ARTICLE IN PRESS

Journal of Hydrology xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Evaluation of a model-based groundwater drought indicator in the conterminous U.S.

Bailing Li [a,b,*], Matthew Rodell [b]

[a] Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, United States
[b] Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, United States

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Groundwater drought indicator
Groundwater persistence
GRACE data assimilation
Groundwater temporal variability

SUMMARY

Monitoring groundwater drought using land surface models is a valuable alternative given the current lack of systematic in situ measurements at continental and global scales and the low resolution of current remote sensing based groundwater data. However, uncertainties inherent to land surface models may impede drought detection, and thus should be assessed using independent data sources. In this study, we evaluated a groundwater drought index (GWI) derived from monthly groundwater storage output from the Catchment Land Surface Model (CLSM) using a GWI similarly derived from in situ groundwater observations. Groundwater observations were obtained from unconfined or semi-confined aquifers in eight regions of the central and northeastern U.S. Regional average GWI derived from CLSM exhibited strong correlation with that from observation wells, with correlation coefficients between 0.43 and 0.92. GWI from both in situ data and CLSM was generally better correlated with the Standard Precipitation Index (SPI) at 12 and 24 month timescales than at shorter timescales, but it varied depending on climate conditions. The correlation between CLSM derived GWI and SPI generally decreases with increasing depth to the water table, which in turn depends on both bedrock depth (a CLSM parameter) and mean annual precipitation. The persistence of CLSM derived GWI is spatially varied and again shows a strong influence of depth to groundwater. CLSM derived GWI generally persists longer than GWI derived from in situ data, due at least in part to the inability of coarse model inputs to capture high frequency meteorological variability at local scales. The study also showed that groundwater can have a significant impact on soil moisture persistence where the water table is shallow. Soil moisture persistence was estimated to be longer in the eastern U.S. than in the west, in contrast to previous findings that were based on models that did not represent groundwater. Assimilation of terrestrial water storage data from the Gravity Recovery and Climate Experiment (GRACE) satellite mission improved the correlation between CLSM based regional average GWI and that based on in situ data in six of the eight regions. Practical issues regarding the application of GRACE assimilated groundwater storage for drought detection are discussed. An important conclusion of this study is that model parameters that control the depth to the water table, including bedrock depth, strongly influence the evolution and persistence of simulated groundwater and require careful configuration for drought monitoring.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Drought is a natural hazard that has a broad range of social and economic impacts, from decreased agricultural productivity to restrictions on residential water use. Long lasting drought or frequent severe drought events in arid and semi-arid regions can lead to even more devastating consequences such as inadequate food supply and desertification (Mishra and Singh, 2010). Monitoring systems capable of detecting and mapping drought over large spatial scales and with temporal continuity are essential for assessing drought severity and extent, and for mitigating its impacts.

Droughts are generally initiated by below-normal precipitation over a period of weeks or longer, and over time they can propagate through different components of the hydrological cycle including groundwater (Changnon, 1987). Groundwater drought, which is a distinct class of drought, not a sub-class of meteorological, agricultural, or hydrological drought (Mishra and Singh, 2010), has profound impacts on ecosystems and on water supply for irrigation and municipal use in regions where surface water stores are

* Corresponding author at: Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, United States. Tel.: +1 301 286 6020; fax: +1 301 614 5808.

E-mail address: Bailing.li@nasa.gov (B. Li).

http://dx.doi.org/10.1016/j.jhydrol.2014.09.027
0022-1694/© 2014 Elsevier B.V. All rights reserved.

Please cite this article in press as: Li, B., Rodell, M. Evaluation of a model-based groundwater drought indicator in the conterminous U.S. J. Hydrol. (2014), http://dx.doi.org/10.1016/j.jhydrol.2014.09.027

inadequate to meet demand. Groundwater's response to hydrometeorological inputs is dampened by the processes of recharge and discharge, which act as a low-pass filter, with significant, persistent events and seasonality retained (Eltahir and Yeh, 1999). The unique temporal characteristics of groundwater storage directly influence drought evolution in related processes. Peters et al. (2005) modeled drought propagation from groundwater recharge to groundwater discharge and found that the groundwater system ignored the occurrence of less intensive drought events but exhibited protracted periods of recovery from severe droughts. Improved monitoring and understanding of groundwater drought and its relationship with other types of drought have the potential to improve groundwater management as well as drought forecasting.

Due to a deficiency of in situ observations, operational drought monitoring activities rely on subsurface wetness information from land surface models (LSM) driven by observation based meteorological forcing fields (e.g., Mo, 2008), in some cases from assimilating satellite data (Houborg et al., 2012). Land surface models can provide continuous and consistent fields of land surface states (e.g., soil moisture) but they are limited by imperfect model physics and uncertainties in parameters and forcing fields. As a result, modeled soil moisture and associated drought indices derived from different combinations of LSMs and forcing fields can exhibit significant variability, which complicates accurate drought quantification (Mo, 2008; Sheffield et al., 2012; Xia et al., 2014a, 2014b). Simplification of often complex hydrogeology and physics related to parameterizations of groundwater, in those models that simulate groundwater at all (e.g., Niu et al., 2007; Koster et al., 2000), leads to additional uncertainty in modeled groundwater and associated drought indices.

Launched in 2002, the NASA/German Space Agency's Gravity Recovery and Climate Experiment (GRACE; Tapley et al., 2004) maps Earth's gravity field with enough accuracy to infer monthly changes in terrestrial water storage (TWS), which includes soil moisture, groundwater, snow, and surface waters, with a maximum spatial resolution of about 150,000 km^2 at mid-latitudes (Rowlands et al., 2005; Swenson et al., 2006). GRACE derived TWS has been used to estimate declines of groundwater storage (Rodell et al., 2009; Wada et al., 2010; Famiglietti et al., 2011; Feng et al., 2013; Voss et al., 2013), and its lows are strongly correlated with drought events (Andersen et al., 2005; Leblanc et al., 2009; Li et al., 2012; Thomas et al., 2014). In recent years, assimilation of GRACE TWS into land surface models has been demonstrated as an effective means of disaggregating GRACE TWS vertically, horizontally, and temporally while improving modeled fluxes and states (Zaitchik et al., 2008; Su et al., 2010). This capability has enabled the application of GRACE observations for operational drought monitoring, which requires timeliness and high spatial and temporal resolutions not achievable by GRACE alone (Houborg et al., 2012; Rodell, 2012).

The objectives of this study were to evaluate a groundwater drought index derived from groundwater storage simulated by the Catchment Land Surface Model (CLSM; Koster et al., 2000) using in situ groundwater observations, and to characterize the temporal variability of the groundwater drought indicator including its persistence and its relationship with precipitation anomalies. In situ groundwater data records spanning 10–30 years were gathered for eight semi-humid to humid regions in the continental U.S. with varying hydrogeological properties. Groundwater drought indices were created based on anomalies (relative to seasonal mean) of monthly groundwater storage estimates from CLSM and in situ measurements which were also standardized. The groundwater drought indices were analyzed alongside the Standardized Precipitation Index and a soil moisture drought index in order to examine the relationships among these indices and the types of droughts they represent. GRACE TWS data (derived from University of Texas, Center for Space Research, release 5 dataset; Swenson and Wahr, 2006; Landerer and Swenson, 2012) was assimilated into CLSM (as in Zaitchik et al., 2008) and the resulting groundwater fields were evaluated using the in situ data. Implications for the application of GRACE assimilated groundwater storage for drought monitoring are discussed.

2. In situ data and model estimates

Fig. 1 shows the locations of groundwater observation wells located in Long Island (New York), New Jersey, Massachusetts, Pennsylvania and four sub-basins of the Mississippi River basin: the Upper-Mississippi, Ohio–Tennessee, Missouri and the combined Red River and Lower Mississippi (hereafter referred to as Red–LM). Depth-to-water measurements were obtained from the USGS Office of Groundwater and the Illinois State Water Survey. Criteria that were applied in selecting these observation wells for this study included (1) they were determined to represent groundwater levels in unconfined or semi-confined aquifers, (2) they were not directly impacted by pumping or injections, and (3) the measurements were frequent enough to capture the seasonal cycle and generally were continuous over a period of ten or more years. To produce a monthly time series, an average value was used when multiple measurements were made in a given month. We converted depth to water table measurements to water level anomalies (deviations from the temporal mean) by taking the additive inverse of the measurements and subtracting the long term mean at each well. We then derived groundwater storage anomalies by multiplying the water level anomalies by the specific yield. Following Rodell et al. (2007), the specific yield was determined for each location based on published studies in which it was estimated for the same aquifer through field experimentation and/or numerical modeling. When no aquifer-specific estimates could be found, a specific yield value was assigned based on the range of values for the geologic material as reported by Johnson (1967) and any other available well-specific information, including the depth-to-water variability itself. Missing monthly groundwater storage anomaly values were filled using linear interpolation. Table 1 presents the periods of the resulting data records (some wells may have shorter records) for each region, the area, the number of wells, and the averaged (over the well locations only) precipitation and well properties.

