Assimilation of GRACE Terrestrial Water Storage Data into a Land Surface Model: Results for the Mississippi River Basin

BENJAMIN F. ZAITCHIK

Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, and Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland

MATTHEW RODELL

Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland

ROLF H. REICHLE

Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore County, Baltimore, and Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

(Manuscript received 13 August 2007, in final form 15 October 2007)

ABSTRACT

Assimilation of data from the Gravity Recovery and Climate Experiment (GRACE) system of satellites yielded improved simulation of water storage and fluxes in the Mississippi River basin, as evaluated against independent measurements. The authors assimilated GRACE-derived monthly terrestrial water storage (TWS) anomalies for each of the four major subbasins of the Mississippi into the Catchment Land Surface Model (CLSM) using an ensemble Kalman smoother from January 2003 to May 2006. Compared with the open-loop CLSM simulation, assimilation estimates of groundwater variability exhibited enhanced skill with respect to measured groundwater in all four subbasins. Assimilation also significantly increased the correlation between simulated TWS and gauged river flow for all four subbasins and for the Mississippi River itself. In addition, model performance was evaluated for eight smaller watersheds within the Mississippi basin, all of which are smaller than the scale of GRACE observations. In seven of eight cases, GRACE assimilation led to increased correlation between TWS estimates and gauged river flow, indicating that data assimilation has considerable potential to downscale GRACE data for hydrological applications.

1. Introduction

Since its launch in March 2002, the Gravity Recovery and Climate Experiment (GRACE) satellite system has provided unprecedented measurements of column-integrated terrestrial water storage (TWS) for the entire globe. These measurements have been applied in novel investigations of river discharge (Syed et al. 2005), regional evapotranspiration (Rodell et al. 2004a; Swenson and Wahr 2006a), climate and teleconnections (Andersen et al. 2005; Crowley et al. 2006), and the changing mass of major glaciers and ice sheets (Luthcke et al. 2006; Tamisiea et al. 2005; Velicogna and Wahr 2006), yielding important insight on regional to global-scale water cycle variability.

To realize the full potential of GRACE for hydrology, the derived regional-scale, column-integrated, monthly water storage anomalies must be disaggregated horizontally, vertically, and in time. Observational estimates of TWS from GRACE are routinely generated on a monthly basis, though techniques have been developed for producing 10-day anomalies (Rowlands et al. 2005). GRACE's horizontal resolution is limited to about 150 000 km^2 (Rowlands et al. 2005; Yeh et al. 2006). Vertically, the GRACE TWS observation is a single number that integrates changes in groundwater, soil moisture, vegetation, surface water, snow, and ice. Skillful disaggregation of GRACE terrestrial water storage anomalies into changes in these individual components would greatly improve their value for hydrological research and applications. For

Corresponding author address: Benjamin F. Zaitchik, Hydrological Sciences Branch, Code 614.3, NASA Goddard Space Flight Center, Greenbelt, MD 20771.
E-mail: ben.zaitchik@nasa.gov

DOI: 10.1175/2007JHM951.1

example, initialization of seasonal forecasts requires accurate estimates of model variables that are closely related to TWS and that contribute to the "memory" of the climate system at monthly scales. While a number of operational satellite platforms provide data on land surface conditions, including skin temperature, surface soil moisture, and vegetation, GRACE is the only remote sensor currently capable of detecting changes in TWS at any depth, under any conditions.

One approach to vertical disaggregation of GRACE data is to use auxiliary information to isolate individual components. Rodell et al. (2007) computed groundwater storage variations averaged over the Mississippi River basin and its four major subbasins by using soil moisture and snow water equivalent output from the Global Land Data Assimilation System (GLDAS; Rodell et al. 2004b) to estimate and remove those components from GRACE TWS, assuming vegetation and surface water contributions to be negligible. The results compared favorably with piezometer-based groundwater storage estimates for the full Mississippi River basin and the two larger subbasins. Similarly, Yeh et al. (2006) used ground-based observations of soil moisture to isolate groundwater storage variations from the GRACE signal, with reasonable success.

A more sophisticated disaggregation method is to merge GRACE-derived TWS with that simulated by a land surface model (LSM) via data assimilation. This approach has a number of advantages. First, the GRACE observations themselves, though coarse, yield reasonably reliable estimates of TWS anomalies (Swenson et al. 2006). Assimilating these data into an LSM, therefore, has the potential to improve the accuracy of TWS in LSM simulations (Ellett et al. 2006), much as assimilation of remotely sensed snow cover (Clark et al. 2006; Rodell and Houser 2004), snow water equivalent (Slater and Clark 2006), soil moisture (Reichle et al. 2007), and skin temperature (Bosilovich et al. 2007) have had a positive impact on LSM simulations. Second, our understanding of hydrological processes, as captured by the model, is used to enhance the satellite observations, providing downscaling and quality control of GRACE observations while enabling synthesis of multiple observation types in a physically consistent manner. Third, an assimilated observation of TWS influences a number of processes within an LSM in addition to water storage. Predictions of water and energy fluxes are thus informed by the GRACE observation, allowing us to quantify the influence of a bulk TWS anomaly on spatially distributed runoff, evaporation, ground heat transfer, etc. This is a primary motivation for data assimilation in general, though it is also a point of caution; assimilating one model state can have a destabilizing impact on other model processes (Walker and Houser 2005).

The unique characteristics of GRACE measurements pose two particular challenges for assimilation into an LSM. First, the assimilation algorithm itself must map very few, coarse-resolution GRACE observations onto the many LSM elements required to simulate land surface processes at a useful resolution. This is an uncertain process at best, and it demands an assimilation algorithm that skillfully distributes the information from a single coarse-scale observation onto the numerous finer-scale model elements to which it is applied. Second, it is necessary to assimilate the GRACE observation into an analogous field in the LSM. This is a problem of disaggregation in its own right, as TWS is divided between several storage components in any advanced LSM. Furthermore, the lack of a groundwater reservoir is a deficiency that makes many current LSMs inappropriate for this task (Niu et al. 2007).

In this study we adapted to these challenges by assimilating GRACE TWS anomalies into the Catchment Land Surface Model (CLSM) using an ensemble Kalman smoother (EnKS). The complete assimilation system is presented as follows. In section 2 GRACE and other data sources are described in greater detail. Section 3 reviews relevant features of the CLSM and presents the EnKS. Results are given in section 4 and conclusions in section 5.

2. Data

The study required near-surface meteorological data to force CLSM, GRACE-derived TWS anomalies for data assimilation, and groundwater and runoff observations for evaluation. The meteorological data were drawn from the GLDAS forcing database (Rodell et al. 2004b) at $2.0° \times 2.5°$ resolution. The other datasets are described next.

a. *GRACE data*

GRACE data used in this study were processed at the University of Texas Center for Space Research (CSR), at the GeoForschungsZentrum Potsdam (GFZ), and at the National Aeronautics and Space Administration (NASA) Jet Propulsion Laboratory (JPL). Each center uses its own processing algorithm, but the essential characteristics of the calculation are the same. Global representations of Earth's gravity field are produced on a near-monthly basis as sets of spherical harmonic coefficients up to degree and order 120, based on highly precise K-band microwave measurements of the distance between two identical satellites orbiting earth

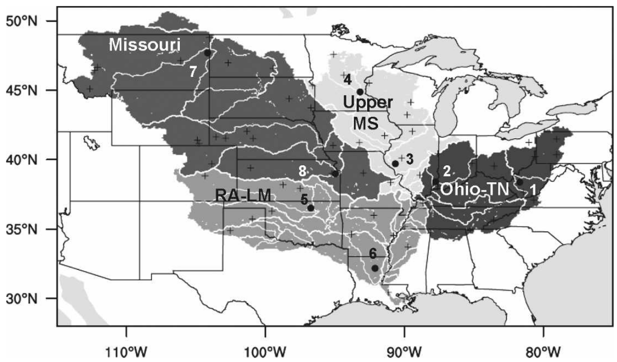

FIG. 1. The four major subbasins of the Mississippi River: the Missouri, Upper Mississippi, Ohio–Tennessee, and the combined Red–Arkansas/Lower Mississippi (RA–LM). Thin white lines indicate the borders of smaller watersheds within each subbasin, including the 1) Kanahwa, 2) Wabash, 3) Illinois, 4) Minnesota, 5) Arkansas, 6) Ouachita, 7) Yellowstone, and 8) Kansas Rivers, which were used in model evaluations (Table 3). Black dots indicate location of river gauges for the smaller watersheds. Crosses indicate locations of the 58 piezometers used to calculate groundwater variability.

