, M., et al. (2004a), The Global Land Data Assimilation System, *Bull. Am. Meteorol. Soc.*, *85*, 381–394, doi:10.1175/BAMS-85-3-381.

- Rodell, M., J. S. Famiglietti, J. Chen, S. I. Seneviratne, P. Viterbo, S. Holl, and C. R. Wilson (2004b), Basin-scale estimates of evapotranspiration using GRACE and other observations, *Geophys. Res. Lett.*, *31*, L20504, doi:10.1029/2004GL020873.
- Rodell, M., J. Chen, H. Kato, J. S. Famiglietti, J. Nigro, and C. Wilson (2007), Estimating ground water storage changes in the Mississippi River basin (USA) using GRACE, *Hydrogeol. J.*, *15*, 159–166, doi:10.1007/s10040-006-0103-7.
- Rowlands, D. D., S. B. Luthcke, S. M. Klosko, F. G. R. Lemoine, D. S. Chinn, J. J. McCarthy, C. M. Cox, and O. B. Anderson (2005), Resolving mass flux at high spatial and temporal resolution using GRACE intersatellite measurements, *Geophys. Res. Lett.*, *32*, L04310, doi:10.1029/2004GL021908.
- Seo, K.-W., C. R. Wilson, J. S. Famiglietti, J. L. Chen, and M. Rodell (2006), Terrestrial water mass load changes from Gravity Recovery and Climate Experiment (GRACE), *Water Resour. Res.*, *42*, W05417, doi:10.1029/2005WR004255.
- Shao, Y., and A. Henderson-Sellers (1996), Validation of soil moisture simulation in land surface parameterization schemes with HAPEX data, *Global Planet. Change*, *13*, 11–46, doi:10.1016/0921-8181(95)00038-0.
- Sorooshian, S., and V. K. Gupta (1983), Automatic calibration of conceptual rainfall-runoff models: The question of parameter observability and uniqueness, *Water Resour. Res.*, *19*(1), 260–268, doi:10.1029/WR019i001p0260.
- Strassberg, G., B. R. Scanlon, and M. Rodell (2007), Comparison of seasonal terrestrial water storage variations from GRACE with groundwater-level measurements from the High Plains aquifer (USA), *Geophys. Res. Lett.*, *34*, L14402, doi:10.1029/2007GL030139.
- Swenson, S. C., and P. C. D. Milly (2006), Climate model biases in seasonality of continental water storage revealed by satellite gravimetry, *Water Resour. Res.*, *42*, W03201, doi:10.1029/2005WR004628.
- Swenson, S. C., P. J.-F. Yeh, J. Wahr, and J. S. Famiglietti (2006), A comparison of terrestrial water storage variations from GRACE with in situ measurements from Illinois, *Geophys. Res. Lett.*, *33*, L16401, doi:10.1029/2006GL026962.
- Swenson, S., J. S. Famiglietti, J. Basara, and J. Wahr (2008), Estimating profile soil moisture and groundwater variations using GRACE and Oklahoma Mesonet soil moisture data, *Water Resour. Res.*, *44*, W01413, doi:10.1029/2007WR006057.
- Syed, T. H., J. S. Famiglietti, V. Zlotnicki, and M. Rodell (2007), Contemporary estimates of Pan-Arctic freshwater discharge from GRACE and reanalysis, *Geophys. Res. Lett.*, *34*, L19404, doi:10.1029/2007GL031254.
- Syed, T. H., J. S. Famiglietti, M. Rodell, J. Chen, and C. R. Wilson (2008), Analysis of terrestrial water storage changes from GRACE and GLDAS, *Water Resour. Res.*, *44*, W02433, doi:10.1029/2006WR005779.
- Syed, T. H., J. S. Famiglietti, and D. Chambers (2009), GRACE-based estimates of terrestrial freshwater discharge from basin to continental scales, *J. Hydrometeorol.*, *10*(1), 22–40, doi:10.1175/2008JHM993.1.
- Szilagy, J. (2004), Heuristic continuous base flow separation, *J. Hydrol. Eng.*, *9*(4), 311–318, doi:10.1061/(ASCE)1084-0699(2004)9:4(311).
