Groundwater Drought: Environmental Controls and Monitoring

Bailing Li[1,2] & Matthew Rodell[2]

[1]ESSIC, University of Maryland

[2]Hydrological Sciences Laboratory, NASA Goddard Space Flight Center

Abstract (maximum 150 words)

Arising from a long-term deficit of precipitation, groundwater drought often lags meteorological drought by months to years, depending on climate and subsurface hydrogeological conditions. Due to the paucity of *in situ* groundwater data in many parts of the world, global monitoring of groundwater storage variations and drought may be best achieved using satellite observations and/or groundwater time series simulated by hydrological models. The Gravity Recovery and Climate Experiment (GRACE) and GRACE Follow On (GRACE-FO) missions have greatly benefitted the modeling and monitoring of groundwater storage changes and drought at the global scale. In this chapter, we first review environmental controls on the temporal variability of groundwater using *in situ* data. We then describe an approach that infuses GRACE and GRACE-FO observations into a land surface model for assessing groundwater storage changes and drought globally. We also discuss characteristics of simulated groundwater drought and the limitations of current groundwater drought monitoring approaches.

1.Introduction

Drought, as initiated by below-normal precipitation over weeks or longer time periods, can propagate through the components of the hydrological cycle including groundwater (Changnon, 1987). Groundwater drought, defined as conditions when water tables drop below their normal levels, is a class of drought distinct from hydrological and agricultural drought (Mishra and Singh, 2010). Integrating surface meteorological conditions over months to years, groundwater variations lag changes in precipitation longer than soil moisture or streamflow (Eltahir and Yeh, 1999; Li and Rodell, 2015). In addition, because overlying layers of soil and rock act as a low-pass filter on near-surface processes, groundwater storage changes reflect mainly the low frequency hydrometeorological variations (Eltahir and Yeh, 1999; Wang, 2012), making it difficult to estimate groundwater drought conditions using indicators designed for other types of drought.

Groundwater has been under immense stress in many parts of the world in recent decades (Mukherjee et al., 2020) Globally, more than 2 billion people use groundwater exclusively for drinking water, and 43% of irrigation water comes from aquifers (UNESCO, https://en.unesco.org/themes/water-security/hydrology/groundwater). Groundwater depletion associated with groundwater withdrawals for irrigation have been reported in many agriculturally important regions (Rodell et al., 2009; Famiglietti et al., 2011; Feng et al., 2013; Voss et al., 2013; Rodell et al., 2018). Changes in the groundwater systems also affect watershed health, as groundwater sustains streamflow in arid and semi-arid regions and in wet regions during dry periods (Eltahir and Yeh, 1999; Hughes et al., 2012). Decreases in water table levels in recent decades have led to significant forest diebacks in the floodplains of Australia's managed rivers (Cunningham et al., 2011; Kath et al., 2014) and altered the landscape of riparian-dependent

43 woodlands in the southwestern U.S. (Stromberg et al., 1992). Climate change may further stress
44 groundwater systems through its impacts on precipitation quantities, types, and intensity,
45 evapotranspiration, and water demand and re-distribution (Green et al., 2011; Taylor et al., 2012;
46 Famiglietti, 2014; Van Loon et al., 2016; Condon et al., 2020).

47 Groundwater drought is generally not well monitored due to limited availability of groundwater
48 observations around the world. The U.S. Geological Survey (USGS) is one of a very few
49 organizations that provide real-time groundwater conditions using well data
50 (https://groundwaterwatch.usgs.gov/default.asp), but it is limited to the United States, and much
51 of the country is under-sampled in real-time groundwater observations. At the time of writing
52 there were 1793 wells in the real-time network. The USGS archives data from thousands more
53 wells, but most of them are either infrequent or have short or discontinuous records. In addition,
54 wells located in confined aquifers are unsuitable for groundwater storage monitoring because
55 changes in hydraulic head are not necessarily caused by storage changes. A handful of other
56 countries including Australia and the Netherlands also make groundwater well data available
57 either directly (e.g., http://www.bom.gov.au/water/groundwater/explorer/;
58 https://www.dinoloket.nl/) or through the International Groundwater Resources Assessment
59 Centre (https://ggmn.un-igrac.org/) but they suffer from the same issues as the USGS data.
60 Groundwater data are unavailable in most of the rest of the world due to lack of observation
61 networks or measurements, records that are not digitized and centralized, or data access
62 restrictions (e.g., Rodell et al., 2009).