in tandem (Tapley et al. 2004). The gravitational effects of changes in atmospheric surface pressure and ocean bottom pressure are removed using numerical model analyses, such that the remaining variability can be attributed primarily to the redistribution of terrestrial water storage. The observed gravity signal degrades at higher degrees and orders, so there is a trade-off between spatial resolution and signal accuracy. The GRACE data used here (Chambers 2006) were smoothed using a Gaussian averaging kernel with 400-km averaging radius and "destriped" following Swenson and Wahr (2006b). Water storage changes were extracted for each of the four major subbasins of the Mississippi River: the Ohio–Tennessee, Upper Mississippi, Missouri, and combined Red–Arkansas/Lower Mississippi (Fig. 1). This approach is consistent with earlier hydrological applications of GRACE (Chen et al. 2005; Syed et al. 2005).

Monthly TWS anomalies were obtained for February 2003–April 2006 from CSR (release 1), for February 2003–May 2006 from GFZ (release 3), and from January 2003–November 2005 from JPL (release 2). Gravitational anomalies were not reported by any of the three processing teams for June 2003, January 2004, or July–October 2004. The data gap in the summer of 2004 was due to a resonance that caused the GRACE satellites to enter near-repeat orbit for several months (Wagner et al. 2006). As all three centers produced reasonably similar TWS anomalies for the period of overlap (Fig. 2), we used an average of available estimates to provide monthly assimilation inputs from January 2003 to May 2006.

Optimal data assimilation requires error estimates for both the model and the observation. This is not straightforward given the multiple sources of uncertainty in GRACE. Following Wahr et al. (2006) we use 20 mm as a conservative estimate of RMS error for midlatitude GRACE TWS measurements. The subbasin average anomalies used in this study span several averaging radii, suggesting that actual error for the entire subbasin may be smaller. For this reason a second assimilation integration was performed for which the error of GRACE subbasin averages was assumed to be 10 mm. Results for this simulation did not differ significantly from the simulation with the conservative error estimate, so they are not presented in this paper.

b. *Groundwater observations*

Time series of groundwater storage anomalies were generated based on unconfined and semiconfined water-level records from 58 piezometers distributed across the Mississippi River basin (Fig. 1). Sources included the U.S. Geological Survey (USGS) Ground-Water Climate Response Network (CRN), the USGS WatStore system, the Illinois State Water Survey, and published reports. The groundwater and GRACE TWS datasets were derived completely independently. Specific yield estimates, used to convert well-water levels to equivalent heights of water, were individually selected based on published information prior to any comparisons with GRACE, and no subsequent tuning was performed on either dataset. Additional details are provided by Rodell et al. (2007).

c. *River discharge*

River discharge data used in this study are daily mean gauged values collected by automated recorders. For the Mississippi River itself, these data were obtained from the U.S. Army Corps of Engineers New Orleans District (http://www.mvn.usace.army.mil/eng/edhd/watercon.htm). For all tributary rivers (including the Upper Mississippi), data were obtained from the USGS National Water Information Service (http://nwis.waterdata.usgs.gov/nwis/sw).

3. Methods

a. *The Catchment Land Surface Model*

The CLSM (Koster et al. 2000) was developed in response to a perceived shortcoming in conventional land surface models: the soil-layer-based vertical discretization of conventional LSMs is not well suited to surface hydrologic processes. In effect, layer-based

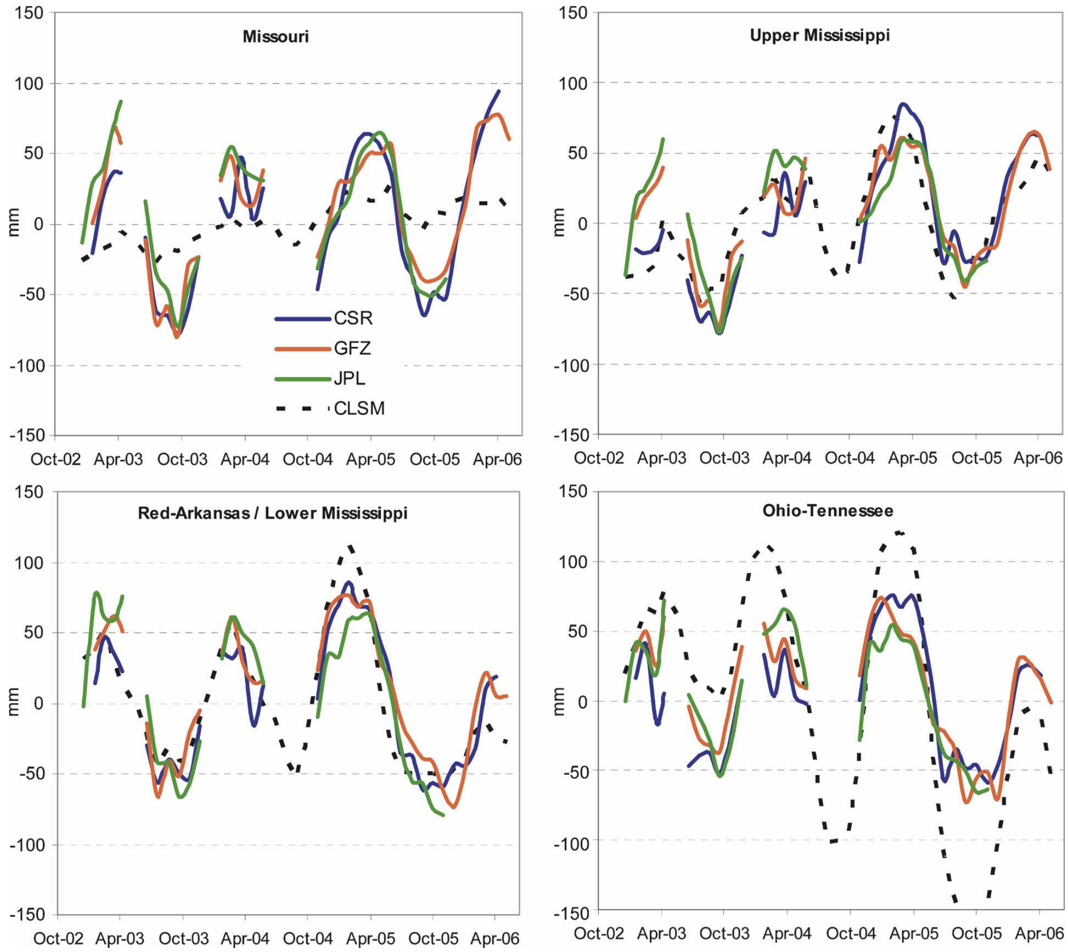

FIG. 2. Monthly TWS anomalies based on CSR, GFZ, and JPL GRACE observational estimates and from an open-loop simulation with the CLSM.

LSMs assume uniform topographic and hydrologic characteristics at the grid scale, typically spanning tens of kilometers. This impairs a model's ability to simulate runoff, which in turn leads to unrealistic fields of soil moisture and evapotranspiration (Koster and Milly 1997). CLSM instead divides the land surface into topographically defined catchments with an average area of approximately 4000 km^2 and models hydrologic processes based on each catchment's topographical statistics. Subcatchment heterogeneity of soil moisture is modeled by dividing the catchment into dynamic fractions of saturated, unsaturated, and wilting areas, each governed by equations appropriate for its soil moisture status.

