Assimilation of GRACE terrestrial water storage data into a land surface model: results for the Mississippi River basin

by

Benjamin F. Zaitchik, Matthew Rodell, and Rolf H. Reichle

Popular Summary:
NASA's GRACE mission has the potential to be extremely valuable for water resources applications and global water cycle research. What makes GRACE unique among Earth Science satellite systems is that it is able to monitor variations in water stored in all forms, from snow and surface water to soil moisture to groundwater in the deepest aquifers. However, the space and time resolutions of GRACE observations are coarse. GRACE typically resolves water storage changes over regions the size of Nebraska on a monthly basis, while city-scale, daily observations would be more useful for water management, agriculture, and weather prediction. High resolution numerical (computer) hydrology models have been developed, which predict the fates of water and energy after they strike the land surface as precipitation and sunlight. These are similar to weather and climate forecast models, which simulate atmospheric processes. We integrated the GRACE observations into a hydrology model using an advanced technique called data assimilation. The results were new estimates of groundwater, soil moisture, and snow variations, which combined the veracity of GRACE with the high resolution of the model. We tested the technique over the Mississippi River basin, but it will be even more valuable in parts of the world which lack reliable data on water availability.

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 & Hydrological Sciences Branch, NASA Goddard Space Flight Center

Matthew Rodell, Hydrological Sciences Branch, NASA Goddard Space Flight Center

Rolf H. Reichle, Goddard Earth Sciences and Technology Center, University of Maryland Baltimore County & Global Modeling and Assimilation Office, NASA Goddard Space Flight Center

*Corresponding author's address:

Hydrological Sciences Branch, Code 614.3

NASA Goddard Space Flight Center

Greenbelt, MD 20771

bzaitchik@hsb.gsfc.nasa.gov

Abstract

Assimilation of data from the Gravity Recovery and Climate Experiment (GRACE) system of satellites yielded improved estimates of groundwater variability in the Mississippi river basin, as evaluated against independent measurements. We assimilated monthly terrestrial water storage anomalies from GRACE into the Catchment Land Surface Model (CLSM) with an ensemble Kalman filter for each of the four major sub-basins of the Mississippi over a 41-month period. Compared with the open loop CLSM simulation, assimilation estimates of groundwater exhibited significant increases in skill for the Missouri sub-basin, the combined Red-Arkansas/Lower-Mississippi sub-basin, and the Mississippi basin as a whole. Assimilation also improved runoff estimates in all four sub-basins, but not significantly. Assimilation was moderately successful at horizontally disaggregating GRACE data. At smaller spatial scales, the performance of the assimilation system was highly sensitive to the level of error ascribed to the GRACE observations. Spurious spikes appeared in the water storage estimates as the GRACE observation error estimate was reduced in the assimilation. The spikes should be addressed in future more sophisticated versions of the assimilation algorithm.

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.

In order to realize the full potential of GRACE for hydrology, the derived regional-scale, column-integrated, monthly water storage anomalies must be disaggregated spatially, vertically, and in time. Observational estimates of TWS from GRACE are routinely generated on a monthly basis, though techniques are under development for producing 10-day estimates (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 example, initialization of seasonal forecasts requires accurate estimates of variables such as TWS that constitute 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. (2006) computed groundwater storage variations averaged over the Mississippi River basin and its four major sub-basins 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 groundwater storage estimates based on piezometer observations over the full Mississippi River basin and the two larger sub-basins. 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 (Ellett et al., 2006). 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, 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 (Margulis et al., 2002; Crow and Wood, 2003; Reichle and Koster, 2005), 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, are used to enhance the satellite observations, providing downscaling and quality control of GRACE observations while enabling synthesis of data from multiple observing systems 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.

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

4

lack of a groundwater reservoir is a deficiency which 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 assimilation algorithm based on an Ensemble Kalman Filter (EnKF). The complete assimilation system is presented as follows. In Section 2 the GRACE data are described in greater detail. Section 3 reviews relevant features of the CLSM. Section 4 presents the EnKF. Results are given in Section 5 and conclusions in Section 6.

2. 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 NASA's

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 a set of spherical harmonics

coefficients up to degree and order 120, based on highly precise K-band microwave

measurements of the distance between two identical satellites orbiting Earth 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, 2007) were smoothed using a Gaussian averaging kernel with 400

km radius and "destriped" following Swenson and Wahr (2006). Water storage changes

were extracted for each of the four major sub-basins of the Mississippi River: the Ohio-

Tennessee, Upper Mississippi, Missouri, and combined Red-Arkansas / Lower

Mississippi (Figure 1).

Monthly TWS anomalies were obtained for February 2003 – April 2006 from CSR, for

February 2003 – May 2006 from GFZ, and from January 2003 – November 2005 from

JPL. 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 (Figure 2), we used an average of available estimates to provide monthly assimilation inputs from January 2003 – May 2006.

Optimal data assimilation requires error estimates for both the model and the observation. This is not straight-forward 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 mid-latitude GRACE TWS measurements. The sub-basin average anomalies used in this study span several averaging radii, suggesting that actual error for the entire sub-basin may be smaller. For this reason a second assimilation integration was performed for which the error of GRACE sub-basin averages was assumed to be 10 mm. Finally, an assimilation experiment was run in which GRACE error was set to 1 mm. This low error estimate is inconsistent with current understanding of GRACE uncertainty, and it should be interpreted as a numerical experiment in which GRACE data were weighted heavily relative to model predictions.

3. The Catchment LSM

The Catchment LSM (CLSM; Koster et al., 2000) was developed in response to a perceived shortcoming in conventional land surface models: the layer-based vertical discretization of conventional LSMs is not well suited to surface hydrologic processes. In effect, layer-based 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 a series of topographically-defined catchments with an average area of approximately 4,000 km^2 and models hydrologic processes based on each catchment's topographical statistics. Sub-catchment 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 (Figure 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 are in communication with the catchment deficit and permit a rough representation of non-equilibrium vertical conditions such as infiltration fronts. Snow is represented in a state-of-the-art three-layer snow physics scheme (Stieglitz et al., 2001).

8

In this study CLSM simulations were performed for the Mississippi Basin. Forcing data were drawn from the GLDAS forcing database (Rodell et al., 2004b) at 2.0°x2.5° resolution. Catchment information was defined on the basis of a 30 arc-second 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 Basin comprises 783 defined catchments, resulting in 1950 tiles under a 1.0°x 1.25° atmospheric grid (Figure 4). The model was spun up for 10 years under 2002 forcing conditions and integrated from 1 January 2003 through the end of available GRACE data, 1 June 2006. The results of four integrations are described in this paper: an open loop simulation without assimilation (OL) and three assimilation integrations with GRACE estimated error set to 20mm (A20), 10mm (A10) and 1mm (A1), respectively. Each integration included 20 ensemble members with statistically perturbed forcing fields; additional integrations with 12 and 100 ensemble members each were performed for comparison, and they yielded similar results.