63 Numerical modeling provides an alternative means for estimating groundwater variations and
64 monitoring groundwater drought (Li and Rodell, 2015). When forced by temporally consistent
65 meteorological data, hydrological models can provide spatially and temporally continuous

groundwater fields at regional and global scales (Döll et al., 2014; Sutanudjaja et al., 2018; Li et al., 2019a). In addition, advanced hydrological models, constrained by water and energy balances, are capable of providing reasonable ET estimates (Mueller et al., 2011; Jimenez et al., 2011) which, along with precipitation, has the largest impact on groundwater temporal variability (Eltahir and Yeh, 1999; Li and Rodell, 2015). In recent decades, satellite observations from the Gravity Recovery and Climate Experiment (GRACE, Tapley, 2004) and GRACE Follow On (GRACE-FO) missions have enabled new methods of monitoring groundwater storage and drought at regional to global scales (Rodell and Famiglietti, 2002; Rodell et al., 2007; Zaitchik et al., 2008; Houborg et al., 2012; Famiglietti and Rodell, 2013; Döll et al., 2014; Kumar et al., 2016; Girotto et al., 2017; Li et al., 2019b).

In this chapter, we first review groundwater temporal variability reflected in *in situ* groundwater observations and then describe an approach that infuses GRACE and GRACE-FO observations into a land surface model to enable groundwater storage and drought monitoring. Finally, we discuss challenges facing groundwater drought monitoring and future research directions.

2.Environmental controls on groundwater

To understand groundwater drought and its environmental controls, we first examine groundwater temporal variability using *in situ* data from 181 wells located in four regions in the northeastern U.S. and four sub-basins of the Mississippi River (Fig.1). Three criteria were used for selecting well records from the larger set of available observations in the USGS archive (Rodell et al., 2007; Li and Rodell, 2015). First, water level measurements must show seasonal variability that indicates communication with the atmosphere. This eliminates wells located in

confined aquifers. Second, to exclude direct anthropogenic impacts such as withdrawals and injections, wells with large abrupt changes in their water level time series were excluded. Third, we only selected wells with at least 10 years of temporally consistent observations (minimum four measurements per year) in order to allow examination of seasonality and inter-annual variability. For the purpose of examining temporal variability, water level measurements were converted to anomalies by subtracting the long-term mean at each location. The resulting water level anomalies were further converted to groundwater storage anomalies by multiplying them by the specific yield, which was estimated individually for each well (Rodell et al., 2007; Li et al., 2015) and provided in Li et al. (2019b). This enables comparison of groundwater storage changes across wells and regions.

2.1 Precipitation

Fig.2 shows monthly groundwater storage anomalies from individual wells (gray lines) and their regional average in the eight regions (dark black lines). The bars at the top represent monthly precipitation from the North America Land Data Assimilation System-2 (NLDAS-2, Xia et al., 2012). NLDAS-2 precipitation is derived by temporally disaggregating daily gauge measurements using hourly radar images and is further spatially interpolated to a 0.125° grid (Cosgrove et al., 2003). Regional groundwater storage and conforming precipitation time series were computed by averaging in situ groundwater data and gridded precipitation data from all well locations within each region. The influence of precipitation on regional groundwater storage anomalies (relative to long-term mean) is apparent during prolonged wet and dry events (Fig.2). For instance, the well-known drought that affected the Midwest and Northern Great Plains during 1987-1989 produced large negative anomalies in the Upper Mississippi and

Missouri basins, and a large positive anomaly is seen in the Upper Mississippi during the summer floods of 1993.

These groundwater measurements exhibit strong seasonal variation peaking in April to May and reaching annual minimum in fall and winter, despite weak seasonality in precipitation such as in the four northeast regions or that out of phase with that of groundwater such as in the Upper Mississippi basin (Fig. 3) (Li et al., 2015). The strong seasonality is shaped by combined influences of precipitation and ET. For instance, in the Upper Mississippi and Missouri where precipitation shows strong seasonality, maximum groundwater and precipitation occurs in the same month (Missouri) or one month apart (Upper Mississippi). In the northern and high altitudes of these regions, seasonal snowpack also stores a significant amount of annual precipitation and releases it through spring snowmelt (Perez-Valdivia et al., 2012), which balances the effect of increasing ET in the spring. On the other hand, in the Ohio-Tennessee and the combined Red River and Lower-Mississippi (Red-LM) basins precipitation increases in late spring to early summer, but evaporative demand and plant root uptake also increase, reducing the water available for groundwater recharge and resulting in groundwater peaking before precipitation peaks.