CLSM simulates subsurface water storage changes within natural hydrological catchments instead of on regular grids (Koster et al., 2000). Within each catchment three subsurface state variables, surface excess, root zone excess and catchment deficit, simulate water storage changes at different vertical depths based on water and energy balance equations. CLSM does not explicitly simulate the groundwater table, but the catchment deficit variable, which is defined as the amount of water, per unit area, needed to fill a catchment to capacity, reflects changes in the shallow unconfined aquifer. Groundwater storage can be derived from the catchment deficit and the maximum capacity for water of each catchment, which is determined by a bedrock depth parameter and soil porosity. Following Houborg et al. (2012), CLSM bedrock depths were increased by 2 m uniformly everywhere so that the dynamic range of simulated TWS would be better aligned with that of GRACE derived TWS, particularly during dry periods. More information on calculating groundwater storage from CLSM catchment deficit can be found in Zaitchik et al. (2008) and Li et al. (2012).

CLSM was forced using the Princeton meteorological dataset (Sheffield et al., 2006), which provides more than 60 years (1948–2010) of 3-hourly, 1° gridded fields of precipitation, solar radiation, wind speed, surface pressure, surface air temperature, and relative humidity. The precipitation data was derived from

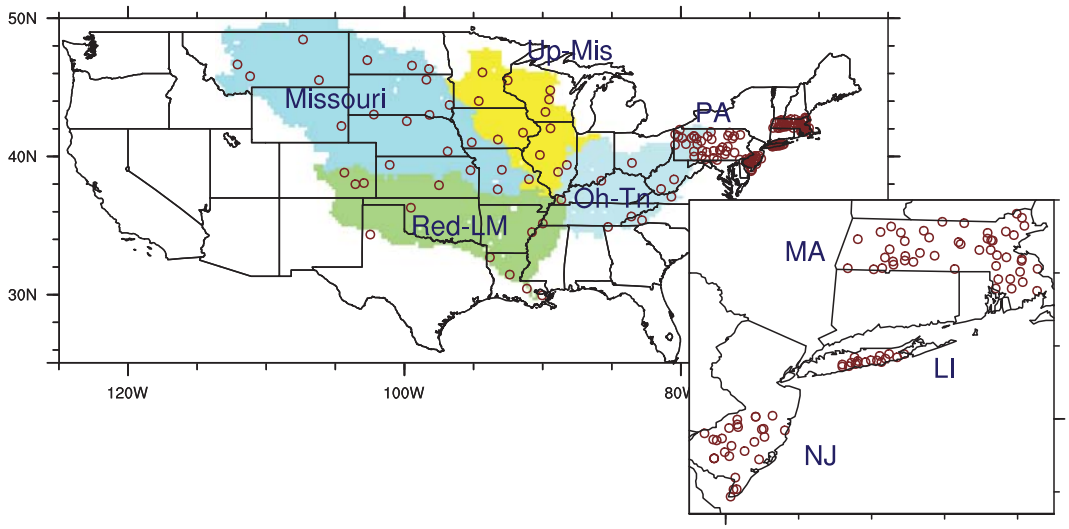

Fig. 1. Locations of groundwater observation wells in Long Island ("LI"), New Jersey ("NJ"), Massachusetts ("MA"), Pennsylvania ("PA") and the four sub-basins of the Mississippi River basin, the Upper Mississippi ("Up-Mis"), the Ohio–Tennessee ("Oh–Tn"), the combined Red River and Lower Mississippi ("Red–LM"), and the Missouri.

Table 1
Data period of in situ groundwater observations, number of groundwater wells, area, average specific yield ($\bar{S}_y$), average well depth ($\bar{d}_{well}$), average depth to water table ($\bar{d}_{gw}$) and average (over well locations only) annual precipitation ($\bar{P}$) for 1948–2010.

ID	Region	Data period	# of Wells	Area (km^2)	$\bar{S}_y$	$\bar{d}_{well}$ (m)	$\bar{d}_{gw}$ (m)	$\bar{P}$ (mm)
1	Long Island	1992–2011	17	2000	0.26	15	8	1147
2	New Jersey	2002–2011	27	14,200	0.17	27	6	1221
3	Massachusetts	1980–2011	48	28,400	0.20	9	4	1165
4	Pennsylvania	2002–2011	35	102,900	0.07	42	10	1045
5	Upper Mississippi	1980–2010	13	491,800	0.17	19	6	849
6	Ohio–Tennessee	1980–2010	10	528,100	0.09	38	7	1200
7	Red–LM	1980–2010	13	903,900	0.16	86	16	970
8	Missouri	1980–2010	19	1,324,000	0.14	30	10	576

ground based observations which were then temporally and spatially disaggregated using statistics derived from radar images. Other fields were based on the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis. These were bias corrected using ground-based observations. The 63 year CLSM simulation driven by the Princeton forcing data was used to generate climatologies of the modeled states.

Because the Princeton dataset ends in 2010, North American Land Data Assimilation System (NLDAS-2) forcing fields (Xia et al., 2012) were used to drive CLSM to present (they are also used in producing near-real time drought indices). NLDAS-2 covers the conterminous U.S. and part of Canada and Mexico, and the fields are posted on a 0.125° grid with an hourly time step. NLDAS-2 precipitation is based on daily precipitation measurements at over 10,000 gauges which are then temporally disaggregated and gap filled using radar images (Cosgrove et al., 2003). Other fields are based on NCEP's North American Regional Analysis. Bias between different forcing data sets can cause discontinuities in model output and impair drought detection. Therefore we bias-corrected the NLDAS-2 forcing fields to the corresponding fields in the Princeton data set before applying them for model simulation.

The assimilation of GRACE derived TWS anomalies into CLSM was conducted using an ensemble Kalman smoother (Zaitchik et al., 2008). Gridded GRACE TWS anomalies (Landerer and Swenson, 2012) were first aggregated into larger natural river basins or combined basins to construct basin-scale GRACE TWS anomaly time series (see Houborg et al., 2012, for basin delineation). To convert the GRACE TWS anomalies to values compatible with CLSM simulated TWS, the temporal mean of TWS (see Zaitchik et al., 2008, for calculation of CLSM TWS) from an open loop (no data assimilation) CLSM simulation was calculated for each basin and added to the basin-scale GRACE TWS anomalies. Further, the assimilation was conducted in two iterations because GRACE provides monthly means as opposed to instantaneous observations. First, the model was propagated forward from the beginning of each month to the end of the month to obtain a monthly TWS forecast, from which the innovation (the difference between the modeled and assimilated estimates) was calculated. In the second iteration, ensemble updates derived from ensemble statistics and monthly innovations were applied to each daily state for each ensemble member and fluxes were re-integrated based on analysis and forcing fields. Both the open loop and GRACE data assimilation simulations were driven by biased-corrected NLDAS-2 forcing fields.

3. Drought indices

The Standardized Precipitation Index (SPI) was employed to indicate precipitation anomalies over different timescales. SPI is expressed as standard deviations from a long term mean of a normal distribution and is convenient for describing drought conditions and for comparing drought severity across different regions (McKee et al., 1993). In addition, SPI can be computed on different timescales to examine drought temporal variability. All SPI values used in this study were computed using Princeton precipitation from 1948 to 2010.

Wetness percentiles based on subsurface state data are frequently employed as indicators of drought and can be easily

derived given a sufficiently long data record (Mo, 2008; Houborg et al., 2012). The seasonality associated with subsurface states may lead to non-Gaussian behaviors which may prevent accurate identification of drought. This issue can be avoided by deriving a separate climatology for each calendar month (Mo, 2008). Since some of our in situ groundwater data records contain only 10 years of data, which is too short to form a month-specific climatology, the groundwater drought index (hereafter referred to as GWI) used in this study was constructed using the following procedures. Because shallow groundwater storage exhibits strong seasonality (e.g., Eltahir and Yeh, 1999), we first removed the seasonal cycle from the monthly time series of in situ and CLSM based groundwater storage at each location. We then standardized the de-seasonalized anomalies by subtracting the temporal mean and dividing by the standard deviation to obtain GWI values for each data type. Standardization, which is also used in calculating SPI, makes average indices more representative of regional behaviors than averaged groundwater storage anomalies, which can be skewed by data at a few locations with larger dynamic ranges. Note that the choice of specific yield affects in situ groundwater storage estimates but has no impact on the associated GWI because of the standardization procedure. A soil moisture drought index (hereafter referred to as SMI) was also generated from CLSM estimated monthly root zone moisture using the same procedure.