4. Data Assimilation Algorithm

The defining characteristics of CLSM make the model particularly appropriate for the assimilation of GRACE-derived TWS anomalies. As described in Section 2, 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. The vertical distribution of soil moisture implicit in CLSM moisture reservoirs is also well-suited for GRACE assimilation. The presence of a variable water table is essential, since it means that the model accounts for at least part of the groundwater variability measured by GRACE. Moreover, CLSM's lack of traditional hydrologic layers in the subsurface is convenient, as the TWS increment can be applied directly to a column-integrated prognostic variable (the *catchment deficit*), 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 sub-basin to numerous model catchments within the sub-basin, but it does furnish a rationale for spatially distributing the effects of assimilation at the sub-catchment scale.

GRACE TWS anomalies were assimilated using an Ensemble Kalman Filter (EnKF) scheme, which is an adaptation of the standard Kalman filter for moderately non-linear systems (Evensen, 1994;Reichle et al., 2002). In the EnKF, 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.

Stated generically, the EnKF update at time step (t) is based on an ensemble of state

vectors $\mathbf{X}_t^i$ of the model predicted variable (X) from each ensemble member (i) that

captures all sources of uncertainty related to model physics, parameters, and forcing data,

such that the ensemble spread of $\mathbf{X}_t^i$ constitutes an estimate of error in model predictions.

The measurement process for observations is expressed as

$$\mathbf{Y}_t = \mathbf{M}_t[\mathbf{X}(t)] + \mathbf{v}_t. \tag{1}$$

That is, the observation vector $\mathbf{Y}_t$ is related to the "true" model state vector $\mathbf{X}(t)$ by a

nonlinear operator ($\mathbf{M}_t$). Uncertainties in the measurement process are reflected in the

error vector $\mathbf{v}_t$ (with $\mathbf{C}_v$ denoting its covariance matrix). Vectors $\mathbf{X}_t^i$ and $\mathbf{Y}_t$ need not be of

the same length: $\mathbf{X}_t^i$ has length n (for n model variables), while $\mathbf{Y}_t$ has length m (for m

observations for the given timestep).

The assimilation update is then computed for each ensemble member as

$$\mathbf{X}_{t+}^i = \mathbf{X}_{t-}^i + \mathbf{K}_t\left[\mathbf{Y}_t - \mathbf{M}_t\left(\mathbf{X}_{t-}^i\right) + \mathbf{v}_t^i\right]. \tag{2}$$

Here, (-) indicates the state vector prior to update and (+) is the state vector after update,

and the vector $\mathbf{v}_t^i$ is a random realization of the observation error. 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,

$$\mathbf{K}_t = \left[\mathbf{C}_{XM} \left(\mathbf{C}_M + \mathbf{C}_v \right)^{-1} \right]_t. \tag{3}$$

The matrix $\mathbf{C}_M$ is the covariance of the transformed measurement predictions ($M_t[\mathbf{X}(t)]$),

and $\mathbf{C}_{XM}$ is the cross-covariance between the state $\mathbf{X}_t$ and $M_t[\mathbf{X}_t]$. The cross-covariance

$\mathbf{C}_{XM}$ is particularly important because it provides the basis for the distribution of

observational information from the coarse sub-basin scale to the finer-scale tile space.

Since $\mathbf{C}_{XM}$ 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).

Here, we apply the EnKF update separately for each sub-basin, that is the model state

matrix $\mathbf{X}_t$ has dimensions N_j by 20, reflecting an ensemble of 20 members for the

catchment deficit in each of the N_j tiles in sub-basin j (j=1,…,4). $\mathbf{Y}_t$, meanwhile, is a

scalar, as only one GRACE observation is available for each major sub-basin at each

assimilation time step. The transformation function $M_t[\mathbf{X}_t]$ includes all calculations

involved in producing TWS anomaly estimates from GRACE orbital data, and we take

the reported error of this process to be the (scalar) observation error variance $\mathbf{C}_v$. GRACE

observations were not scaled for assimilation except that the 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.

In the past, the EnKF has been applied successfully to observations with roughly the

same spatial and temporal scales as the land surface model. GRACE data, however,

present additional challenges due to their coarse temporal and spatial resolutions. In this

early study we took a simple approach to these challenges. Temporally, we apply the entire increment at the mid-date of the observation averaging period (typically the 15th of the calendar month). This simple approach seems appropriate for TWS, which is expected to vary slowly on the average, and which will be distributed between communicating water reservoirs in the CLSM over time. Spatially, the perturbations that were added to the model forcings and prognostic variables of each ensemble member during the assimilation integration were generated with a horizontal correlation scale of 2 degrees, which very roughly represents error scales in global-scale precipitation fields (Reichle and Koster, 2003). Note again that the horizontal error correlations contained in C_{XM} dictate the horizontal distribution of GRACE observational information from the sub-basin scale onto the catchments that are contained within the sub-basin.

Vertically, for catchments that were modeled to be snow-free at the time of assimilation, the entire increment was applied to the *catchment deficit*. This is justified because the *catchment deficit* defines the equilibrium soil water storage in CLSM, and equilibrium is the only reasonable assumption for the application of a monthly averaged observation. For catchments with snow cover, positive increments (i.e., "wetting") were applied entirely to snow. Negative increments ("drying") were applied first to snow and then, if all snow was removed, to the *catchment deficit*. Each of these simple approaches can be refined in future research.

5. Results

5.1 Terrestrial Water Storage

As seen in Figure 2, the open loop simulation with the CLSM captured the general

seasonal cycle of terrestrial water storage (TWS) for the four major sub-basins of the

Mississippi. All four basins experience a wintertime peak in TWS followed by a summer

trough (see also Figure 5). The open loop simulation differed substantially from GRACE

estimates, however, in the magnitude of the seasonal cycle, in inter-annual variability,

and, more subtly, in the timing of the wetting and drying that occurs in each year.

Notably, seasonal TWS variations in the Missouri Basin were only 20-40 mm in the open

loop simulation, but ranged from 70 to 130 mm in GRACE retrievals. In the Ohio-

Tennessee, the open loop simulation returned a large and growing annual TWS cycle,

while the annual cycle of GRACE 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 (Figure 5). The A10 simulation more

closely resembled GRACE TWS estimates than the A20 on account of the greater

prescribed confidence in the GRACE data. Both A10 and A20 experienced some

unrealistic spikes due to the once-per-month updates, which is discussed further in the

next section. Only during the four month data gap in the summer of 2004 did the

assimilation runs match the open loop simulation.