The primary prognostic variable in the CLSM is the catchment deficit, defined as the average depth of water that would need to be added to bring the catchment to saturation (Fig. 3). The equilibrium vertical distribution of soil moisture is then diagnosed on the basis of the catchment deficit and soil parameters. This distribution includes an implicit water table, located at the depth of equilibrium saturation. In addition to the catchment deficit, CLSM prognostics include reservoirs of root zone excess moisture and surface excess moisture that permit a rough representation of nonequilibrium vertical conditions such as infiltration fronts. The surface excess moisture reservoir is small relative to both root zone excess and catchment deficit. Snow is represented in a state-of-the-art three-layer snow physics scheme (Stieglitz et al. 2001).

In this study CLSM simulations were performed for the Mississippi River basin. Catchment information was defined on the basis of a 30 arc-s digital elevation model from the USGS (Verdin and Verdin 1999). For computational reasons, the fundamental modeling element in CLSM is the "tile," defined by the intersection of a catchment with the overlying atmospheric grid. The Mississippi River basin comprises 783 defined catch-

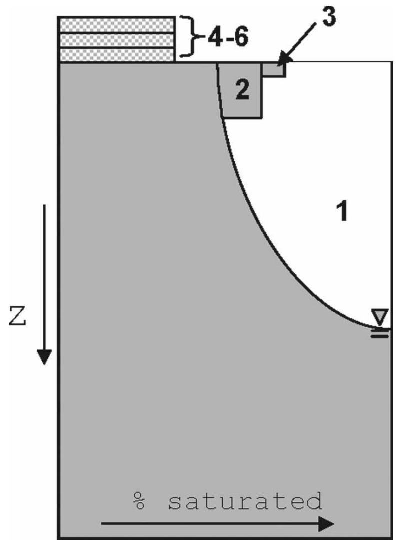

FIG. 3. Prognostic hydrologic variables in the CLSM: 1) catchment deficit, 2) root zone excess, 3) surface excess, 4)–6) three snow layers.

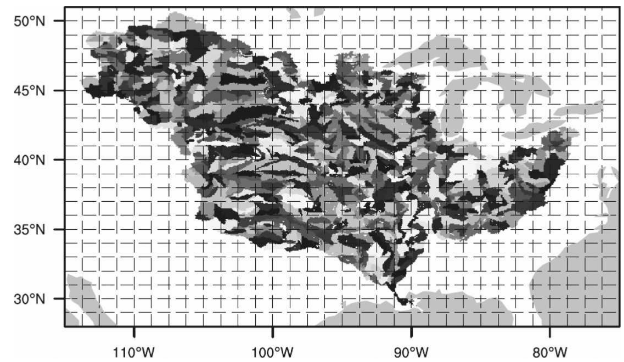

FIG. 4. CLSM modeling domain for the Mississippi River basin. Shading indicates topographically defined catchment units, and the dashed grid indicates atmospheric forcing. The CLSM model unit is the tile, defined by any unique combination of a catchment with an atmospheric grid cell.

ments, resulting in 1950 tiles under a $1.0° \times 1.25°$ atmospheric grid (Fig. 4). The model was spun up for 10 yr under 2002 forcing conditions and integrated from 1 January 2003 through 1 June 2006. The results of two integrations are described in this paper: an open-loop (OL) simulation without assimilation and a GRACE data assimilation (GRACE DA) integration.

The defining characteristics of CLSM make it particularly appropriate for the assimilation of GRACE-derived TWS anomalies. As described in section 2a, GRACE TWS anomalies can be extracted for a watershed of arbitrary shape. By pairing watershed-defined GRACE estimates with a watershed-defined CLSM domain, it is possible to perform area-accurate assimilation for hydrologically defined basins. More importantly, the presence of a variable water table is essential, since it means that the model accounts for most of the groundwater variability measured by GRACE. CLSM's lack of traditional hydrologic layers in the subsurface is convenient: we apply increments (based on TWS) directly to the column-integrated prognostic variable (the catchment deficit) and the primary nonequilibrium prognostic (the root zone excess moisture), without need for arbitrary vertical disaggregation. Finally, the subdivision of each catchment into saturated, unsaturated, and wilting fractions provides a physically based mechanism for weighting the hydrologic effects of an assimilated GRACE observation across a morphologically diverse modeling unit. This does not solve the problem of applying a single GRACE observation of a given subbasin to numerous model catchments within the subbasin, but it does furnish a rationale for spatially distributing the effects of assimilation at the subcatchment scale.

b. Data assimilation

In a data assimilation system, the modeled fields are corrected toward observational estimates, with the degree of correction determined by the levels of error associated with each. Data assimilation algorithms can be divided into filters, including the widely used ensemble Kalman filter (EnKF), and smoothers. Filters assimilate observations as they become available, update only the most recent model state estimates, and are well suited for forecast applications that require estimates of present conditions only. By contrast, smoothers use information from a series of observations to update model fields over a window of time and are often preferable in reanalyses, where there is an interest in accurate representation of the evolving system (Dunne and Entekhabi 2005; Evensen and van Leeuwen 2000). Generally, smoothing estimates are at least as good as filtering estimates. Smoothing algorithms, however, carry a considerable computational cost and their efficient implementation is not trivial (Dunne and Entekhabi 2005; Ravela and McLaughlin 2007).

Because GRACE TWS observations are time averaged and our focus is on retrospective analysis, we use an ensemble smoother for our application. Our approach is similar (but not identical) to previous ensemble smoothing studies (e.g., Dunne and Entekhabi 2005; van Leeuwen 2001). A few practical modifications were necessary on account of the time-averaged nature of GRACE observations: this is not a case of distributing instantaneous observations over a number of model time steps, but of temporally disaggregating a month-long observation onto the finer temporal scale of the LSM. The next section describes the ensemble smoother, and section 3d gives details on its implementation.

c. The ensemble Kalman smoother

The ensemble smoother used here is similar to the EnKF (Reichle et al. 2002). In the ensemble approach, conditional probability densities for predicted states are approximated by a finite number of model trajectories—the ensemble—with a covariance that reflects uncertainties in the model physics, parameters, and forcing data. Assimilation increments are calculated based on the relative uncertainty in the model and the observations, described by the (sample) error covariance matrices.

In our application we subdivide the experiment period into data assimilation (or smoothing) windows T that coincide with calendar months. Each GRACE observation is assimilated only in the month in which the observation was collected. Stated generically, the update for each ensemble member and for smoothing window T can be written as

$$\mathbf{X}^i_{T+} = \mathbf{X}^i_{T-} + \mathbf{K}_T[\mathbf{Y}^i_T - \mathbf{M}_T(\mathbf{X}^i_{T-})]. \tag{1}$$

Here, $\mathbf{X}^i_{T-}$ and $\mathbf{X}^i_{T+}$ denote the ith ensemble member of the state vector before and after the update, respectively. The state vector is a collection of model prognostic variables at one or more times in the assimilation window T. The vector $\mathbf{Y}^i_T$ contains the observations (suitably perturbed) and $\mathbf{M}_T$ is the measurement operator that maps the model fields into observation space. Vectors $\mathbf{X}^i_T$ and $\mathbf{Y}^i_T$ need not be of the same length: $\mathbf{X}^i_T$ has length n (for n model variables), while $\mathbf{Y}^i_T$ has length m (for m observations).