4. Results

Statistics for GWI derived from the Princeton forced CLSM simulation (1948–2010) are presented first to provide general characteristics of groundwater drought. Results from NLDAS forced simulations (2003-2011) are presented next with a focus on evaluating the impact of GRACE data assimilation on modeled groundwater storage. In both cases the model simulations are evaluated using in situ groundwater data. Finally, a method for reconciling the Princeton-forced and NLDAS-forced simulations is presented.

4.1. GWI based on Princeton forced simulation

Fig. 2 shows region averaged (over well locations only) GWI time series derived from in situ groundwater observations and from the Princeton-forced CLSM simulation (among others), in comparison with the 12 month SPI (SPI12). In general, both in situ and Princeton-forced CLSM GWI represent major drought events indicated by SPI12 with similar degrees of severity. For instance, GWI and SPI12 identify the severe drought that occurred in the late 1980s in the Upper Mississippi and Ohio–Tennessee basins, with index values near −2. SPI12 indicates that the same drought affected the Missouri River basin until relief came in 1989, while GWI, based on both the well data and CLSM, suggests that the effects on groundwater lingered into the early 1990s. These lagged effects likely reflect the requirement for surface and shallow subsurface water stores to be replenished before groundwater recharge returns to normal, and the fact that the sustained above-normal precipitation needed for recharge to accelerate in a drier region (see Table 1 for mean annual precipitation) was not seen until 1993. Note that GWI and SPI12 in large regions such as the Mississippi sub-basins have smaller dynamic ranges than those in smaller regions, due to the effect of spatial averaging.

On regional average, GWI derived from CLSM correlates well with GWI from groundwater well data with the strongest correlation observed in New Jersey and Upper-Mississippi and the lowest in Long Island and Red–LM (Table 2). The deeper aquifers in Red–LM (see well depths in Table 1) are not well represented by the simple groundwater formulation in CLSM and respond differently to atmospheric forcing than shallow aquifers do. The small size of Long Island may have contributed to the low correlation because discrepancies between the large scale Princeton forcing and what occurred at individual well sites were not averaged out as much as they would have been in a larger area. Table 2 also lists coefficients of correlation between GWI based on CLSM and that based on in situ data computed at individual well locations and averaged for each region. These correlation coefficients are significantly lower, suggesting that model estimates and their associated drought indices are more reliable at regional scales than at local scales.

SPI is based on the accumulation of precipitation over a specified timescale. Thus, the correlations between GWI and SPI computed for different timescales reflect the response rate of groundwater storage to precipitation inputs. For instance, strong correlation between GWI and 1-month SPI would suggest that changes in groundwater levels are tightly coupled to short term precipitation variability. Fig. 3 shows the correlation between CLSM based GWI and SPI at scales of 6, 12 and 24 months. Mean annual precipitation, CLSM bedrock depths and correlation between CLSM derived SMI and SPI are also presented for comparison. In general, GWI exhibits stronger correlation with SPI12 and SPI24 than with SPI6, reflecting the slow or lagged response of groundwater to surface wet and dry events. The GWI–SPI correlations are strongly influenced by the depth to the water table. In general, deeper water tables tend to be correlated with deeper bedrock and lower annual precipitation, as is the case for the area of the Great Plains that stands out in the two top right panels of Fig. 3. Deeper water tables in the Great Plains cause more attenuation of high frequency atmospheric events and allow multi-year cycles of groundwater variability to appear as seen in Fig. 2 for Red–LM and Missouri. Not surprisingly, these are better correlated with SPI of longer timescales. On the other hand, the shallower groundwater in the Tennessee–Alabama area responds more quickly to atmospheric conditions and is dominated by high frequency variability (as seen in Fig. 2 for Ohio–Tenn) which is better correlated with shorter timescale SPI.

Despite deep bedrock in the Pacific Northwest, GWI is better correlated with SPI6 than with SPI12 and SPI24, because the high rate of precipitation in the region sustains shallow water tables in CLSM. Adjustments to model parameters such as those related to runoff could cause changes in mean depth to groundwater in CLSM, which would affect temporal variability of groundwater (and GWI) and hence the detection of drought. Thus comparisons such as this one between modeled and observed timescales of groundwater response to various climatic conditions could be a new useful form of model calibration for those models that simulate groundwater.

The SMI based on CLSM root zone soil moisture generally shows stronger correlation with SPI in the eastern U.S. than in the west at the scales examined here, reflecting stronger dependency of soil moisture on precipitation in wetter environments. In addition, the correlation (including spatial patterns and magnitudes) becomes increasingly similar to the correlation between GWI and SPI in the east as the timescale increases, suggesting that the long term temporal variability of soil moisture is controlled in part by shallow groundwater. In the west, SMI correlates more strongly with SPI6 than with SPI12 and SPI24. In dry western regions, this can be explained by soils that dry quickly between precipitation events without regard to long term precipitation totals. In addition, the influence of groundwater on soil moisture, which would have provided longer scale temporal variability and thus stronger correlation with longer term SPI, is also weak in dry climates due to deeper water tables. In the wet Pacific Northwest, the influence of a shallow water table on the soil wetness profile and the lack of drying between individual rainfall events made the correlation of SMI–SPI similar to that of GWI–SPI. SMI exhibits the lowest

Please cite this article in press as: Li, B., Rodell, M. Evaluation of a model-based groundwater drought indicator in the conterminous U.S. J. Hydrol. (2014), http://dx.doi.org/10.1016/j.jhydrol.2014.09.027

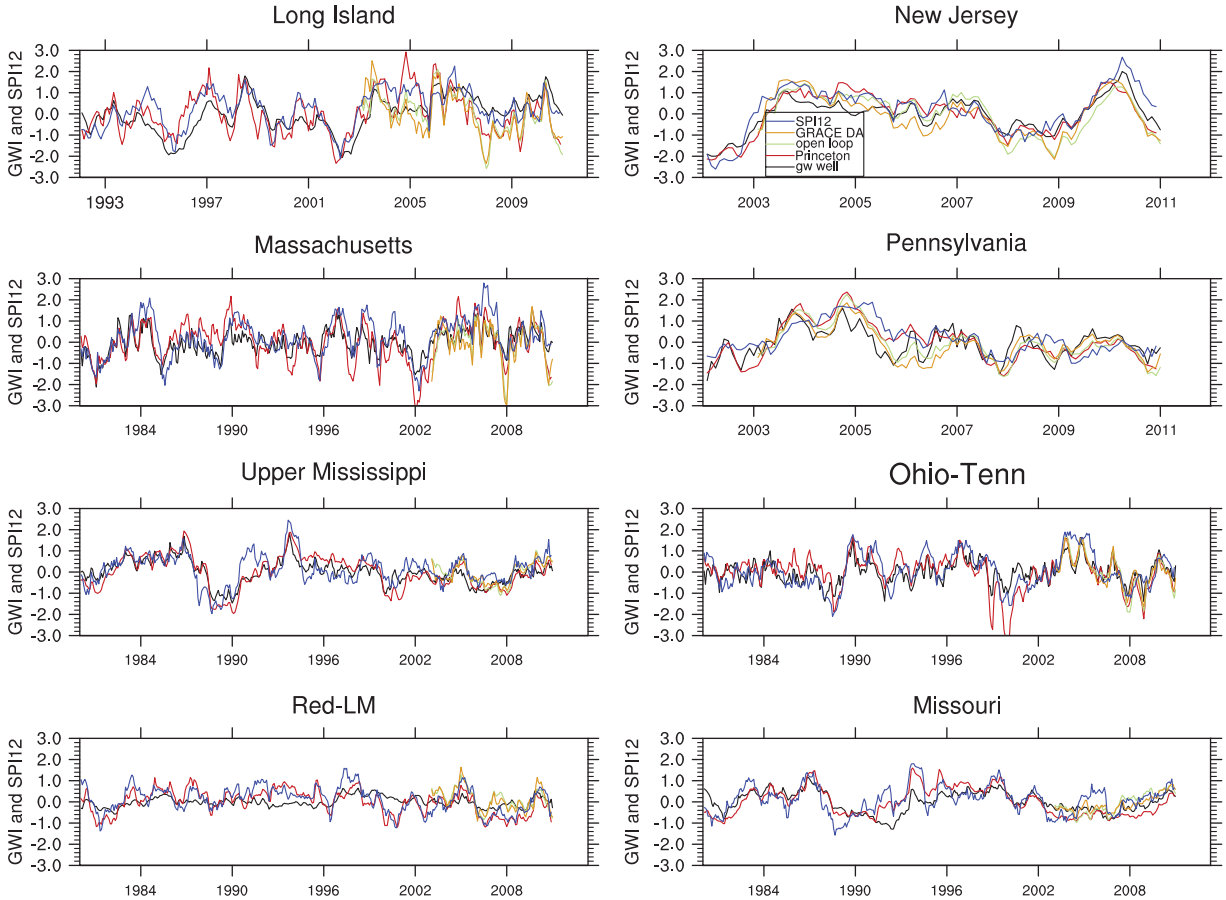

Fig. 2. Region-averaged GWI time series for in situ groundwater data (in black), the Princeton forced simulation (red), the NLDAS-2 forced open loop (in green), and the NLDAS-2 forced GRACE data assimilation (in orange), in comparison with SPI12 (in blue).