There was one systematic exception in which CLSM assimilation simulations did not

track GRACE estimates of TWS. During observed dry periods in the Missouri Basin

14

none of the simulations dried as much as GRACE observations would indicate. This occurred because at these points the dry anomaly observed by GRACE exceeded the maximum possible catchment deficit of the CLSM. When assimilation increments attempt to dry the model beyond this model limit the increment is truncated. One possible interpretation of this discrepancy between model and observation is that the CLSM effectively capped the dynamic range of TWS according to physical principles that are not considered in GRACE observations. Another interpretation is that the GRACE observation captures TWS components not included in the CLSM, such as surface water and confined groundwater aquifers. Yet another interpretation is that GRACE has revealed an inaccurate parameterization in the CLSM, and that the model should be adjusted to allow greater drying in the Missouri Basin. In any case, it is clear that data assimilation provides information which is potentially valuable for refining both models and observing systems. Work is ongoing to reveal the source of this discrepancy, which may lead to adjustments in CLSM, our interpretation of GRACE observations, or both. An important consequence of CLSM's maximum possible catchment deficit in this study is that it caused the mean assimilation increment to be greater than zero. This contributed to wetter soils and greater runoff in the assimilation simulations relative to open loop.

5.2 Vertical disaggregation

One of the strongest motivations for assimilating GRACE data is the potential to vertically disaggregate the GRACE observation. Figure 6 breaks down estimated TWS into its components (shallow groundwater, soil moisture, and snow) and demonstrates the

influence of data assimilation on these components. Also shown in Figure 6 are independent observations of groundwater that were derived from unconfined and semi-confined water level records from 58 piezometers distributed across the Mississippi River basin (Rodell et al., 2006). 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.

Figure 6 shows that GRACE data 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 is more strongly controlled by atmospheric forcing. Figure 6 also reveals spikes in the temporal evolution of groundwater and soil moisture that are artifacts of the assimilation procedure, because the entire GRACE increment was applied at the observation mid-point. Future assimilation development should address this issue.

Table 1 quantifies the agreement between estimated and observed groundwater for the Mississippi as a whole and its four sub-basins and shows that GRACE data assimilation

significantly improved estimates of the amplitude and phase of the seasonal cycle of groundwater. For all sub-basins and the Mississippi as a whole, RMS errors in groundwater estimates decreased for all assimilation integrations when compared to the open loop simulation, resulting in positive skill scores. The only exception is for the A1 integration in the Missouri sub-basin, which will be discussed below. This result clearly demonstrates the positive impact of GRACE observations on the amplitude of groundwater estimates derived from data assimilation.

To assess the phase of the seasonal cycle, Table 1 shows time series correlations between modeled and observed groundwater. Averaged over the entire Mississippi basin, correlations were larger for the assimilation integrations than for the open loop simulation. This result was statistically significant at the 5 % level for the A10 and A1 simulations, and marginally significant (at the 10 % level) for the A20 simulation. The Missouri and Red-Arkansas/Lower-Mississippi sub-basins also experienced statistically significant improvements. We found a statistically significant decrease in correlation for the Ohio-Tennessee, but note again that assimilating GRACE data did improve RMSE, most dramatically in the A10 simulation.

Assimilation did not significantly change the poor correlation with observed groundwater in the Upper Mississippi sub-basin. 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 (Rodell et al., 2004b). There is no obvious

17

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.

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 brevity of the GRACE data period (less than four years) makes it difficult to assess the impact of assimilation on CLSM's performance with respect to inter-annual variability. We found that correlations calculated for anomalies of monthly groundwater (after removing the seasonal cycle) were statistically indistinguishable for all simulations (not shown).

5.3 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. For example, increasing soil moisture via data assimilation may cause the LSM to compensate by over-estimating drainage or evaporation. Here we consider the impact of GRACE assimilation on runoff and evapotranspiration in CLSM simulations. Evaluating Table 2 shows that accumulated runoff and runoff variability in CLSM were similar for the open loop, A20, and A10 simulations, and that correlation with river gauge data for the major sub-basins of the Mississippi was correspondingly similar. For the Mississippi River basin as a whole, total runoff was similar in all

18

simulations: A20 produced 3% less runoff than the open loop simulation, and A10

produced 2% more. For the Ohio-Tennessee, GRACE data assimilation reduced

simulated runoff. For the Missouri and Upper Mississippi, assimilation increased runoff.

Nonetheless, (one-month-lagged) correlations between monthly gauge based and CLSM

runoff demonstrate that the model captures runoff variability reasonably well. Marginal

improvements to runoff correlation provided by data assimilation are not statistically

significant, but the results indicate that assimilation of GRACE TWS data does not

degrade CLSM runoff at the basin scale. Note that the A1 simulation typically exhibits

higher bias and poorer correlation with gauged runoff when compared with the other

experiments. This result indicates that in A1 GRACE observations are weighted too

heavily in the assimilation scheme, and that doing so negatively impacts the simulation of

hydrologic fluxes even when hydrologic state variables are generally improved (Table 1).

Effects on simulated evapotranspiration (ET) were qualitatively similar to those for

runoff. Assimilation of GRACE data tended to cause a slight increase in simulated ET

for all basins (Table 3). The changes in ET from data assimilation are much smaller than

typical differences in ET when different LSMs or different precipitation forcing data sets

are used (Kato et al., 2007). Nonetheless, the influence of GRACE assimilation on ET is

sometimes dramatic at smaller spatial scales. In 2003, for example, the A10 simulation

yielded only slightly larger ET than the open loop simulation for each of the four major

sub-basins, but the difference in ET was concentrated geographically in the Great Plains

(Figure 7). This region is noted for strong land-atmosphere coupling (Koster et al.,

2004), hence the result, if confirmed, could have implications for short term to seasonal precipitation forecasting.

5.4 Horizontal disaggregation

While the CLSM represents sub-catchment partitioning between saturated, unsaturated, and wilting fractions, 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 EnKF, which uses the error information that is modeled in the ensemble to determine the horizontal disaggregation of information from GRACE observations. Recall from Section 4 (Equations (2) and (3)) 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 fine-scale model resolution (in tile space). This means that the increment for a given tile will be large if (1) the model and GRACE disagree at the sub-basin scale of the observation, (2) the confidence in the model (at the observation scale) is low relative to the observation, and (3) there is a strong (error) correlation between the observed variable (at the coarse scale) and the model states in the given tile.

Here, the ensemble produced considerable spatial variability in the magnitude of assimilation increments within the four sub-basins for which GRACE TWS anomalies were computed. As a result, output states and fluxes from the open loop and assimilation

20

simulations differed substantially in certain regions (e.g., Figure 7). This demonstrates a

basic principle of data assimilation—that observations can compensate for model

uncertainty—but there is no deterministic hydrological basis for the distribution of

increments at scales finer than the observations. Thus, care must be taken to represent

model and observation error characteristics as accurately as possible (Section 4). Since it

is impossible to capture the model and observation error characteristics perfectly, errors

at the tile scale may grow in some cases even as errors at the sub-basin scale of the

GRACE observations are reduced.