The time-dependent Kalman gain matrix determines the relative weights of the model versus the observations during the update, and is defined on the basis of their respective covariance matrices,

$$\mathbf{K}_T = [\mathbf{C}_{\mathbf{XM}}(\mathbf{C}_{\mathbf{M}} + \mathbf{C}_\upsilon)^{-1}]_T. \tag{2}$$

Uncertainties in the observations are described with the covariance matrix $\mathbf{C}_\upsilon$. The matrix $\mathbf{C}_{\mathbf{M}}$ is the error covariance of the corresponding model predictions ($\mathbf{M}_T[\mathbf{X}_T]$), and $\mathbf{C}_{\mathbf{XM}}$ is the error cross covariance between the state $\mathbf{X}_t$ and the model predictions $\mathbf{M}_T[\mathbf{X}_T]$. The cross covariance $\mathbf{C}_{\mathbf{XM}}$ is particularly important because it provides the basis for the distribution of observational information from the coarse subbasin scale to the finer-scale tile space. Since $\mathbf{C}_{\mathbf{XM}}$ is diagnosed from the ensemble, the perturbations that are added to the forcings and state variables of each ensemble member must include realistic horizontal correlations (see below).

d. Implementation of the GRACE EnKS

Here, we apply the EnKF update separately for each subbasin, such that the model state matrix $\mathbf{X}_T$ has dimensions $2N_j \times 20$, reflecting an ensemble of 20 members for the catchment deficit and the root zone excess moisture at the beginning of the month in each of the N_j tiles in subbasin j ($j = 1, \ldots, 4$). (See below for tiles with snow cover.) Meanwhile, $\mathbf{Y}_T$ is a scalar, as only one GRACE observation is available for each major subbasin in each assimilation window. GRACE TWS anomalies were converted to absolute TWS values by adding the corresponding time-mean TWS from an open-loop CLSM simulation for the assimilation period. GRACE observations were not otherwise scaled for assimilation. The (scalar) observation error variance $\mathbf{C}_\upsilon = (20\text{ mm})^2$ includes errors involved in producing TWS anomaly estimates from GRACE orbital data. Each CLSM integration included 20 ensemble members; additional integrations with 12 and 100 ensemble members each were performed for comparison, and they yielded similar results.

In the past, both the EnKF and the EnKS have been applied successfully to the assimilation of land surface observations with roughly the same temporal and spatial scales as the land surface model. GRACE data, however, present additional challenges due to their coarse temporal and spatial resolutions. Temporally, each GRACE observation can be described as a monthly estimate of relative TWS that is informed by approximately three satellite overpasses, depending on region size, shape, and orientation (Tapley et al. 2004). We therefore calculate the model-predicted TWS—denoted $\mathbf{M}_T[\mathbf{X}_T]$ above—as follows (and schematically depicted in Fig. 5). First, the CLSM ensemble is propagated forward one month without any data assimilation. During this first pass, model estimates of subbasin average TWS are stored in memory at three specific times—on the 5th, 15th, and 25th days of the month—roughly mimicking the GRACE observation frequency (Fig. 5[1]). The model-predicted TWS, $\mathbf{M}_T[\mathbf{X}_T]$, is then computed as the average TWS over these three times (Fig. 5[2]). Next, the assimilation increments are calculated using Eq. (1) for a state $\mathbf{X}_T$ that consists of the catchment deficit and root zone excess values on the 1st of the month, one value each for every catchment within the subbasin and every ensemble member (Fig. 5[3]). Finally, the ensemble is reinitialized at the beginning of the month and, on the second iteration, the assimilation increments (Fig. 5[3]) are applied evenly

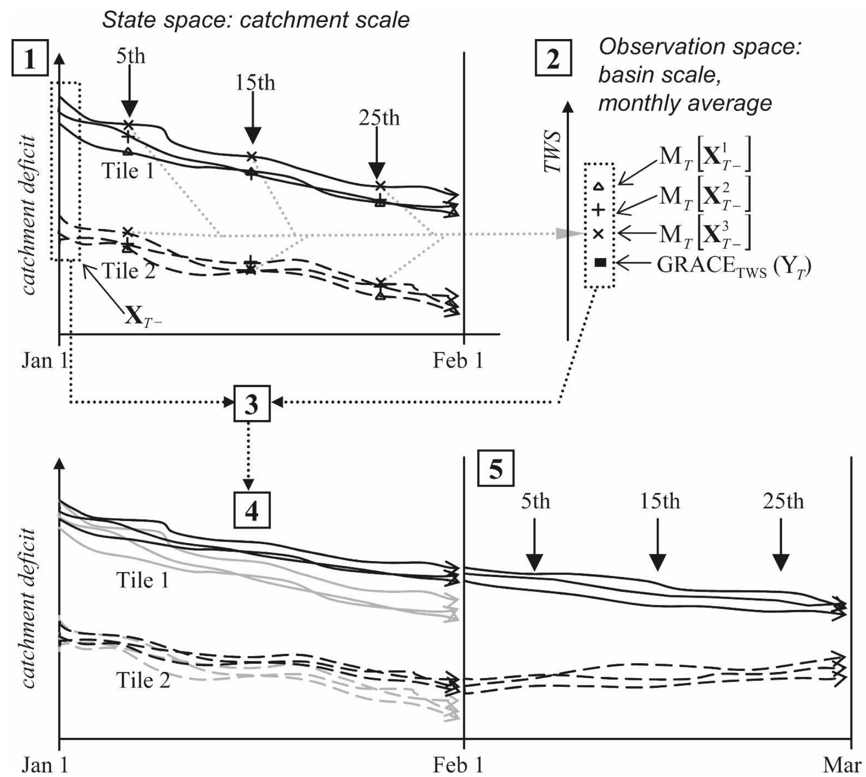

FIG. 5. Ensemble smoother. Consider one subbasin with two snow-free CLSM tiles and three ensemble members. For simplicity, root zone excess moisture is not included in this schematic. [1] One-month forecast ensemble integration without assimilation. Store catchment deficit for the 5th, 15th, and 25th of the month. [2] Calculate model prediction of GRACE observation—$\mathbf{M}_T[\mathbf{X}^i_{T-}]$—by converting stored catchment deficit values into basin-scale, time-average TWS. [3] Use Eq. (1) to compute analysis increments for catchment deficits on the 1st of the month (state vector $\mathbf{X}^i_T$). [4] Integrate CLSM again from the 1st of the month and apply analysis increments evenly distributed over all days of the month. [5] Proceed with ensemble forecast and repeat process.

over each day of the month (incremental analysis update; Fig. 5[4]). This uniform smoother is appropriate for TWS, which exhibits low-frequency variability as well as high-frequency responses to atmospheric events (Rodell and Famiglietti 2001). The updates are redistributed between water reservoirs in the CLSM over time, while the atmospheric forcing data enables the model to depict high-frequency variability. The efficacy with which CLSM redistributes water between catchment deficit, root zone excess moisture, and surface excess moisture is demonstrated by the fact that it makes very little difference whether or not root zone excess moisture is included in the assimilation update. A simulation that used only catchment deficit in the update yielded results that were very similar, though not entirely identical, to those reported here (results not shown).

Spatially, each GRACE observation spans tens to hundreds of CLSM catchments. In an EnKF-based updating system, the horizontal error correlations contained in $\mathbf{C_{XM}}$ dictate the horizontal distribution of GRACE observational information from the subbasin scale onto the many catchments that are contained within the subbasin. These correlations are determined by the spatial structure of the perturbations that we add to model forcings and prognostic variables of each ensemble member. For the simulations presented in this paper, perturbations were added to the precipitation and radiation forcing fields and to the catchment deficit and root zone excess moisture CLSM prognostics. All perturbations were generated with a horizontal correlation scale of 2°, which very roughly represents error scales in global-scale precipitation fields (Reichle and Koster 2003), with a temporal correlation of 72 h for the forcing perturbations and 24 h for perturbations to the model prognostics, and with probability distributions and covariances drawn from earlier experience with CLSM data assimilation (Reichle et al. 2007). At the same time, CLSM parameters of soil depth, drainage, and porosity inform the dynamic range of hydrologic prognostic variables in each catchment, adding additional spatial structure to $\mathbf{C_{XM}}$. Taken together, meteorologically justified perturbations to forcing and catchment-specific limits on ensemble variability in TWS provide a statistically and physically informed basis for distributing GRACE increments at scales smaller than the observation.

Finally, assimilation requires water storage updates to be distributed between snow and subsurface water. For catchments that were modeled to be snow-free at the time of an assimilation update, the entire increment was applied to the catchment deficit and root zone excess moisture. Surface excess moisture was neglected, due to its transience and small volume relative to the other reservoirs. For catchments with snow cover, positive increments (i.e., "wetting") were applied entirely to snow. Negative increments ("drying") were applied first to snow water equivalent and then, if all snow was removed, to the catchment deficit. This rule-based scheme was adopted as a simple, physically reasonable approach to updating snow fields, which were zero for most of the domain for the majority of the simulation. Alternative methods for updating snow fields in cold regions are the subject of continued research.