Table 2
Correlation coefficients between region-averaged GWI from the Princeton forced CLSM simulation and from in situ groundwater observations. Numbers in parentheses represent regional averages of correlation coefficients at individual well locations between the two sets of GWI. All correlations were calculated for the common period of in situ data (listed in Table 1) and Princeton forced CLSM simulation (1948–2010).

Long Island	New Jersey	Massachusetts	Pennsylvania	Upper Mississippi	Ohio–Tennessee	Red–LM	Missouri
0.58 (0.42)	0.87 (0.67)	0.71 (0.38)	0.75 (0.52)	0.90 (0.59)	0.69 (0.38)	0.43 (0.17)	0.74 (0.49)

correlation with SPI24 (compared to that with SPI6 and SPI12) in the Great Plains and the southwestern U.S., in contrast to the behavior of GWI, which reflects large inter-annual variability of precipitation and the minimal influence of (deeper) groundwater on soil moisture.

The correlation between CLSM-based GWI and SPI can be evaluated using the correlation between GWI derived from in situ data and SPI. Fig. 4 compares regional average (over individual well locations) correlation between SPI and GWI based on both in situ observations and CLSM output (from three different simulations), for each of the eight regions. GWI based on CLSM-Princeton exhibits the highest correlation with SPI, which is mainly due to the fact that SPI was also derived from the Princeton precipitation data. Simplified model physics may also cause model estimates to be strongly correlated with the precipitation input. However, Fig. 4 shows that the timescale of maximum correlation for both modeled and in situ GWI is similar in each region. In addition, the relationship between the correlation and scales takes similar form, in particular, the quick peak followed by a gradual decline in the wetter regions (the four northeastern regions and Ohio–Tennessee), suggesting diminishing impact of precipitation variability on groundwater after about a year; and a logarithmic function for the two drier regions, the Upper-Mississippi and Missouri, suggesting sustained impact of long term precipitation variability on groundwater. These results suggest that the model, to a certain degree, properly represented the correlation between in situ groundwater and precipitation in the study regions.

The persistence of drought can be quantified using the characteristic time, a weighted sum of autocorrelations, R, at different lags (Mo and Schemm, 2008):

$$T_0 = 1 + 2\sum_{i=1}^{N}(1 - i/N)R(i) \quad (1)$$

where i is the ith lag of the total number of lags (month), N, which is 30 for all results. Fig. 5 shows T_0 of GWI based on CLSM versus T_0 of GWI based on groundwater observations at individual well locations. Since auto-correlation is sensitive to data lengths, locations where the in situ data record is shorter than 120 months (4 times the maximum lag) were excluded from this graph. CLSM derived

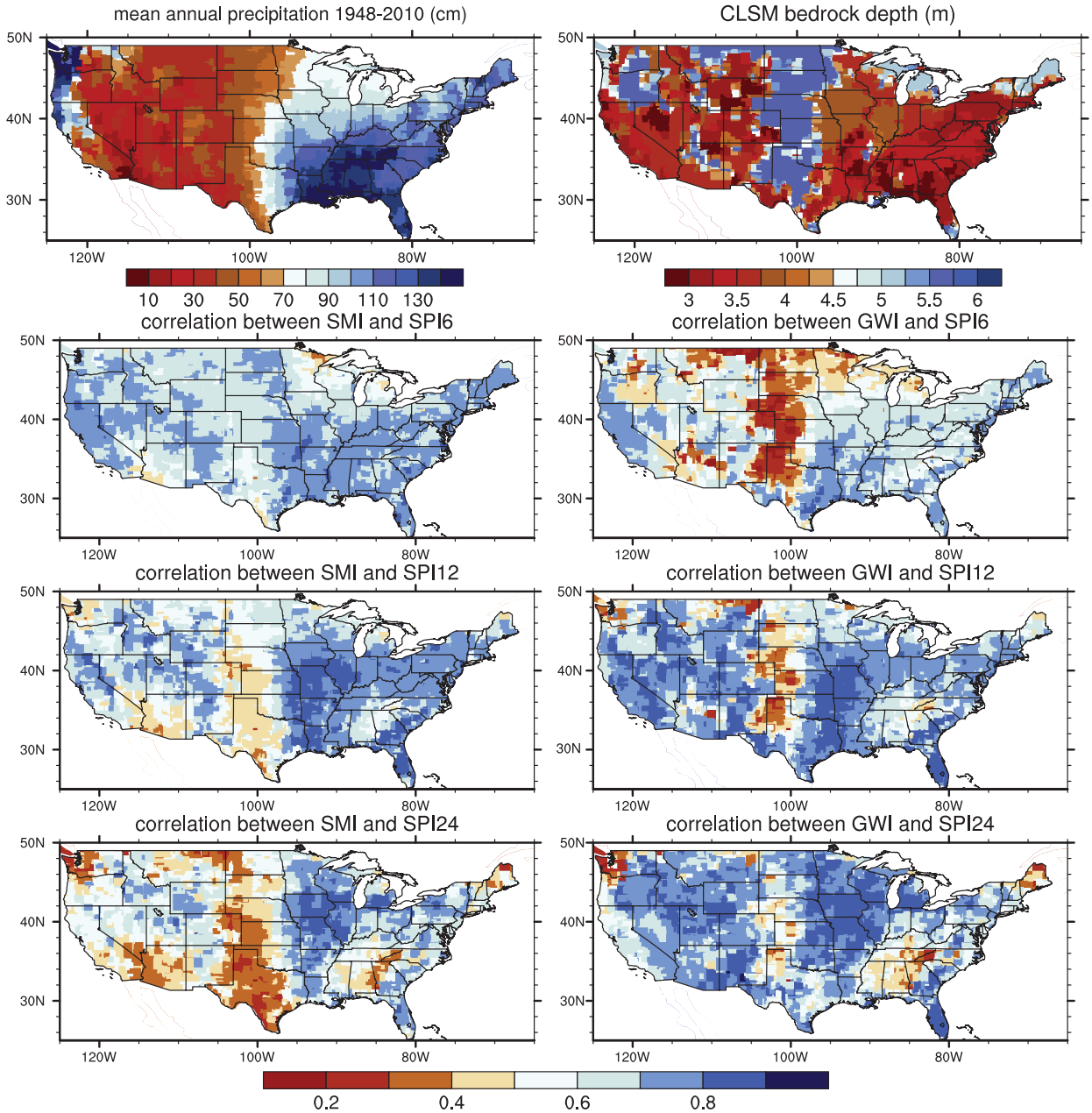

Fig. 3. Mean annual precipitation (top left panel), CLSM bedrock depth (top right panel), correlations between the CLSM derived SMI and SPI at the 6, 12 and 24 month timescale (lower left three panels) and correlations between the CLSM derived GWI and SPI (lower right three panels). SPI, SMI, and GWI were derived from Princeton precipitation and the Princeton-forced CLSM simulation for 1948–2010.

GWI generally exhibits longer characteristic time than observation based GWI. The spatial and temporal interpolation required to generate forcing data and the use of climatology based data sets such as vegetation greenness likely caused groundwater and soil moisture to vary more smoothly, leading to stronger auto-correlation. Nevertheless, regional average characteristic time, represented by the *x*–*y* coordinates of number symbols (representing region id; see Table 1), based on CLSM agrees well with that based on observations. It can be seen in Fig. 5 that groundwater in the wetter regions has shorter characteristic times than that in drier regions. In situ groundwater data in the Red–LM basin show stronger persistence than model-based groundwater does, likely due to the deeper water table, which, as previously noted, attenuates the high frequency events and allows groundwater to develop stronger long term variability and thus stronger auto-correlations at longer timescales. In Pennsylvania, auto-correlation of observed groundwater declines quickly with increasing time lags (not shown). This may be related to more complex hydrogeology that is not represented in CLSM, as many of the wells are sited in semi-confined aquifers.