Simulated runoff, which is essentially a residual of a model's soil water budget, is prone

to absorbing errors introduced by hydrological data assimilation. Evaluation of model

results against river gauge data from eight smaller watersheds within the Mississippi

(Figure 1) indicates that this was not a problem for either the A20 or A10 simulations

(Table 4). Data assimilation had either a small positive or a small negative impact on

correlations between CLSM runoff and gauge data for these smaller watersheds, and the

effect was statistically insignificant in either direction. In contrast, the A1 assimilation

caused a reduction in the correlation between simulated runoff and observed data in six of

the eight watersheds, and this reduction was significant in two cases. This tendency was

somewhat evident at the scale of major sub-basins (Table 2), but aggregation hid the

significance of the error. For the Missouri, A1 actually produced one of the strongest

correlations with gauge data at the basin scale, even though simulated runoff in two

interior watersheds (the Kansas and Yellowstone) was severely degraded. The small-

scale comparison provides compelling evidence that GRACE data should not be

assimilated by direct insertion or other aggressive updating algorithms; imposing coarse observations on a higher resolution model can have undesirable effects when observation error is not taken into account properly.

On the other hand, there are some regions in which the A10 and A20 simulations seem to give inadequate weight to GRACE observations. High quality groundwater (Changnon et al., 1988) and in situ soil moisture (Hollinger and Isard, 1994) measurements in the state of Illinois make it possible to evaluate statewide TWS variability (Rodell and Famiglietti, 2001; Yeh et al., 1998; Swenson et al., 2006). Results are shown in Table 5. Compared with the in situ data, the A1 simulation showed the strongest correlation and highest skill relative to open loop of any simulation. It is also the only assimilation run that achieved a significant increase in correlation relative to the open loop simulation. It would seem, then, that in some regions GRACE observations have a greater potential to improve CLSM simulations than is realized in the A20 or A10 simulations. A more sophisticated error characterization in the assimilation system could perhaps take advantage of this.

6. Conclusions

Variability in terrestrial water storage is a source of considerable uncertainty in current LSMs (Dirmeyer et al., 2006). The paucity of direct observations of TWS, either *in situ* or remotely sensed, makes it difficult either to improve model formulations of storage dynamics or to evaluate models against independent data. As the only remote sensing system capable of measuring water storage changes at all levels below and on the land surface, GRACE provides an unprecedented opportunity to understand and improve 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 can add 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 Filter. The results were encouraging, including (i) decreases in RMS errors and significant increases in correlation between simulated and measured groundwater in the Mississippi Basin, (ii) improved skill in simulating TWS variability in Illinois, and (iii) a small increase in correlation between simulated runoff and gauged river flow. Generally, however, flux estimates were neither improved nor worsened significantly by the assimilation of GRACE observations. Evaluation of the assimilation results at scales finer than the GRACE data revealed no added uncertainty when an appropriate level of error was ascribed to the GRACE data.

This assigned level of GRACE observation error did have a significant impact on assimilation results. The best model results were obtained when the observation error standard deviation was set to 10 mm, which is lower than estimated by Wahr et al. (2006) but not unreasonable, given that the sub-basin scale of the observations considered here covers several GRACE averaging radii. When GRACE error was intentionally underestimated (1 mm) the skill of the assimilation estimates typically suffered, particularly at small spatial scales, except in terms of TWS in Illinois. These results suggest that assimilation with an estimate of 20 mm for GRACE TWS error standard deviation under-utilizes the information contained in GRACE data, while simulations with overly aggressive assimilation schemes introduce errors in locations where the water storage condition is not aligned with the regional mean. A more sophisticated assimilation algorithm with better representation of model errors, simultaneous assimilation of soil moisture observations, or time-varying GRACE uncertainty estimates could perhaps more effectively utilize the information contained in GRACE TWS observations while minimizing the introduction of error at sub-observation scales.

Acknowledgements

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.

Geophysical Research Letters, **32**.

Bosilovich, M., J. D. Radakovich, and A. da Silva, 2007: Skin temperature analysis and

bias correction in a coupled land-atmosphere data assimilation system. *Journal of

the Meteorological Society of Japan*, **In Press**.

Chambers, D. P., 2007: Evaluation of new GRACE time-variable gravity data over the

ocean. *Geophysical Research Letters*, **In Press**.

Changnon, S. A., F. Z. Huff, and C. F. Hsu, 1988: Relations between precipitation and

shallow groundwater in Illinois. *Journal of Climate*, **1**, 1239-1250.

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. *Advances in Water Resources*, **29**, 1209-

1221.

Crow, W. T. and E. F. Wood, 2003: The assimilation of remotely sensed soil brightness

temperature imagery into a land surface model using ensemble Kalman filtering:

A case study based on ESTAR measurements during SGP97. *Advances in Water

Resources*, **26**, 137-149.

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. *Geophysical Research Letters*, **33**.

Dirmeyer, P. A., X. A. Gao, M. Zhao, Z. C. Guo, T. K. Oki, and N. Hanasaki, 2006:
GSWP-2 - Multimodel anlysis and implications for our perception of the land
surface. *Bulletin of the American Meteorological Society*, **87,** 1381-1397.

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

Evensen, G., 1994: Sequential Data Assimilation with a Nonlinear Quasi-Geostrophic
Model Using Monte-Carlo Methods to Forecast Error Statistics. *Journal of
Geophysical Research-Oceans*, **99,** 10143-10162.

Hollinger, S. E. and S. A. Isard, 1994: A Soil-Moisture Climatology of Illinois. *Journal
of Climate*, **7,** 822-833.

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. *Journal of the Meteorological
Society of Japan*, **In Press**.

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

Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar, 2000: A
catchment-based approach to modeling land surface processes in a general
circulation model 1. Model structure. *Journal of Geophysical Research-
Atmospheres*, **105,** 24809-24822.

Koster, R. D., P. A. Dirmeyer, Z. C. Guo, G. Bonan, E. Chan, P. Cox, C. T. Gordon, S. Kanae, E. Kowalczyk, D. Lawrence, P. Liu, C. H. Lu, S. Malyshev, B. McAvaney, K. Mitchell, D. Mocko, T. Oki, K. Oleson, A. Pitman, Y. C. Sud, C. M. Taylor, D. Verseghy, R. Vasic, Y. K. Xue, and T. Yamada, 2004: Regions of strong coupling between soil moisture and precipitation. *Science*, **305,** 1138-1140.

Luthcke, S. B., H. J. Zwally, W. Abdalati, D. D. Rowlands, R. D. Ray, R. S. Nerem, F. G. Lemoine, J. J. McCarthy, and D. S. Chinn, 2006: Recent Greenland ice mass loss by drainage system from satellite gravity observations. *Science*, **314,** 1286-1289.

Margulis, S. A., D. McLaughlin, D. Entekhabi, and S. Dunne, 2002: Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment. *Water Resources Research*, **38.**

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. *Journal of Geophysical Research*, **In Press**.

Reichle, R. H. and R. D. Koster, 2003: Assessing the impact of horizontal error correlations in background fields on soil moisture estimation. *Journal of Hydrometeorology*, **4,** 1229-1242.

——, 2005: Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model. *Geophysical Research Letters*, **32.**

Reichle, R. H., D. B. McLaughlin, and D. Entekhabi, 2002: Hydrologic data assimilation with the ensemble Kalman filter. *Monthly Weather Review*, **130,** 103-114.