In summary, the GRACE EnKS has two distinguishing characteristics relative to other EnKS assimilation schemes described in the literature (e.g., Dunne and Entekhabi 2005; van Leeuwen 2001). First, we have fixed the smoother window at one month, a period comprising several GRACE overpasses that are used to inform a single, monthly estimate of TWS. One could describe this as a modified EnKF update, in that $\mathbf{Y}_T$ includes only a single value, but we believe that the use of an aggregated set of satellite overpasses to inform a temporally distributed application of assimilation increments is best described as an ensemble smoother. It would be possible to extend the smoother to a period longer than one month in order to utilize multiple GRACE estimates in each update. However, doing so would delay the availability of the TWS estimates from the assimilation system and hence diminish their value for operational applications such as drought monitoring and forecast initialization, both of which are primary motivations for the present work.

The second distinctive element of the GRACE EnKS is that it is iterative, with assimilation increments applied during the second model pass. This approach takes full advantage of the observation's potential to inform all model states and fluxes, while avoiding the spurious modification of fluxes that can arise in the context of a retroactive smoother update (Dunne and Entekhabi 2006).

4. Results

a. Terrestrial water storage

As seen in Fig. 2, the open-loop simulation with the CLSM captured the general seasonal cycle of TWS for the four major subbasins of the Mississippi. All four basins experience a wintertime peak in TWS followed by a summer trough (see also Fig. 6). The open-loop simulation differed substantially from GRACE estimates, however, with respect to the magnitude and (more subtly) the phase of the seasonal water cycle, and its interannual variability. Notably, the seasonal amplitude of TWS in the Missouri basin was only 20–40 mm in the open-loop simulation, but varied from 70 to 130 mm in GRACE retrievals. In the Ohio–Tennessee basin, the open-loop simulation returned a large and variable annual TWS cycle, while GRACE indicated that it was smaller and more stable across years.

By design, data assimilation produced TWS time series that were intermediate between the open-loop simulation and GRACE observations (Fig. 6). Only during the 4-month data gap in the summer of 2004 did the assimilation and open-loop simulation results match. There was also one systematic exception in which CLSM assimilation output did not track GRACE estimates of TWS. During observed dry periods in the Missouri basin none of the simulations dried as much as GRACE observations indicated they should, because the dry anomaly observed by GRACE exceeded the maximum possible catchment deficit of the CLSM, and hence the increments were truncated. One possible explanation of this discrepancy between model and observation is that the GRACE observation captures TWS components not included in the CLSM, such as surface water. Another interpretation is that GRACE has revealed a potential flaw in the CLSM, and that model parameters should be adjusted to allow greater drying in this region. In either case, it is clear that data assimilation provides information that is potentially valuable for refining models and observing systems.

b. Vertical disaggregation

One of the strongest motivations for assimilating GRACE data is the need to separate the contributions of the TWS components, which individually are more useful for scientific and social applications. Figure 7 plots anomalies of simulated shallow groundwater, soil moisture, and snow from the open-loop and assimilation runs, alongside the GRACE TWS and independently derived groundwater time series. GRACE data

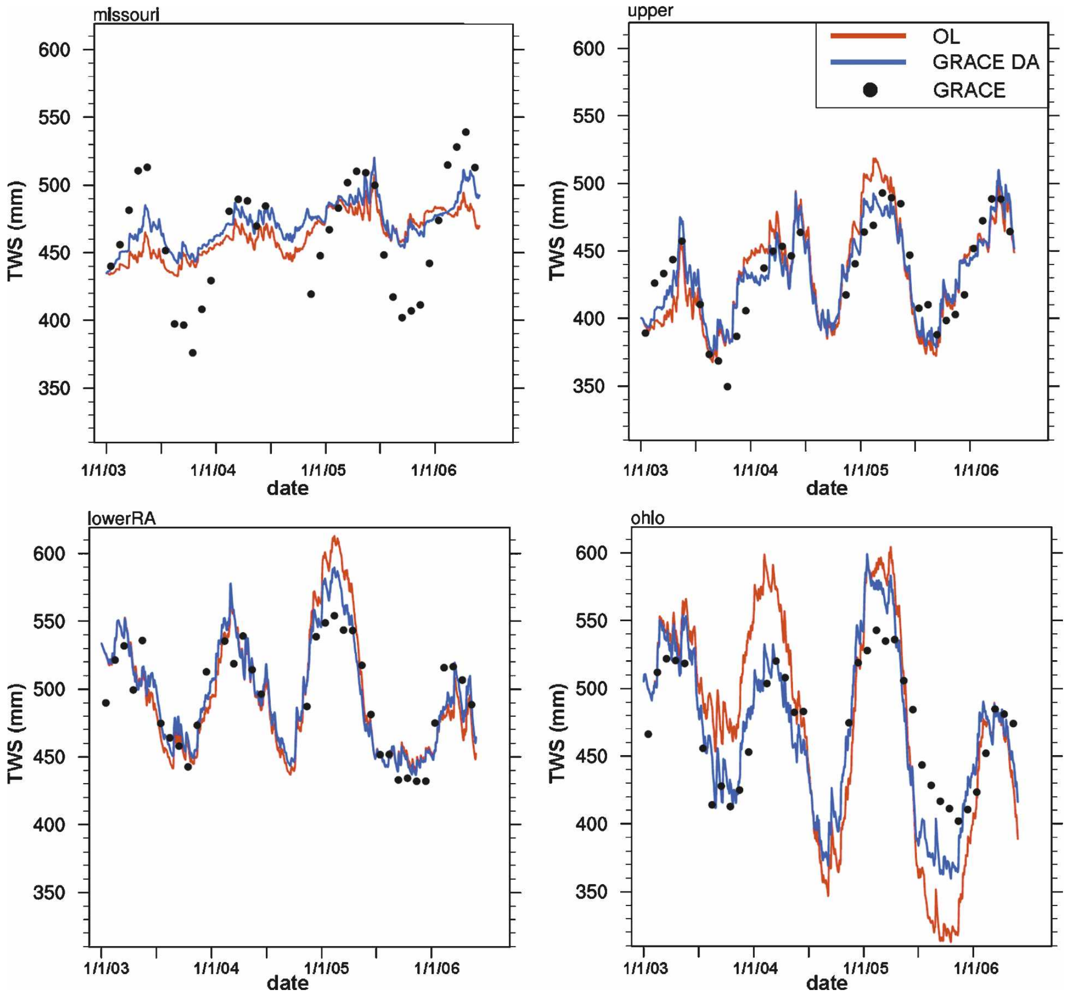

FIG. 6. Daily average TWS (mm), January 2003–May 2006, from OL and GRACE DA CLSM simulations, for (clockwise from the upper left) the Missouri, Upper Mississippi, Ohio–Tennessee, and Lower Mississippi/Red–Arkansas subbasins. Also shown are monthly GRACE TWS anomalies, shifted to the CLSM mean.

assimilation brought the seasonal cycle of TWS over the Mississippi basin into closer agreement with the GRACE anomalies (relative to the open-loop simulation), particularly in 2005. As a result, GRACE data assimilation shifted the annual peak in basin-average groundwater levels to later in the season, improving agreement with observational groundwater data. Assimilation had no detectable impact on soil moisture's phase, but the magnitude of annual soil moisture variability was reduced relative to open-loop simulations, primarily in the Ohio–Tennessee basin (not shown). Hence the assimilation of monthly GRACE data has a greater influence on groundwater, which varies slowly, than it does on soil moisture, which responds more quickly to atmospheric forcing.

Table 1 quantifies the agreement between estimated and observed groundwater for the Mississippi River basin as a whole and for its four subbasins. GRACE data assimilation significantly improved estimates of the amplitude and phase of the seasonal cycle of groundwater. In all cases, the GRACE DA integration exhibited smaller RMS errors than OL relative to measured groundwater variability, resulting in positive skill scores.

To assess the phase of the seasonal cycle, Table 1 shows time series correlations between modeled and observed groundwater. Averaged over the entire Mississippi River basin, correlations were larger for GRACE DA than for OL. This result was statistically significant at the 5% level [Fisher Z-transform test for correlation coefficients, applied to daily data (n = 1247)]. The Missouri and Ohio–Tennessee subbasins also experienced statistically significant improvements. Improvement in the combined Red–Arkansas and Lower Mississippi basins was marginally significant.