Across the continental U.S., the characteristic time of CLSM based GWI exhibits larger spatial variability than that of other drought indices (Fig. 6). The spatial pattern of groundwater characteristic time does not resemble that of SPI6 or SPI12. In fact, in the Tennessee–Alabama and Northwest coastal areas where SPI6 and SPI12 show the longest persistence, groundwater exhibits the shortest persistence. The shallow groundwater tables in these areas cause groundwater to respond quickly to precipitation. In the Great Plains, where the water table is relatively deep and precipitation rates are lower, the shallow aquifers are mainly recharged during infrequent episodes of significant precipitation,

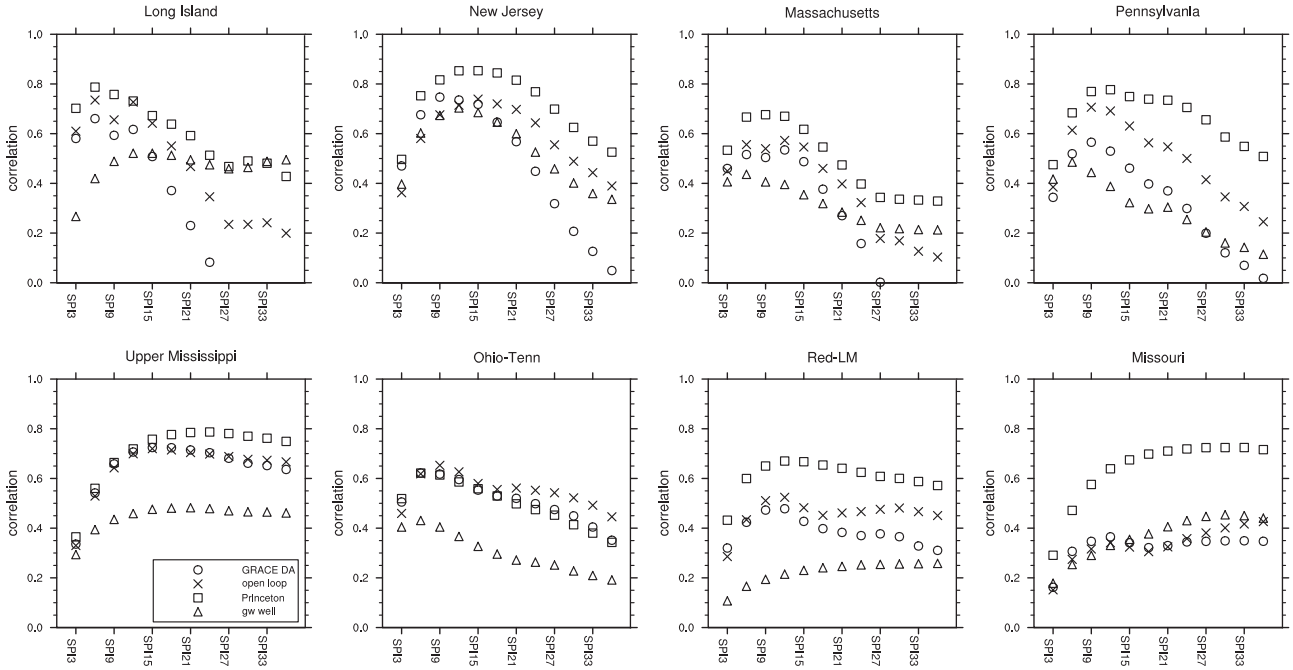

Fig. 4. Regional average (over well locations) correlation coefficients between SPI and GWI of in situ groundwater ("gw well"), Princeton forced CLSM estimates, the open loop ("OL"), and GRACE data assimilation ("GRACE DA").

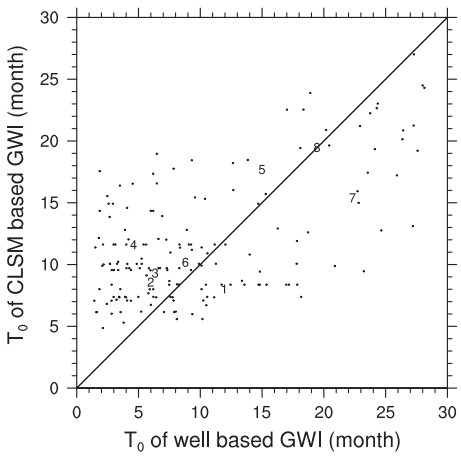

Fig. 5. The characteristic time (T_0) of GWI based on Princeton-forced CLSM output versus that based on in situ groundwater data at individual wells. Regional average characteristic time is represented by the x and y coordinates of number symbols (representing region id). CLSM estimates from the same period of in situ groundwater were used in calculating auto-correlation, and locations with less than 120 months of in situ data were excluded.

and natural discharge happens more slowly in response to increased water levels, hence groundwater anomalies tend to persist. These results suggest that depth to groundwater is a key factor in determining groundwater persistence and its correlation with SPI. Thus model parameters and physics that influence the water table depth should be set with care if modeled groundwater is to be used as an indicator of drought.

Persistence of SMI exhibits spatial patterns similar to those of GWI in the eastern half of the U.S., reflecting the strong influence of groundwater on soil moisture. Soil moisture generally shows greater persistence in the east than in the west, which is opposite to the findings by Sheffield et al. (2012) and Mo and Schemm (2008), who relied on soil moisture estimates from models without any groundwater component. This result suggests that shallow groundwater can significantly affect soil moisture and thus its persistence in wet climates. In the Mountain West, soil moisture exhibits longer persistence than the surrounding areas, similar to that of precipitation. Reduced evapotranspiration at high elevations and seasonal snow's dominance of the water cycle result in longer persistence of soil moisture. In the Great Plains, soil moisture exhibits the lowest persistence due to high rates of evapotranspiration and the lack of groundwater influences.

4.2. GWI based on NLDAS-forced simulations

As indicated earlier, the NLDAS forcing fields, bias corrected to match the Princeton dataset, were used for the open loop and GRACE data assimilation simulations which are also plotted in Fig. 2. GWI from these two simulations fluctuates more than Princeton based GWI, due to the higher spatial and temporal resolutions of NLDAS. As shown in Table 3 (2nd and 3rd rows), these higher quality forcing fields improved the correlation between CLSM based GWI and observation based GWI in the four northeast regions, but they had little impact on the four Mississippi subbasins. It is possible that the benefits of higher resolution forcing fields would have been more apparent had the in situ observations been more densely spaced.

More importantly, Table 3 (3rd and 4th rows) shows that GRACE data assimilation improved the correlation between regional average model-based GWI and GWI based on in situ data in all regions except Upper Mississippi and Long Island. Long Island is too small, narrow, and close to the ocean for GRACE to measure effectively. Similarly, due to the coarse resolution of GRACE observations, GRACE data assimilation did not consistently improve the correlation between CLSM estimates and in situ data at individual well locations (not shown).

Fig. 4 shows that the timescale of maximum correlation between the GWI of NLDAS-forced simulations and SPI is similar to what has been discussed earlier. In general, GRACE data assimilation lowered the correlation with SPI of longer scales while exerting minimal impacts on the correlation with SPI of shorter