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 Resources Research*, **37,** 1327-1339.

Rodell, M. and P. R. Houser, 2004: Updating a land surface model with MODIS-derived snow cover. *Journal of Hydrometeorology*, **5,** 1064-1075.

Rodell, M., J. Chen, H. Kato, J. Famiglietti, J. Nigro, and C. Wilson, 2006: Estimating ground water storage changes in the Mississippi River basin (USA) using GRACE. *Hydrogeology Journal*, **doi:10.1007/s10040-006-0103-7**.

Rodell, M., 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. *Geophysical Research Letters*, **31**.

Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C. J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll, 2004b: The global land data assimilation system. *Bulletin of the American Meteorological Society*, **85,** 381-+.

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. *Geophysical Research Letters*, **32**.

Slater, A. G. and M. P. Clark, 2006: Snow data assimilation via an ensemble Kalman filter. *Journal of Hydrometeorology*, **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. *Journal of Hydrometeorology*, **2,** 228-242.

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

——, 2006b: Post-processing removal of correlated errors in GRACE data. *Geophysical Research Letters*, **33.**

Swenson, S., 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. *Geophysical Research Letters*, **33.**

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. *Geophysical Research Letters*, **32.**

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. *Geophysical Research Letters*, **32.**

Tapley, B. D., S. Bettadpur, M. Watkins, and C. Reigber, 2004: The gravity recovery and climate experiment: Mission overview and early results. *Geophysical Research Letters*, **31.**

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. *Journal of Hydrology*, **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. *Journal of Geodesy*, **80,** 94-103.

Wahr, J., S. Swenson, and I. Velicogna, 2006: Accuracy of GRACE mass estimates.

    Geophysical Research Letters, **33**.

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. *Journal of Geophysical Research-Atmospheres*, **103,**

    19823-19837.

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 Resources Research*, **42**.

Figure Captions

Figure 1: The four major sub-basins of the Mississippi: 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 sub-basin, including the (1) Kanahwa, (2) Wabash, (3) Illinois, (4) Minnesota, (5) Canadian, (6) Ouachita, (7) Yellowstone, and (8) Kansas Rivers, which were used in model evaluation (Table 4). Black dots indicate location of river gauges for the smaller watersheds.

Figure 2: Monthly TWS anomalies based on CSR, GFZ, and JPL GRACE observational estimates and from an open loop simulation with the Catchment Land Surface Model (CLSM).

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

Figure 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.

Figure 5: Five-day average TWS (mm), January 2003 – May 2006, from OL, A20, and A10 CLSM simulations, for (clockwise from the upper left) the Missouri, Upper

Mississippi, Ohio-Tennessee, and Lower Mississippi-Red-Arkansas sub-basins.

Also shown are monthly GRACE TWS anomalies, shifted to the CLSM mean.

Figure 6: Groundwater, soil moisture, and snow water equivalent for the Mississippi river basin for (A) open loop and (B) A10 simulation. 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.