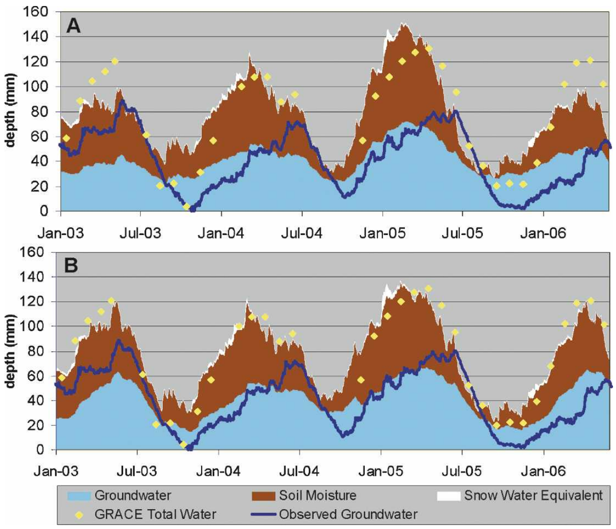

FIG. 7. Groundwater, soil moisture, and snow water equivalent for the Mississippi River basin for (a) OL and (b) GRACE DA simulations. Also shown are area-averaged daily groundwater observations and monthly GRACE-derived TWS anomalies. GRACE and modeled TWS are adjusted to a common mean, as are observed and modeled groundwater.

Assimilation did not improve CLSM's poor correlation with observed groundwater in the Upper Mississippi subbasin. This result is consistent with the fact that the observed seasonal cycle of groundwater in the Upper Mississippi is out of phase with GRACE-derived TWS as well as (not shown) soil moisture simulated by the four LSMs included in the GLDAS suite of models. There is no obvious explanation for this discrepancy, but interestingly, observed groundwater correlates well with variations in the elevation of Lake Michigan (not shown), despite the fact that none of the piezometers was closer than 100 km from the lake shore. As discussed in the following section, CLSM performs much better when evaluated against gauged discharge of the upper Mississippi River.

It should be noted that the correlations shown in Table 1, and in all subsequent tables, are calculated from area-averaged time series data that contain a seasonal cycle. The magnitude of the correlation coefficient is determined primarily by the accuracy with which the model reproduces the timing of seasonal groundwater variability. The value of reported significance tests should not be overstated, as data are temporally autocorrelated and spatially aggregated. The purpose of reporting these results is simply to demonstrate the detectable improvement in CLSM provided by the assimilation of GRACE data. The brevity of the GRACE data period (less than 4 yr) makes it difficult to assess the impact of assimilation on CLSM's performance with respect to interannual variability. We found that correlations calculated for anomalies of monthly groundwater (after removing the seasonal cycle) were statistically indistinguishable for all simulations (not shown).

TABLE 1. Evaluation of groundwater estimates from open-loop and assimilation integrations against measured groundwater. Correlation coefficient *r* and RMSE (mm) are calculated with respect to daily average groundwater storage based on observations from 58 piezometers. Skill due to assimilation is calculated relative to the open-loop simulation ("skill" = $1 - \mathrm{RMS}_{\mathrm{Assim}}/\mathrm{RMS}_{\mathrm{OL}}$). Bold font indicates a significant increase in *r* relative to OL at the 5% significance level. Italics indicate increase in *r* at the 10% significance level. Significance was assessed using a Student's *t* test.

	OL		GRACE DA		
	r	RMSE	*r*	RMSE	Skill
Mississippi	0.59	23.5	**0.70**	18.5	0.21
Ohio–Tennessee	0.78	62.8	**0.82**	40.4	0.36
Upper Mississippi	0.29	42.6	0.27	39.6	0.07
Red–Arkansas/Lower Mississippi	0.69	30.9	*0.72*	26.4	0.15
Missouri	0.41	24.5	**0.66**	19.7	0.20

c. *Hydrologic fluxes*

Comparisons with groundwater are useful but do not constitute a complete evaluation. Despite improved simulation of hydrologic states, data assimilation may in fact degrade simulated hydrologic fluxes because of the way in which simulated processes were calibrated during model development. For example, increasing soil moisture via data assimilation may cause an LSM to compensate by overestimating drainage or evaporation (Walker and Houser 2005). Here we consider the impact of GRACE assimilation on runoff and evapotranspiration in CLSM simulations. Table 2 shows that accumulated runoff in GRACE DA was slightly less than that in OL in the Ohio–Tennessee subbasin, due primarily to drier wintertime conditions, and slightly wetter in the Missouri and combined Red–Arkansas/Lower Mississippi basins, due to wetter wintertime conditions. It is difficult to assess the relative accuracy of the simulations in this regard, as CLSM does not include a runoff routing scheme that would allow for direct comparisons with river gauge data, but it is reassuring that assimilation did not lead to massive changes in total simulated runoff.

With respect to hydrologic variability, GRACE DA exhibits significantly stronger correlation between TWS and monthly accumulated gauged river flow in all three of the closed major subbasins and for the Mississippi River on the whole (Table 2). TWS was chosen as the variable of comparison because it is expected to correlate with total basin runoff and the correlation does not

TABLE 2. Mean runoff (mm yr^{-1}) and coefficient of linear correlation between CLSM TWS, CLSM runoff (R), and monthly average gauged river flow. Gauged runoff is derived from measurements at USGS and Army Corps of Engineers gauging stations. TWS was used for correlations because CLSM lacks a routing module, reducing the power of the direct runoff comparison. Correlations were not calculated for the Red–Arkansas/Lower Mississippi basin because of upstream contributions to gauged flow. Bold font indicates significance at the 5% level.

	OL			GA		
	Mean	r_{TWS}	r_R	Mean	r_{TWS}	r_R
Mississippi	34.1	0.75	0.67	34.0	**0.79**	0.68
Ohio–Tennessee	120.2	0.66	0.83	113.7	**0.71**	0.86
Upper Mississippi	7.6	0.61	0.75	8.7	**0.69**	0.74
Red–Arkansas/Lower Mississippi	47.1	—	—	49.0	—	—
Missouri	2.8	0.43	0.71	3.6	**0.57**	0.75

involve a routing-dependant time lag. Correlations with simulated runoff are shown as well, though the lack of a routing scheme in CLSM limits the interpretation of this comparison. Taken in combination, the data in Table 2 indicate that the assimilation of GRACE data produced no major artifacts in CLSM simulation of runoff quantity, and that it did improve the phase of the annual cycle.

Assimilation had a detectable effect on simulated evapotranspiration (ET) in some regions. It is difficult to evaluate this effect quantitatively, as the difference between OL and GRACE DA simulations is small relative to the differences between ET produced by running differing land surface models or changing atmospheric forcing datasets (Kato et al. 2007). Over the 41 months of simulation, differences between OL and GRACE DA were also small relative to differences between model simulation and point-scale surface measurements available from flux towers. Nonetheless, variability in soil moisture and in groundwater are known to cause changes in surface energy partitioning that are relevant to climate (Fan et al. 2007; Koster et al. 2004). As such, it is worth noting that data assimilation influenced the spatial patterning of both soil moisture and surface fluxes throughout the simulation. Figure 8 shows an example of this for April 2005, a month that was typical of springtime differences between OL and GRACE DA integrations. Data assimilation led to wetter conditions for much of the Ohio–Tennessee, Upper Mississippi, and combined Red–Arkansas/Lower Mississippi subbasins, and produced somewhat drier conditions for portions of the Missouri subbasin. These changes in moisture led to substantial differences in latent heat flux for all subbasins except the Ohio–Tennessee, where evaporation was not moisture limited during this period. For portions of the Southern Great Plains, the mean monthly difference in latent heat flux was as large as 20.6 W m^{-2}. Previous studies have indicated that a change in energy partitioning on the order of 15–20 W m^{-2} can have a significant impact on boundary layer processes and precipitation (Schar et al. 1999), and the Southern Great Plains is known as a region with strong land–atmosphere coupling.