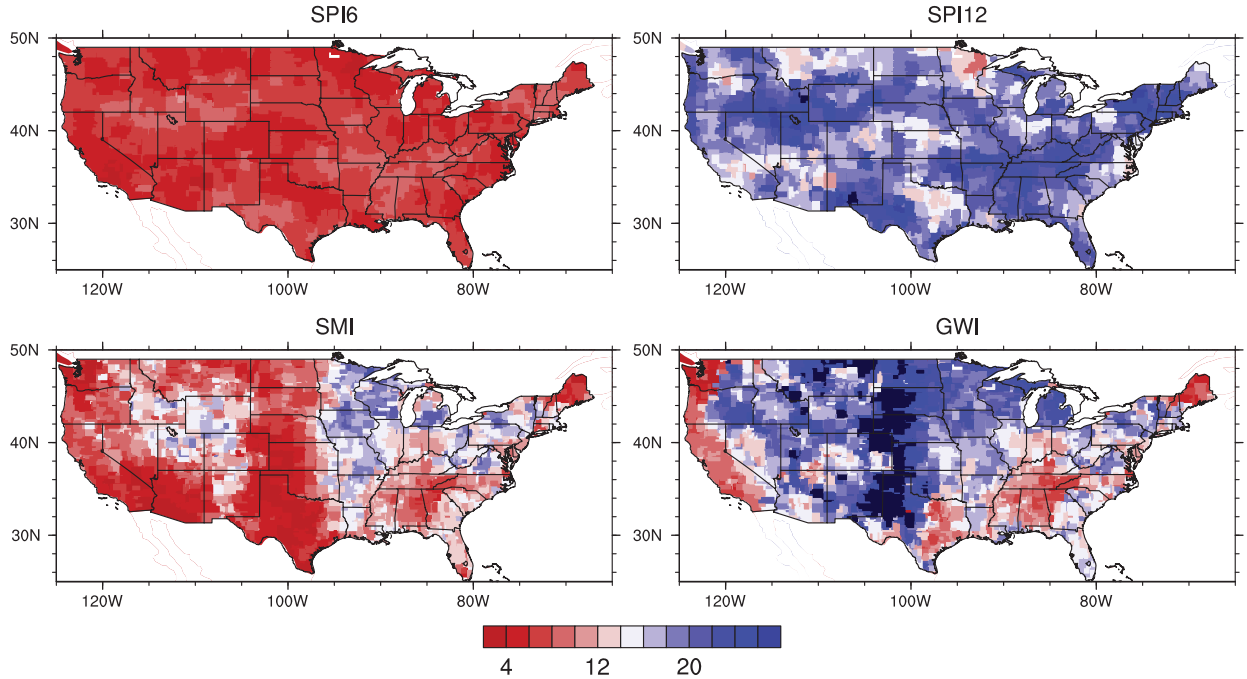

Fig. 6. The characteristic times (month) of SPI6, SPI12, SMI, and GWI derived from the Princeton-forced CLSM simulation mapped over the conterminous U.S. All calculations were based on data from 1948 to 2010.

Table 3
Correlation coefficients between region-averaged GWI derived from Princeton forced simulation (2nd row), the open loop (3rd row), the GRACE data assimilation (4th row) and the average GWI based on in situ groundwater. The correlation was calculated for 2003–2010, the common period of all data sets.

	Long Island	New Jersey	Massachusetts	Pennsylvania	Upper Mississippi	Ohio–Tennessee	Red–LM	Missouri
Princeton	0.43	0.77	0.69	0.72	0.72	0.92	0.69	0.77
NLDAS OL	0.60	0.85	0.72	0.78	0.73	0.89	0.55	0.73
GRACE DA	0.56	0.88	0.76	0.80	0.73	0.92	0.67	0.84

scales. This is due to the fact that GRACE data assimilation often exerts the largest impact on maxima and minima of groundwater storage and thus has a stronger influence on inter-annual variability. This effect can be seen in Fig. 7 which shows regional average (over well locations only) groundwater storage anomalies (relative to temporal means of each data set) from the open loop and the GRACE data assimilation simulations and in situ groundwater data. The open loop exhibits larger dynamic ranges than in situ groundwater in all regions except in Red–LM and Missouri. In all cases, GRACE data assimilation nudged model estimates towards in situ observations, demonstrating the value of GRACE TWS in conjunction with data assimilation for improving regional scale groundwater storage estimates and for drought monitoring.

GRACE data assimilation reduced groundwater persistence slightly in all regions (not shown), which may be a side effect of data assimilation, in general. When assimilation increments are applied, the continuity of estimated states is disrupted and thus the auto-correlation and characteristic time are reduced. The period of GRACE data assimilation (2003–2011) may also be too short to derive reliable auto-correlation values at larger lags where GRACE data assimilation has the most impact. This issue should be revisited when more GRACE data (and in situ data for comparison) become available.

4.3. Reconciliation of model estimates from different forcing data sets

To produce the weekly, GRACE data assimilation based drought indicators that are distributed by the National Drought Mitigation Center (NDMC; see http://drought.unl.edu/MonitoringTools/NASA-GRACEDataAssimilation.aspx), a climatology was derived from Princeton forced CLSM output, while the NLDAS forced GRACE data assimilation is used for near-real time model simulation and for generating drought indices. As indicated earlier, differences between the two forcing data sets can lead to discrepancies in estimated states and affect the accuracy of wetness rankings. This can be seen in Fig. 8, where groundwater storage from CLSM forced by unaltered NLDAS meteorology is biased low relative to groundwater from the Princeton climatology run (only data from 1990 to 2010 are shown). Bias correction of the NLDAS forcing fields to Princeton reduced the output groundwater bias (as shown by the open loop simulation) but left intact a dynamic range that was significantly larger than that of the Princeton forced output. GRACE data assimilation created additional biases, which would have led to wetter drought indices. The solution is to scale the statistics of the GRACE data assimilation output fields to be consistent with those of the Princeton forced climatology simulation:

$$g_{scaled} = (g_N - \bar{g}_N)\sigma_P/\sigma_N + \bar{g}_P \quad (2)$$

where g represents groundwater storage at any given location; subscripts P and N represent Princeton and NLDAS-forced (in our case, with GRACE data assimilation) estimates; the upper bar and σ represent the temporal mean and standard deviation of groundwater storage, respectively. Eq. (2) ensures that the mean and standard deviation of GRACE data assimilation based groundwater storage matches those of the Princeton forced simulation during the overlapping period (2003–2010). Fig. 8 shows that applying this

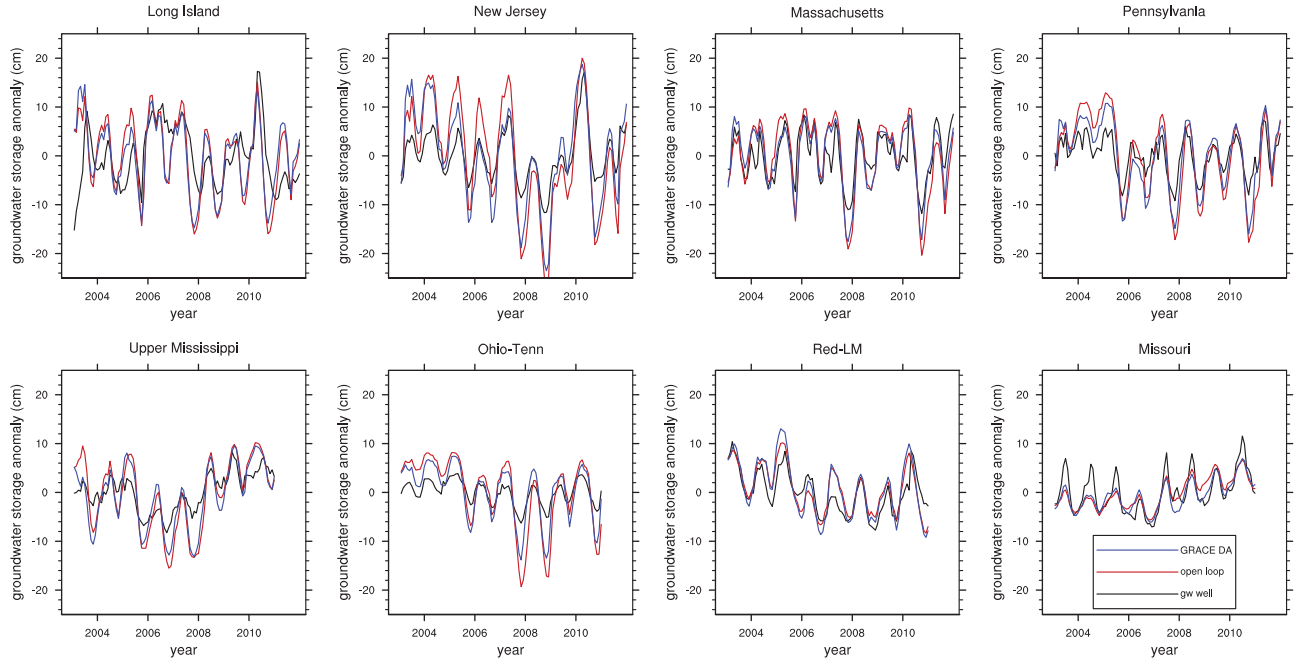

Fig. 7. Regional average (over well locations only) monthly groundwater storage anomalies (relative to temporal mean) from the NLDAS-forced open loop, GRACE data assimilation simulation, and groundwater observation wells.