- Tapley, B., S. Bettapur, M. Watkins, and C. Reigber (2004), The Gravity Recovery and Climate Experiment: Mission overview and early results, *Geophys. Res. Lett.*, *31*, L09607, doi:10.1029/2004GL019920.
- Wahr, J., S. Swenson, V. Zlotnicki, and I. Velicogna (2004), Time-variable gravity from GRACE: First results, *Geophys. Res. Lett.*, *31*, L11501, doi:10.1029/2004GL019779.
- Werth, S., A. Güntner, S. Petrovic, and R. Schmidt (2009), Integration of GRACE mass variations into a global hydrological model, *Earth Planet. Sci. Lett.*, *227*(1–2), 166–173, doi:10.1016/j.epsl.2008.10.021.
- Wu, W., and R. E. Dickinson (2004), Time scales of layered soil moisture memory in the context of land-atmosphere interaction, *J. Clim.*, *17*, 2752–2764, doi:10.1175/1520-0442(2004)017<2752:TSOLSM>2.0.CO;2.
- Wu, W., M. A. Geller, and R. E. Dickinson (2002), The response of soil moisture to long-term variability of precipitation, *J. Hydrometeorol.*, *3*, 604–613, doi:10.1175/1525-7541(2002)003<0604:TROSMT>2.0.CO;2.
- Xie, Z., F. Yuan, Q. Duan, J. Zheng, M. Liang, and F. Chen (2007), Regional parameter estimation of the VIC land surface model: Methodology and application to river basins in China, *J. Hydrometeorol.*, *8*(3), 447–468, doi:10.1175/JHM568.1.
- Yang, Z.-L., R. E. Dickinson, A. Robock, and K. Y. Vinnikov (1997), Validation of the snow sub-model of the Biosphere-Atmosphere Transfer Scheme with Russian snow cover and meteorological observational data, *J. Clim.*, *10*, 353–373, doi:10.1175/1520-0442(1997)010<0353:VOTSSO>2.0.CO;2.
- Yapo, P. O., H. V. Gupta, and S. Sorooshian (1998), Multi-objective global optimization for hydrologic models, *J. Hydrol.*, *204*, 83–97, doi:10.1016/S0022-1694(97)00107-8.
- Yeh, P. J.-F., and E. A. B. Eltahir (2005a), Representation of water table dynamics in a land surface scheme: part I. Model development, *J. Clim.*, *18*, 1861–1880, doi:10.1175/JCLI3330.1.
- Yeh, P. J.-F., and E. A. B. Eltahir (2005b), Representation of water table dynamics in a land surface scheme: part II. Subgrid variability, *J. Clim.*, *18*, 1881–1901, doi:10.1175/JCLI3331.1.
- Yeh, P. J.-F., and J. S. Famiglietti (2009), Regional groundwater evapotranspiration in Illinois, *J. Hydrometeorol.*, *10*(2), 464–478, doi:10.1175/2008JHM1018.1.
- 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, doi:10.1029/98JD01721.
- Yeh, P. J.-F., S. C. Swenson, J. S. Famiglietti, and M. Rodell (2006), Remote sensing of groundwater storage changes in Illinois using the Gravity Recovery and Climate Experiment (GRACE), *Water Resour. Res.*, *42*, W12203, doi:10.1029/2006WR005374.
- Zaitchik, B. F., M. Rodell, and R. Reichle (2008), Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi River basin, *J. Hydrolometeorol.*, *9*, 535–548, doi:10.1175/2007JHM951.1.
- Zeng, X., and M. Decker (2009), Improving the numerical solution of soil moisture-based Richards equation for land models with a deep or shallow water table, *J. Hydrometeorol.*, *8*(3), 447–468.

J. S. Famiglietti and M.-H. Lo, Department of Earth System Science, University of California, Irvine, CA 92697, USA. (jfamigli@uci.edu; mlo@uci.edu)

T. H. Syed, Department of Applied Geology, Indian School of Mines University, Dhanbad, India. (tsyed.ismu@gmail.com)

P. J.-F. Yeh, Institute of Industrial Science, University of Tokyo, Tokyo 153-8505, Japan. (patyeh@rainbow.iis.u-tokyo.ac.jp)