Figure 7: A10 minus OL difference in accumulated evapotranspiration (mm) for 2003.
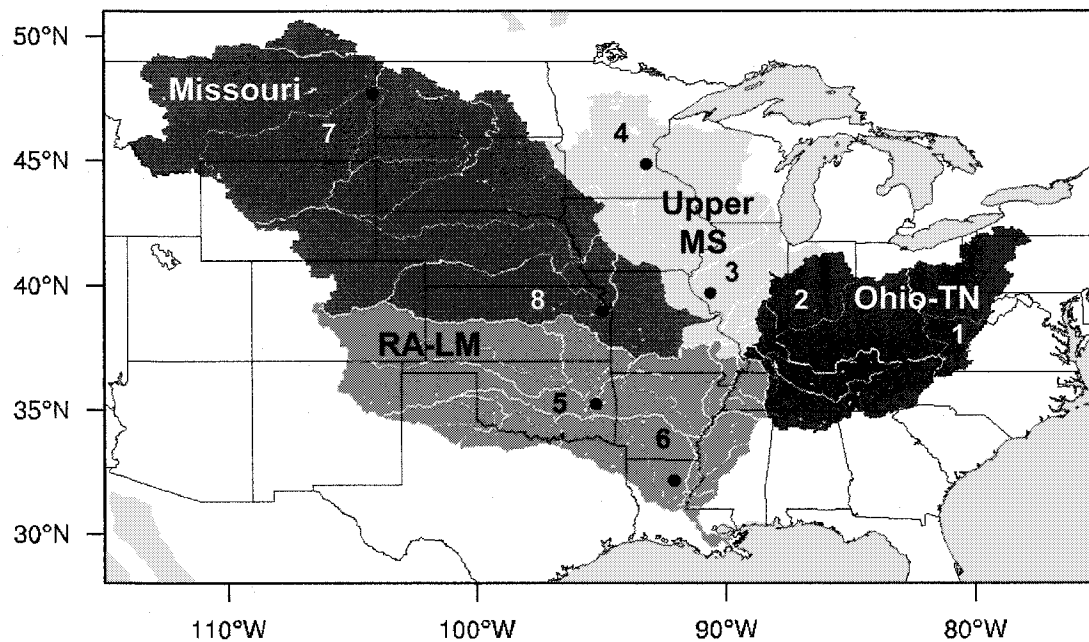
Figure 1: The four major sub-basins of the Mississippi: the Missouri, Upper Mississippi, Ohio-Tennessee, and the combined Red-Arkansas / Lower Mississippi (RA-LM). Thin white lines indicate the borders of smaller watersheds within each sub-basin, including the (1) Kanahwa, (2) Wabash, (3) Illinois, (4) Minnesota, (5) Canadian, (6) Ouachita, (7) Yellowstone, and (8) Kansas Rivers, which were used in model evaluation (Table 4). Black dots indicate location of river gauges for the smaller watersheds.
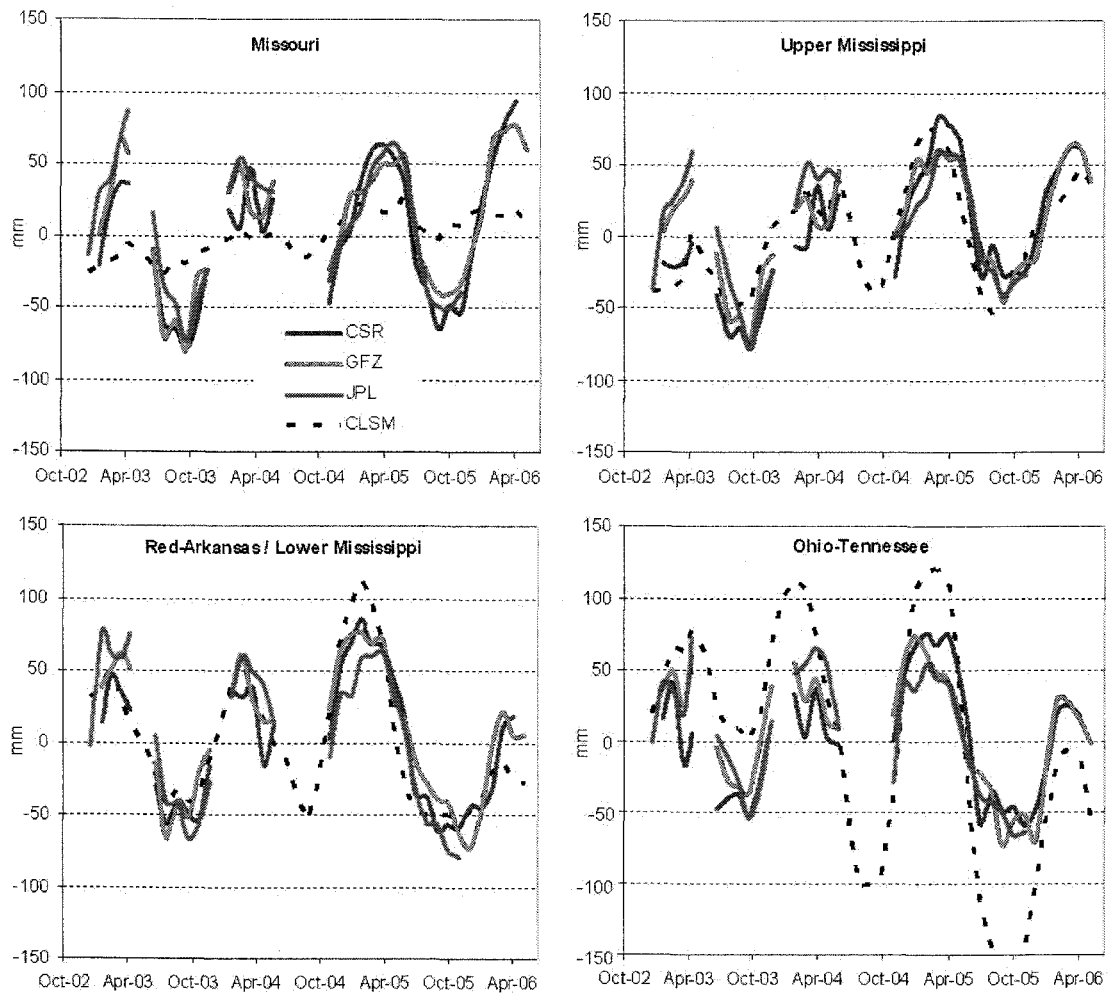
Figure 2: Monthly TWS anomalies based on CSR, GFZ, and JPL GRACE observational estimates and from an open loop simulation with the Catchment Land Surface Model (CLSM).
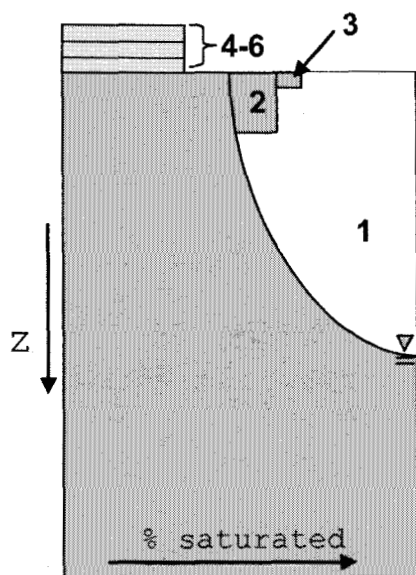
Figure 3: Prognostic hydrologic variables in the CLSM. (1) catchment deficit, (2) root
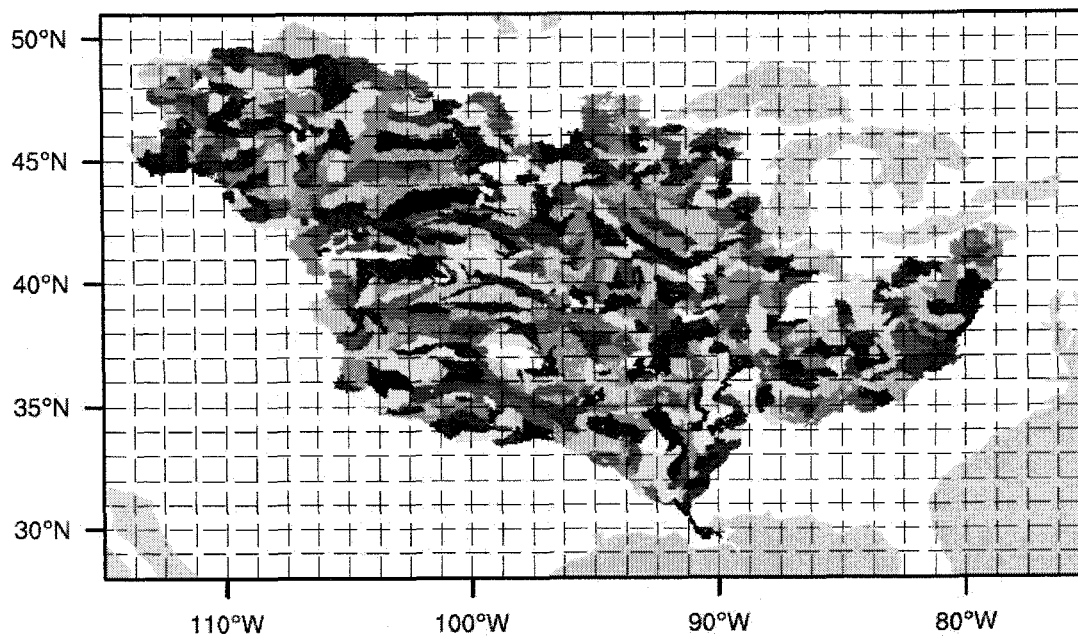zone excess, (3) surface excess, (4-6) three snow layers.