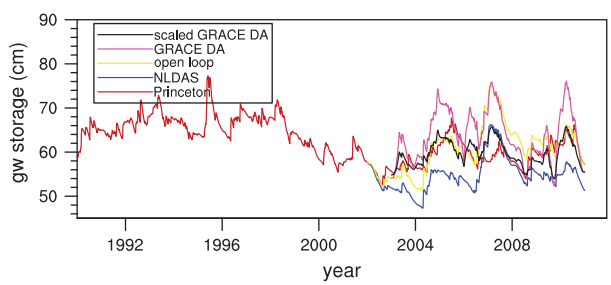

Fig. 8. Daily groundwater storage estimates at a northwestern Colorado catchment from Princeton and NLDAS forced simulations, the open loop (forced by NLDAS bias corrected to Princeton), GRACE data assimilation (forced by NLDAS bias corrected to Princeton), and scaled GRACE data assimilation output, for 1990–2010.

scaling method reduced the dynamic range of estimates from GRACE data assimilation but did not change their temporal variability. Alternatively, estimates from the Princeton forced simulation and GRACE data assimilation can be standardized, separately, before a climatology and the indices are created.

5. Summary and discussions

We evaluated a groundwater drought index that is based on groundwater storage simulated by the Catchment Land Surface Model. We found that the model based GWI was strongly correlated with GWI derived from in situ groundwater storage estimates averaged over most of the study regions. Averaged over individual well locations, the correlations between the two GWIs were significantly lower, probably because of the coarseness of the model input forcing and parameter fields. We conclude that the CLSM based GWI is more reliable at regional scales than at local scales.

The CLSM based GWI generally exhibits stronger correlation with SPI at longer timescales such as SPI12 and SPI24, reflecting the significantly lagged response of groundwater to precipitation anomalies. The correlation is influenced by the bedrock depth, which controls, along with mean annual precipitation, the depth to the water table. In the Great Plains, where the bedrock is deep, GWI is less correlated with SPI6 than in other regions. In the eastern U.S. (especially the wetter Tennessee–Alabama area), GWI is more strongly correlated with shorter timescale SPI where the water table is shallow, which in turn tends to occur where annual precipitation is greater and bedrock depths are shallow. Model based GWI was found to be best correlated with SPI on roughly the same timescales as in situ based GWI correlating with SPI, which provides some confidence in that CLSM represents groundwater variability with enough accuracy to be useful for drought monitoring.

Similarly, the characteristic time of CLSM based groundwater is largely controlled by the depth to water, which itself is related to bedrock depth, porosity and annual precipitation. The longest characteristic times were found in the Great Plains, which indicates that groundwater drought tends to persist there longer than in other regions. Further examination revealed that CLSM groundwater storage estimates in the Great Plains are often dominated by multi-year or decades-long cycles (not shown). Although similarly long cycles were also observed in some in situ groundwater observations in the region, the co-incidence of long characteristic times with deep bedrock in the model suggests that further study is warranted to ensure that the bedrock depth is appropriately set in CLSM and that the resulting characteristic times are realistic. Similarly, a model that does not generate enough baseflow may accumulate too much groundwater storage during wet years (Li et al., 2012), leading to longer term variability and thus amplified persistence in groundwater. Lo et al. (2010) showed that water table dynamics can be better simulated by calibrating model parameters using a combination of gauged stream flows and GRACE TWS. Recalling that the CLSM bedrock depth was increased by 2 m to better match the dynamic range of modeled TWS to that of GRACE TWS, it is possible that modeled groundwater variations and persistence would be further improved by adjusting the parameters that control base flow generation.

In any case, when the persistence of groundwater is very long, it raises the question of groundwater's value as a drought indicator. What is the value of an indicator that lags years behind the onset

and conclusion of a drought or pluvial event? It is possible that the percentile based groundwater index is not a useful indicator in regions, such as the Great Plains, where groundwater's response to meteorological factors is particularly slow.

This study also showed that groundwater affects soil moisture in wet regions where the water table is shallow. In those situations, the soil moisture drought index (SMI) derived from CLSM often correlates strongly with longer timescale SPI, and soil moisture persistence exhibits spatial patterns similar to those of groundwater persistence. It was also found that CLSM soil moisture persists longer in the eastern U.S. than in the west, which is the opposite of the conclusion of Mo and Schemm (2008) and Sheffield et al. (2012), both of whom used models that lacked a groundwater scheme. Establishing with certainty whether soil moisture persists longer in the eastern or western U.S. would require an analysis of in situ soil moisture, which is beyond the scope of this study.

In most regions, GRACE data assimilation improved the temporal correlation between regional average CLSM based GWI and GWI derived from in situ data. With coefficients of correlation ranging from 0.56 to 0.92, it provides some justification for the use of GRACE data assimilation within CLSM for large scale drought monitoring. While the improvements due to GRACE data assimilation were small in several cases, it should be considered that the open loop simulation was forced by high quality NLDAS-2 data, resulting in correlation coefficients that were already high. Larger improvements would be expected in areas where the forcing data were not as good (while the quality of the GRACE observations is fairly uniform within a given latitude band). Further, we anticipate that direct assimilation of gridded GRACE TWS data fields (as opposed to averaging and assimilation over river basins) would preserve more information in the data and thus improve the model estimates more.

A shortcoming of this study is that the CLSM based GWI was not evaluated in an arid climate, owing to the scarcity of groundwater observations that meet our criteria (see Section 2) in the southwestern U.S. That would have been a good test, considering that uncertainties in forcing fields and sensitivities of modeled states to forcing errors are generally higher in dry regions (Gottschalck et al., 2005). Further, the importance of groundwater as a resource, particularly during droughts, is heightened in arid and semi-arid environments where other sources of water (rainfall and surface waters) are strained by the needs of people and agriculture (e.g., Castle et al., 2014). Groundwater abstractions are not simulated by CLSM, which can be problematic where the rates of abstraction are high. GRACE detects groundwater depletion caused by abstractions, thus GRACE data assimilation helps to mitigate related model errors to some extent. However, when abstractions cause massive groundwater depletion over an extended period of time, the dynamic range of modeled groundwater maybe insufficient to accommodate the assimilated values (Zaitchik et al., 2008). In those cases, percentile based groundwater drought indices would lose their meaning even if the impacts of abstractions were represented by the model, and it may in fact be preferable to ignore them and instead isolate the groundwater variability effected by meteorological conditions. For example, in the groundwater drought indicator maps distributed by the NDMC, the southern High Plains aquifer in the Texas panhandle would be always red (exceptional drought) if the model simulated human-induced depletion of that aquifer, which would be worthless for drought monitoring.

In conclusion, groundwater simulation through CLSM in conjunction with GRACE data assimilation is capable of providing information for groundwater drought monitoring that is much needed for various water management, agricultural, economic, and social applications. The value of a distributed model like CLSM lies in its ability to simulate groundwater variability with sufficient realism, as shown in this study, by driving a set of physical equations that represent water and energy cycle processes including evaporation and infiltration with high quality meteorological fields. Groundwater variability is controlled by a combination of meteorology, topography, hydrogeology, land cover, and in some cases water management. Hence groundwater drought is not a sub-class of meteorological or hydrological drought and it should be monitored independently and used to augment drought severity and impact assessments. Model based groundwater variations are imperfect, but in most of the world such information is unavailable at resolutions finer than what is provided by GRACE. We believe that combining the higher resolution of a model such as CLSM with the realism of GRACE through data assimilation is currently the best solution for global groundwater monitoring. Finally, we emphasize the importance of evaluating model estimates using in situ data including soil moisture, groundwater, and runoff, and applying the results to perform regional refinement of the model and the data assimilation approach, which will lead to improved groundwater drought quantification.

Acknowledgements

This study was supported by NASA's Terrestrial Hydrology Program and GRACE Science Team, and by the NASA Centers National Climate Assessment. We thank the valuable comments made by Yoshihide Wada and another reviewer which have helped improve the quality of this paper.

References

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

Castle, S.L., Thomas, B.F., Reager, J.T., Rodell, M., Swenson, S.C., Famiglietti, J.S., 2014. Groundwater depletion during drought threatens future water security of the Colorado River basin. Geophys. Res. Lett. (in press). doi:10.1002/2014GL061055.

Changnon, S.A., 1987. Detecting Drought Conditions in Illinois (No. 163–170). Illinois State Water Survey.

Cosgrove, B., Lohmann, D., Mitchell, K.E., Houser, P.R., Wood, E.F., Schaake, J.C., Robock, A., Marshall, C.R., Sheffield, J., Duan, Q., Luo, L., Higgns, R.W., Pinker, R.T., Tarpley, J.D., Meng, J., 2003. Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. J. Geophys. Res. 108 (D22), 8842. http://dx.doi.org/10.1029/2002JD003118.