d. Horizontal disaggregation

While the CLSM represents subcatchment partitioning between saturated, unsaturated, and wilting frac-

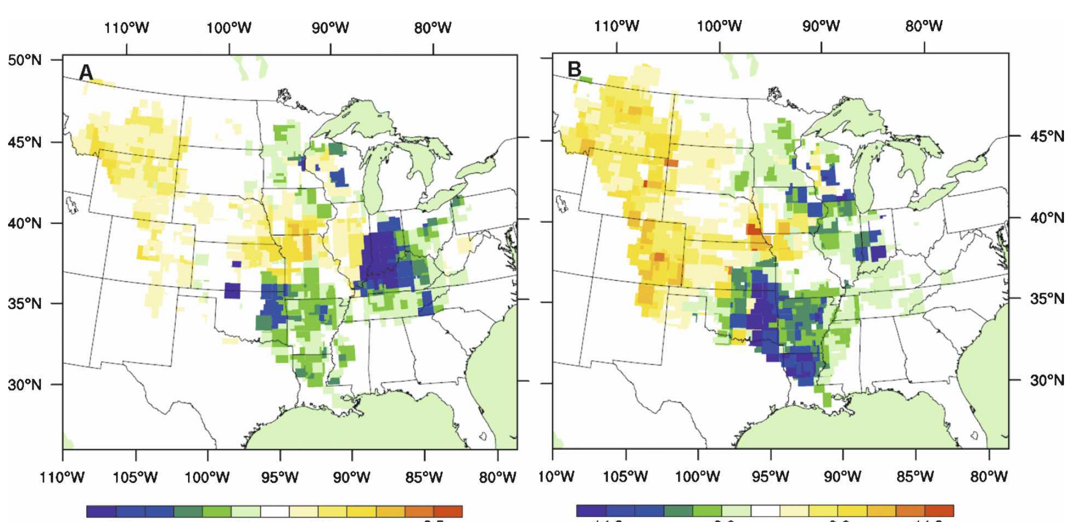

FIG. 8. Influence of assimilation on (a) root zone moisture (%) and (b) latent heat flux (W m^{-2}), plotted as GRACE DA − OL for the month of April 2005.

tions, as described in section 3, there is no physical interaction between different catchments. The horizontal distribution of GRACE-derived assimilation increments, then, is controlled entirely by the EnKS, which uses the variance in the ensemble to determine the horizontal disaggregation of information from GRACE observations. Equations (1) and (2) specify that the assimilation increments are based on (i) the difference between the observation and the corresponding model estimate at the observation scale, (ii) the error variances of the observation and of the corresponding model estimate, and (iii) the modeled (error) cross covariance between the observed variable (at the observation scale) and the model states at the finescale model resolution (in tile space). This means that the increment for a given tile will be large if (i) the model and GRACE disagree at the subbasin scale of the observation, (ii) the confidence in the model (at the observation scale) is low relative to the observation, and (iii) there is a strong (error) correlation between the observed variable (at the coarse scale) and the model states in the given tile. There is no deterministic hydrological basis for the distribution of increments at scales finer than the observations, so care must be taken to represent model and observation error characteristics as accurately as possible (section 3). Because it is impossible to do that perfectly, errors at the tile scale may increase at some locations even as errors at the subbasin scale of the GRACE observations are reduced.

To evaluate the effects of assimilation at subobservational scales, we assessed the correlation between simulated TWS variability and gauged river flow in eight smaller watersheds within the Mississippi (Fig. 1). For five of eight watersheds, GRACE DA yielded significantly higher correlation than OL (Table 3). For these smaller basins, correlations between simulated runoff and gauge data can be viewed with greater confidence, as the absence of a routing scheme is less important. The comparison is imperfect, but the fact that correlations between simulated and gauged runoff were greater for GRACE DA than OL in seven of eight basins (marginally significantly in two) is encouraging. Based on these comparisons, it appears that the assimilation of spatially coarse GRACE observations into a higher-resolution CLSM simulation has the potential to disaggregate the observations with a fair degree of skill.

TABLE 3. Gauged discharge ($m^3 s^{-1}$) for eight smaller watersheds in the Mississippi River basin (locations mapped in Fig. 1). Also shown are correlations between gauged discharge and both TWS and runoff (R) in the OL and GRACE DA CLSM simulations, calculated using daily data for January 2003 through May 2006. Bold font indicates improvement in correlation relative to open-loop simulation at the 5% significance level. Italics indicate improvement at the 10% significance level.

		rTWS		rR	
River	Discharge	OL	GA	OL	GA
Kanawha	537	0.41	0.42	0.53	0.53
Wabash	1001	0.55	0.63	0.31	0.32
Illinois	527	0.68	0.71	0.40	0.41
Minnesota	160	0.61	**0.68**	0.44	0.45
Arkansas	240	0.19	**0.28**	0.37	0.39
Ouachita	83	0.37	0.35	0.04	0.04
Yellowstone	212	0.24	0.26	0.34	*0.41*
Kansas	107	0.4	**0.49**	0.63	*0.68*

5. Discussion

To date, variability in terrestrial water storage has not been well defined by traditional observation techniques or land surface models (Dirmeyer et al. 2006). As the only remote sensing system capable of measuring water storage changes at all levels on and below the land surface, GRACE provides an unprecedented opportunity to improve quantification, understanding, and simulation of TWS variability. Yet the fact that GRACE measures water at all depths simultaneously is also a challenge, and its spatial and temporal resolutions are coarse by any standard of earth science data. Data assimilation shows much promise for effective vertical, horizontal, and temporal disaggregation of the monthly, basin-scale, integrated water column observations provided by GRACE, and thus adds value to these unique observations for research and applications.

In this study, GRACE-derived TWS anomalies were assimilated to the Catchment LSM by means of an ensemble Kalman smoother. The results were encouraging, including (i) decreases in RMS errors and significant increases in correlation between simulated and measured groundwater in the Mississippi River basin, (ii) improved simulation of hydrologic variability at the subobservation scale, and (iii) a small increase in correlation between simulated runoff and gauged river flow in the majority of test watersheds. Evaluation of the assimilation results at scales finer than the GRACE products revealed no degradation of model performance due to the assimilation of the coarse data.

These results indicate that a GRACE data assimilation system can contribute to large-scale drought monitoring. Drought monitors at the national and continental scale currently suffer from a paucity of data on groundwater variability. Data on soil moisture variability is also limiting. GRACE DA results provide estimates of both components, and, using observation-based forcing data, these fields can be extended to near–real time. The power of data assimilation to disaggregate, downscale, and extrapolate spatially coarse GRACE observations has the potential to improve the accuracy and objectivity with which drought conditions are identified.

Another potential application of GRACE assimilation pertains to the influence of data assimilation on simulated evapotranspiration. Recent studies (Bierkens and van den Hurk 2007; York et al. 2002) have demonstrated that variability in groundwater can be relevant to simulations of weather and climate, largely due to its influence on soil moisture and evapotranspiration. At present, GRACE is the best technology available for near-real-time, global monitoring of TWS, and the assimilation algorithm presented here skillfully partitions the GRACE measurement into components of groundwater and soil moisture. The application of the assimilation scheme to coupled land–atmosphere modeling systems would be expected to modify surface fluxes at a magnitude relevant to climate in some regions. Integration with coupled models will require more rigorous evaluation of surface water and energy fluxes produced by the GRACE–CLSM assimilation system. This evaluation is the subject of ongoing research.

Acknowledgments. This research was funded by multiple grants from NASA's Terrestrial Hydrology Program and Solid Earth and Natural Hazards Program. GRACE data were processed by Don P. Chambers, supported by NASA's REASoN Program. We thank the USGS and the Army Corps of Engineers for furnishing most of the evaluation data. We also thank Randy Koster and Sarith Mahanama at NASA GSFC for assisting our implementation of the Catchment Land Surface Model, John Wahr and Sean Swenson at the University of Colorado for guidance on estimating GRACE errors, and Wade Crow of the USDA–HRSL for valuable discussions on the ensemble Kalman smoother.