Figure 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.
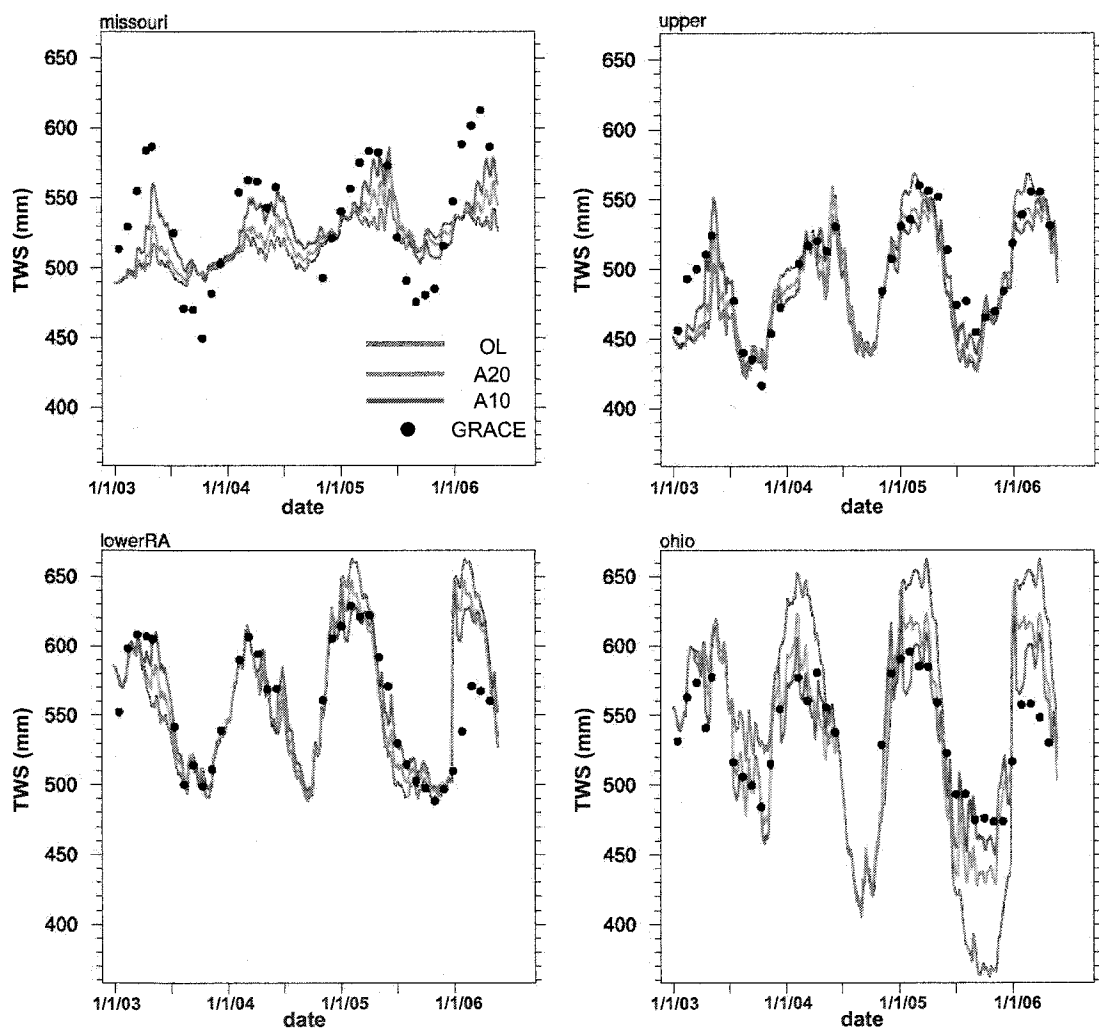
Figure 5: Five-day average TWS (mm), January 2003 – May 2006, from OL, A20, and A10 CLSM simulations, for (clockwise from the upper left) the Missouri, Upper Mississippi, Ohio-Tennessee, and Lower Mississippi-Red-Arkansas sub-basins. Also shown are monthly GRACE TWS anomalies, shifted to the CLSM mean.

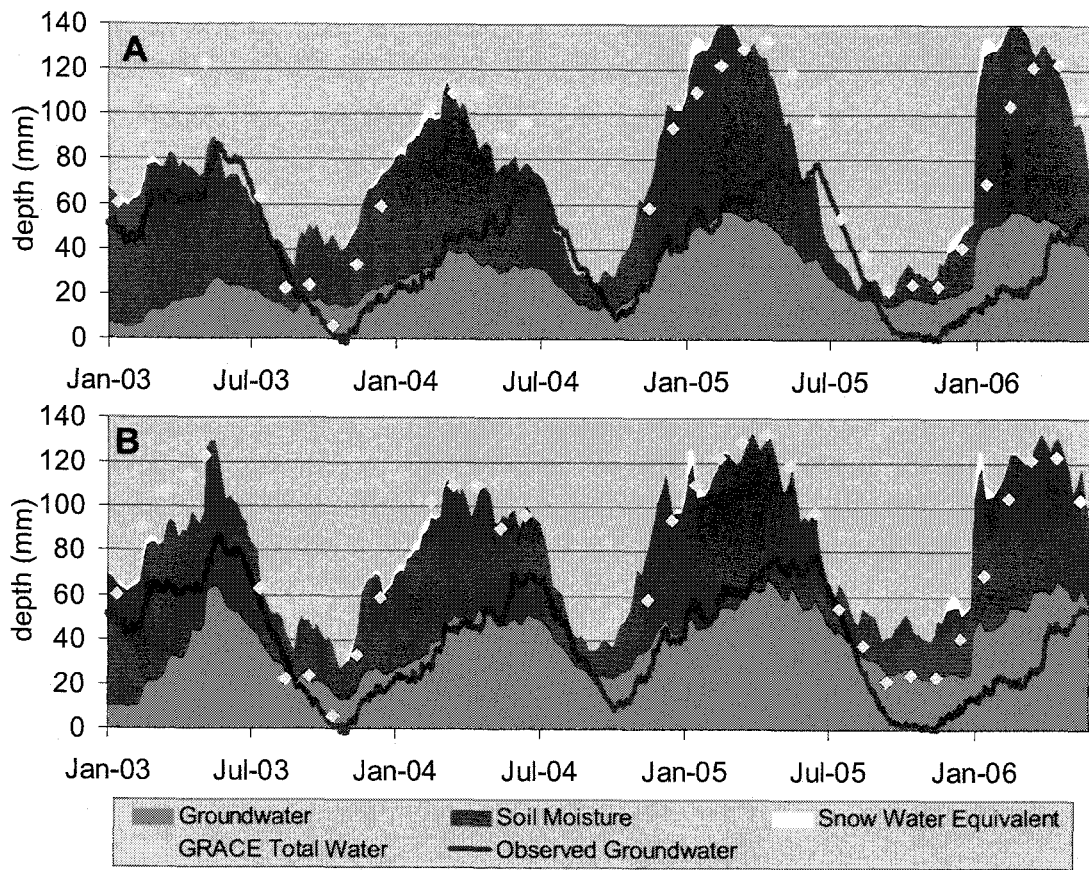Figure 6: Groundwater, soil moisture, and snow water equivalent for the Mississippi river
basin for (A) open loop and (B) A10 simulation. Also shown are area averaged
daily groundwater observations and monthly GRACE-derived TWS anomalies.
GRACE and modeled TWS are adjusted to a common mean, as are observed and
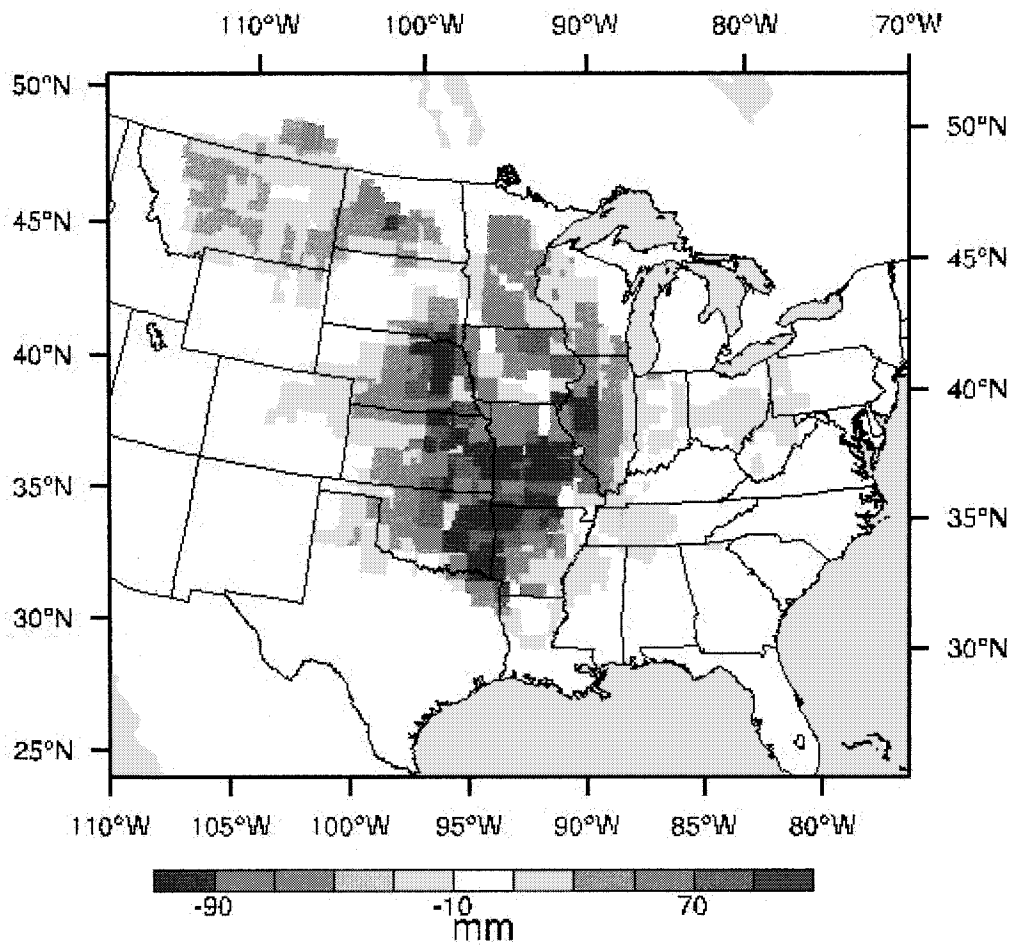modeled groundwater.