Eltahir, E.A.B., Yeh, P.J.-F., 1999. On the asymmetric response of aquifer water level to floods and droughts in Illinois. Water Resour. Res. 35 (4), 1199–1217.

Famiglietti, J.S., Lo, M., Ho, S.L., Bethune, J., Anderson, K.J., Syed, T.H., Swanson, S.C., deLineage, C.R., Rodell, M., 2011. Satellites measure recent rates of groundwater depletion in California's Central Valley. Geophys. Res. Lett. 38, LO3403. http://dx.doi.org/10.1029/2010GL046442.

Feng, W., Zhong, M., Lemoine, J.-M., Biancale, R., Hsu, H.-T., Xia, J., 2013. Evaluation of groundwater depletion in North China using the Gravity Recovery and Climate Experiment (GRACE) data and ground-based measurements. Water Resour. Res. 49, 2110–2118. http://dx.doi.org/10.1002/wrcr.20192.

Gottschalck, J.J., Meng, M.Rodell, Houser, P., 2005. Analysis of multiple precipitation products and preliminary assessment of their impact on global land data assimilation system land surface states. J. Hydro. Meteor. 6 (5), 573–598.

Houborg, R., Rodell, M., Li, B., Reichle, R., Zaitchik, B.F., 2012. Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. Water Resour. Res. 48, W07525. http://dx.doi.org/10.1029/2011WR011291.

Johnson, A.I., 1967. Specific Yield – Compilation of Specific Yields for Various Materials. U.S. Geological Survey Water Supply Paper 1662-D, 75 pp.

Koster, R.D., Suarez, M.J., Ducharne, A., Stieglitz, M., Kumar, P., 2000. A catchment-based approach to modeling land surface processes in a general circulation model, 1, Model structure. J. Geophys. Res. 105, 24809–24822.

Landerer, F.W., Swenson, S.C., 2012. Accuracy of scaled GRACE terrestrial water storage estimates. Water Resour. Res. 48, W04531. http://dx.doi.org/10.1029/2011WR011453.

Li, B., Rodell, M., Zaitchik, B.F., Reichle, R.H., Koster, R.D., van Dam, T.M., 2012. Assimilation of GRACE terrestrial water storage into a land surface model: evaluation and potential value for drought monitoring in western and central Europe. J. Hydrol., 103–115. http://dx.doi.org/10.1016/j.jhydrol.2012.04.035.

Leblanc, M.J., Tregoning, P., Ramillien, G., Tweed, S.O., Fakes, A., 2009. Basin-scale, integrated observations of the early 21st century multiyear drought in southeast Australia. Water Resour. Res. 45, W04408. http://dx.doi.org/10.1029/2008WR007333.

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

McKee, T.B., Doesken, N.J., Kleist, J., 1993. The relationship of drought frequency and duration of time scales. Eighth Conference on Applied Climatology, American Meteorological Society, Jan17–23, 1993, Anaheim CA, pp. 179–186.

Mishra, A.K., Singh, V.P., 2010. A review of drought concepts. J. Hydrol. 391, 202–216. http://dx.doi.org/10.1016/j.jhydrol.2010.07.012.

Mo, K.C., 2008. Model-based drought indices over the United States. J. Hydrometeor. 9, 1212–1230. http://dx.doi.org/10.1175/2008JHM1002.1.

Mo, K.C., Schemm, J.E., 2008. Droughts and Persistent Wet Spells over the United States and Mexico. J. Climate 21, 980–994. http://dx.doi.org/10.1175/2007JCLI1616.1.

Niu, G.-Y., Yang, Z.-L., Dickinson, R.E., Gulden, L.E., Su, H., 2007. Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data. J. Geophys. Res. 112, D07103. http://dx.doi.org/10.1029/2006JD007522.

Peters, E., van Lanen, H.A.J., Torfs, P.J.J.F., Bier, G., 2005. Drought in groundwater—drought distribution and performance indicators. J. Hydrol. 306, 302–317.

Rodell, M., Chen, J., Kato, H., Famiglietti, J., Nigro, J., Wilson, C., 2007. Estimating ground water storage changes in the Mississippi River basin (USA) using GRACE. Hydrogeol. J. 15, 159–166. http://dx.doi.org/10.1007/s10040-006-0103-7.

Rodell, M., Velicogna, I., Famiglietti, J.S., 2009. Satellite-based estimates of groundwater depletion in India. Nature 460, 999–1002. http://dx.doi.org/10.1038/nature08238.

Rodell, M., 2012. Satellite gravimetry applied to drought monitoring. In: Wardlow, B., Anderson, M., Verdin, J. (Eds.), Remote Sensing of Drought: Innovative Monitoring Approaches. CRC Press/Taylor and Francis, Boca Raton, pp. 261–280.

Rowlands, D.D., Luthcke, S.B., Klosko, S.M., Lemoine, F.G.R., Chinn, D.S., McCarthy, J.J., Cox, C.M., Anderson, O.B., 2005. Resolving mass flux at high spatial and temporal resolution using GRACE intersatellite measurements. Geophys. Res. Lett. 32 (4).

Sheffield, J., Doteti, G., Wood, E.F., 2006. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. J. Climate 19 (7), 3088–3111.

Sheffield, J., Livneh, B., Wood, E., 2012. Representation of terrestrial hydrology and large scale drought of the continental US from the North American regional reanalysis. J. Hydrometeor.. http://dx.doi.org/10.1175/JHM-D-11-065.1.

Su, H., Yang, Z.-L., Dickinson, R.E., Wilson, C.R., Niu, G.-Y., 2010. Multisensor snow data assimilation at the continental scale: The value of Gravity Recovery and Climate Experiment terrestrial water storage information. J. Geophys. Res. 115, D10104. http://dx.doi.org/10.1029/2009JD013035.

Swenson, S., Wahr, J., 2006. Post-processing removal of correlated errors in GRACE data. Geophys. Res. Lett. 33 (8), L08402. http://dx.doi.org/10.1029/2005GL025285.

Tapley, B.D., Bettadpur, S., Ries, J.C., Thompson, P.F., Watkins, M.M., 2004. GRACE measurements of mass variability in the Earth system. Science 305, 503–505.

Swenson, S., Yeh, P.J.F., Wahr, J., Famiglietti, J., 2006. A comparison of terrestrial water storage variations from GRACE with in situ measurements from Illinois. Geophys. Res. Lett. 33 (16).

Thomas, A.C., Reager, J.T., Famiglietti, J.S., Rodell, M., 2014. A GRACE-based water storage deficit approach for hydrological drought characterization. Geophys. Res. Lett. 41, 1537–1545. http://dx.doi.org/10.1002/2014GL05932.

Voss, K.A., Famiglietti, J.S., Lo, M., de Linage, C., Rodell, M., Swenson, S.C., 2013. Groundwater depletion in the Middle East from GRACE with implications for transboundary water management in the Tigris-Euphrates-Western Iran region. Water Resour. Res. 49. http://dx.doi.org/10.1002/wrcr.20078.

Wada, Y., van Beek, L.P.H., van Kempen, C.M., Reckman, J.W.T.M., Vasak, S., Bierkens, M.F.P., 2010. Global depletion of groundwater resources. Geophys. Res. Lett. 37, L20402. http://dx.doi.org/10.1029/2010GL044571.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., Livneh, B., Lettenmaier, D., Koren, V., Duan, Q., Mo, K., Fan, Y., David Mocko, 2012. Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. J. Geophys. Res. 117, D03109. http://dx.doi.org/10.1029/2011JD016048.

Xia, Y., Sheffield, J.S., Ek, M.B., Dong, J., Chaney, N., Wei, H., Meng, J., Wood, E.F., 2014a. Evaluation of multi-model simulated soil moisture in NLDAS-2. J. Hydrol. V512 (May), 107–125. http://dx.doi.org/10.1016/j.jhydrol.2014.02.027.

Xia, Y., Ek, M., Mocko, D., Peters-Lidard, C., Sheffield, J., Dong, J., Wood, E., 2014b. Uncertainties, correlations, and optimal blends of drought indices from the NLDAS multiple land surface model ensemble. J. Hydrometeor. http://dx.doi.org/10.1175/JHM-D-13-058.1.

Zaitchik, B.F., Rodell, M., Reichle, R.H., 2008. Assimilation of GRACE terrestrial water storage data into a land surface model: results for the Mississippi river basin. J. Hydrometeor. 9, 535–548. http://dx.doi.org/10.1175/2007JHM951.1.

Please cite this article in press as: Li, B., Rodell, M. Evaluation of a model-based groundwater drought indicator in the conterminous U.S. J. Hydrol. (2014), http://dx.doi.org/10.1016/j.jhydrol.2014.09.027