REFERENCES

Andersen, O. B., S. I. Seneviratne, J. Hinderer, and P. Viterbo, 2005: GRACE-derived terrestrial water storage depletion associated with the 2003 European heat wave. *Geophys. Res. Lett.,* **32,** L18405, doi:10.1029/2005GL023574.

Bierkens, M. F. P., and B. van den Hurk, 2007: Groundwater convergence as a possible mechanism for multi-year persistence in rainfall. *Geophys. Res. Lett.,* **34,** L02402, doi:10.1029/2006GL028396.

Bosilovich, M., J. D. Radakovich, and A. da Silva, 2007: Skin temperature analysis and bias correction in a coupled land-atmosphere data assimilation system. *J. Meteor. Soc. Japan,* **85A,** 99–116.

Chambers, D. P., 2006: Evaluation of new GRACE time-variable gravity data over the ocean. *Geophys. Res. Lett.,* **33,** L17603, doi:10.1029/2006GL027296.

Chen, J., M. Rodell, C. R. Wilson, and J. S. Famiglietti, 2005: Low degree spherical harmonic influences on Gravity Recovery and Climate Experiment (GRACE) water storage estimates. *Geophys. Res. Lett.,* **32,** L14405, doi:10.1029/2005GL022964.

Clark, M. P., A. G. Slater, A. P. Barrett, L. E. Hay, G. J. McCabe, B. Rajagopalan, and G. H. Leavesley, 2006: Assimilation of snow covered area information into hydrologic and land-surface models. *Adv. Water Resour.,* **29,** 1209–1221.

Crowley, J. W., J. X. Mitrovica, R. C. Bailey, M. E. Tamisiea, and J. L. Davis, 2006: Land water storage within the Congo Basin inferred from GRACE satellite gravity data. *Geophys. Res. Lett.,* **33,** L19402, doi:10.1029/2006GL027070.

Dirmeyer, P. A., X. A. Gao, M. Zhao, Z. C. Guo, T. K. Oki, and N. Hanasaki, 2006: GSWP-2—Multimodel analysis and implications for our perception of the land surface. *Bull. Amer. Meteor. Soc.,* **87,** 1381–1397.

Dunne, S., and D. Entekhabi, 2005: An ensemble-based reanalysis approach to land data assimilation. *Water Resour. Res.,* **41,** W02013, doi:10.1029/2004WR003449.

——, and ——, 2006: Land surface state and flux estimation using the ensemble Kalman smoother during the Southern Great Plains 1997 field experiment. *Water Resour. Res.,* **42,** W01407, doi:10.1029/2005WR004334.

Ellett, K. M., J. P. Walker, A. W. Western, and M. Rodell, 2006: A framework for assessing the potential of remote-sensed gravity to provide new insight on the hydrology of the Murray-Darling basin. *Aust. J. Water Resour.,* **10,** 89–101.

Evensen, G., and P. J. van Leeuwen, 2000: An ensemble Kalman smoother for nonlinear dynamics. *Mon. Wea. Rev.,* **128,** 1852–1867.

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 equilibrium water table simulations. *J. Geophys. Res.,* **112,** D10125, doi:10.1029/2006JD008111.

Kato, H., M. Rodell, F. Beyrich, H. Cleugh, E. van Gorsel, H. Liu, and T. P. Meyers, 2007: Sensitivity of land surface simulations to model physics, land characteristics, and forcings, at four CEOP sites. *J. Meteor. Soc. Japan,* **85A,** 187–204.

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

——, 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.

——, and Coauthors, 2004: Regions of strong coupling between soil moisture and precipitation. *Science,* **305,** 1138–1140.

Luthcke, S. B., and Coauthors, 2006: Recent Greenland ice mass loss by drainage system from satellite gravity observations. *Science,* **314,** 1286–1289.

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 GRACE data. *J. Geophys. Res.,* **112,** D07103, doi:10.1029/2006JD007522.

Ravela, S., and D. McLaughlin, 2007: Fast ensemble smoothing. *Ocean Dyn.,* **57,** 123–134.

Reichle, R. H., 2007: Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR). *J. Geophys. Res.,* **112,** D09108, doi:10.1029/2006JD008033.

——, and R. D. Koster, 2003: Assessing the impact of horizontal error correlations in background fields on soil moisture estimation. *J. Hydrometeor.,* **4,** 1229–1242.

——, D. B. McLaughlin, and D. Entekhabi, 2002: Hydrologic data assimilation with the ensemble Kalman filter. *Mon. Wea. Rev.,* **130,** 103–114.

——, R. D. Koster, P. Liu, S. P. P. Mahanama, E. G. Njoku, and M. Owe, 2007: Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR). *J. Geophys. Res.,* **112,** D09108, doi:10.1029/2006JD008033.

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.

——, and P. R. Houser, 2004: Updating a land surface model with MODIS-derived snow cover. *J. Hydrometeor.,* **5,** 1064–1075.

——, J. S. Famiglietti, J. Chen, S. I. Seneviratne, P. Viterbo, S. Holl, and C. R. Wilson, 2004a: Basin scale estimates of evapotranspiration using GRACE and other observations. *Geophys. Res. Lett.,* **31,** L20504, doi:10.1029/2004GL020873.

——, and Coauthors, 2004b: The global land data assimilation system. *Bull. Amer. Meteor. Soc.,* **85,** 381–394.

——, J. Chen, H. Kato, J. 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.

Schar, C., D. Luthi, U. Beyerle, and E. Heise, 1999: The soil-precipitation feedback: A process study with a regional climate model. *J. Climate,* **12,** 722–741.

Slater, A. G., and M. P. Clark, 2006: Snow data assimilation via an ensemble Kalman filter. *J. Hydrometeor.,* **7,** 478–493.

Stieglitz, M., A. Ducharne, R. Koster, and M. Suarez, 2001: The impact of detailed snow physics on the simulation of snow cover and subsurface thermodynamics at continental scales. *J. Hydrometeor.,* **2,** 228–242.

Swenson, S., and J. Wahr, 2006a: Estimating large-scale precipitation minus evapotranspiration from GRACE satellite gravity measurements. *J. Hydrometeor.,* **7,** 252–270.

——, and ——, 2006b: Post-processing removal of correlated errors in GRACE data. *Geophys. Res. Lett.,* **33,** L08402, doi:10.1029/2005GL025285.

——, P. J. F. Yeh, J. Wahr, and J. 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.

Syed, T. H., J. S. Famiglietti, J. Chen, M. Rodell, S. I. Seneviratne, P. Viterbo, and C. R. Wilson, 2005: Total basin discharge for the Amazon and Mississippi River basins from GRACE and a land-atmosphere water balance. *Geophys. Res. Lett.,* **32,** L24404, doi:10.1029/2005GL024851.

Tamisiea, M. E., E. W. Leuliette, J. L. Davis, and J. X. Mitrovica, 2005: Constraining hydrological and cryospheric mass flux in southeastern Alaska using space-based gravity measurements. *Geophys. Res. Lett.,* **32,** L20501, doi:10.1029/2005GL023961.

Tapley, B. D., S. Bettadpur, 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.

van Leeuwen, P. J., 2001: An ensemble smoother with error estimates. *Mon. Wea. Rev.,* **129,** 709–728.

Velicogna, I., and J. Wahr, 2006: Measurements of time-variable gravity show mass loss in Antarctica. *Science,* **311,** 1754–1756.

Verdin, K. L., and J. P. Verdin, 1999: A topological system for delineation and codification of the earth's river basins. *J. Hydrol.,* **218,** 1–12.

Wagner, C., D. McAdoo, J. Klokocnik, and J. Kostelecky, 2006: Degradation of geopotential recovery from short repeat-cycle orbits: Application to GRACE monthly fields. *J. Geod.,* **80,** 94–103.

Wahr, J., S. Swenson, and I. Velicogna, 2006: Accuracy of GRACE mass estimates. *Geophys. Res. Lett.,* **33,** L06401, doi:10.1029/2005GL025305.

Walker, J. P., and P. R. Houser, 2005: Hydrologic data assimilation. *Advances in Water Science Methodologies,* A. Aswathanarayana, Ed., A. A. Balkema, 230 pp.

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.

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.