Figure 7: A10 minus OL difference in accumulated evapotranspiration (mm) for 2003.

Tables

Table 1: Evaluation of groundwater estimates from model and assimilation integrations against measured groundwater. Correlation

coefficient r and RMSE (mm) are calculated with respect to five day average groundwater storage based on observations from

58 piezometers. Skill improvement through assimilation is calculated relative to the open loop simulation ("skill" = 1 –

RMS_{Assim}/RMS_{OL}). Bold fonts (italics) indicate a significant increase (decrease) in r relative to OL at the 5 % level.

	Mississippi			Ohio-Tennessee			Upper Mississippi			Red-Ark/Lower MS			Missouri		
	r	RMSE	skill	r	RMSE	skill	r	RMSE	skill	r	RMSE	skill	r	RMSE	skill
OL	0.42	31.2	--	0.83	69.6	--	0.13	49.0	--	0.47	45.1	--	0.44	23.8	--
A20	0.53	25.4	0.2	0.79	45.7	0.3	0.16	45.1	0.1	0.55	37.1	0.2	**0.63**	20.5	0.1
A10	**0.68**	21.5	0.3	*0.75*	38.0	0.5	0.17	44.9	0.1	**0.61**	33.3	0.3	**0.77**	17.0	0.3
A1	**0.73**	25.6	0.2	*0.69*	39.9	0.4	0.21	48.6	0.0	0.57	37.2	0.2	**0.72**	25.4	-0.1

Table 2: Mean runoff (mm yr^{-1}) and coefficient of linear correlation against monthly average gauged river flow. Gauged runoff is derived from measurements at USGS and Army Corps of Engineers gauging stations, and correlations were calculated with a 1-month lag to account for channel flow time. Correlations were not calculated for the Red-Arkansas/Lower-Mississippi Basin because of upstream contributions to flow.

	Mississippi		Ohio-Tennessee		Upper Mississippi		Red-Ark/LMS		Missouri	
	mean	r	mean	r	mean	r	mean	r*	mean	r
OL	30.6	0.72	108.2	0.50	5.4	0.59	44.0	--	1.9	0.52
A20	29.7	0.73	101.4	0.53	5.7	0.67	44.4	--	2.2	0.57
A10	31.2	0.73	105.8	0.51	6.5	0.70	45.6	--	2.8	0.59
A1	35.9	0.67	120.8	0.41	9.3	0.53	49.5	--	4.9	0.57

Table 3: Average and range of ET (mm day^{-1}) for January 2003 to May 2006. Range refers to the difference between maximum and minimum monthly ET over the length of the simulation.

	MS		OH		UP		LR		MO	
	mean	range	mean	range	mean	range	mean	range	mean	range
OL	1.82	3.38	2.85	4.76	1.97	3.76	2.07	3.28	1.23	3.28
A20	1.87	3.54	2.90	4.71	2.01	3.92	2.12	3.35	1.27	3.39
A10	1.90	3.66	2.90	4.67	2.04	4.21	2.17	3.51	1.32	3.40
A1	1.94	3.73	2.91	4.80	2.05	4.29	2.18	3.77	1.38	3.35

Table 4: Coefficient of linear correlation between CLSM simulated runoff and gauged river flow for eight small watersheds in the Mississippi River Basin (locations mapped on Figure 1). Correlations are calculated using monthly data for January 2003 through May 2006. No lag was applied for channel flow time for these relatively small basins. Italics indicate significant reduction in correlation relative to open loop simulation at the 5% level.

	Kanawha	Wabash	Illinois	Minnesota	Canadian	Ouachita	Yellowstone	Kansas
OL	0.64	0.62	0.72	0.61	0.76	0.44	0.27	0.54
A20	0.65	0.72	0.73	0.64	0.77	0.41	0.34	0.63
A10	0.66	0.65	0.68	0.64	0.69	0.35	0.34	0.53
A1	0.65	0.55	0.63	0.72	*0.49*	0.29	0.17	*0.02*

Table 5: Evaluation of TWS estimates for Illinois. Correlation coefficient r and RMSE

(mm) are calculated against area-averaged monthly TWS based on *in situ* data,

January 2003 – December 2005. Skill and bold face as in Table 1.

	r	RMSE	skill
OL	0.68	30.2	--
A20	0.68	26.1	0.13
A10	0.74	22.9	0.24
A1	**0.84**	19.5	0.35