Groundwater exhibits more noticeable lagged responses to precipitation when the seasonal cycles are removed from both time series (Fig.4) than in Fig.2. The maximum lagged correlation is greater than 0.5 in seven of the eight regions, with the lag of maximum correlation ranging from 2 to 7 months. In the Missouri basin, which is the driest among all regions/basins, groundwater storage reached the lowest level in 1992, much later than the precipitation minimum in 1988 (note precipitation in Fig.4 was smoothed with a 6-month running average). The long lag likely reflects the delayed response and long recovery time needed in a dry climate where

recharge is low and the water table is deep (Fig.2). Such lagged groundwater responses to precipitation were postulated by Changnon (1987). Further, Eltahir and Yeh (1999), based on nearly 30-years of *in situ* data in Illinois, showed that groundwater drought persisted longer with higher intensity than pluvial events. This behavior was attributed to the non-linear dependency of groundwater discharge on water levels which allows faster dissipation of groundwater during floods.

2.2 Subsurface hydrogeological conditions

Hydrogeological conditions such as water table depths and aquifer properties may also impact groundwater temporal variability and groundwater drought through their impacts on recharge rates. Broomfield et al. (2015) examined water level measurements from three principle aquifers in a small region in England where climate conditions are relatively uniform. They found that the autocorrelation scale of groundwater time series increases with increasing water table depth and that longer autocorrelation scale is associated with longer duration and more intensive groundwater drought. The former can be attributed to the previously mentioned low-pass filtering effect of overlying soil and rock on hydrometeorological conditions (Wu et al., 2002). Also based on well data from England, Broomfield and Marchant (2013) showed that groundwater autocorrelation scales are negatively correlated with soil diffusivity, more so in granular aquifers than in fractured geological formations where the unsaturated zone depth plays a stronger role on groundwater temporal variability.

The impacts of hydrogeological properties and conditions are not easily observable in the groundwater data from the eight U.S. regions discussed above, where climate variability likely dominates groundwater temporal variability. However, some results may be attributed to hydrogeological conditions. For instance, the low correlation (0.17 at 7 months lag) between

non-seasonal groundwater and precipitation in the Red-LM basin may be due to the thickness of the unsaturated zone above the water table (Fig.2), whereas the correlation in Massachusetts, where the water table is shallow, was strong (0.85 at 3 months lag). In Pennsylvania, where the regional water table is also deep, non-seasonal groundwater reached a maximum correlation with precipitation, 0.74, at a 2-month lag. Most wells in Pennsylvania are located in fractured rock formations (Li and Rodell, 2015) which allow fast recharge during precipitation, hence strong correlation at a short lag. These results suggest complex and intertwined environmental controls on groundwater at the seasonal to inter-annual scales.

2.3 large-scale climate phenomena

As with other types of drought, groundwater drought is often caused by large-scale climate signals and multi-year oscillations (Asong et al., 2018). These phenomena influence groundwater through their impacts on precipitation, mainly on inter-annual scales. Among them, the Pacific Decadal Oscillation (PDO) and the El Niño Southern Oscillation (ENSO) are known to have the largest impacts on multi-year cyclic behaviors of groundwater in the western U.S. (Hanson et al., 2006; Barco et al., 2010; Velasco et al., 2015). Perez-Valdivia et al. (2012) observed effects of ENSO and PDO on shallow groundwater in the Canadian prairies. Anderson and Emanuel (2008) reported that ENSO affected winter baseflow in North Carolina. Asoka et al. (2017) theorized that groundwater decline in northern Indian was partly attributable to Indian Ocean warming. Several studies have investigated the potential effects of climate change on groundwater storage and recharge (Green et al., 2011; Taylor et al., 2012; Kløve et al., 2014; Meixner et al., 2016; Niraula et al., 2017).

Subsurface properties and hydrogeological conditions have been shown to modulate groundwater responses to large-scale climate signals. For instance, Anderson and Emanuel (2008) found that

groundwater in highly permeable aquifers shows stronger responses to climate signals than that in low permeability aquifers in the same climate regime. Hanson et al. (2006) showed that groundwater variations in wells close to mountain fronts or rivers where hydraulic gradients are large exhibit stronger influences of climate phenomena than those in other areas. These results reflect the more dominant control of climate on groundwater in aquifers where the overlying geology does not inhibit hydrological communication with the surface and atmosphere.

3.Groundwater drought monitoring

Because of the scarcity of *in situ* groundwater observations in most of the world, global-scale groundwater drought monitoring has to rely on information from hydrological models and/or satellite observations. In particular, land surface models (LSMs) have been widely applied for drought monitoring at continental to global scales (e.g., Mo, 2008; Sheffield et al., 2012; Houborg et al, 2012; Li et al., 2019b). When forced by observation-based meteorological fields, LSMs are capable of simulating spatially and temporally continuous groundwater conditions suitable for drought monitoring (Houborg et al., 2012; Li et al., 2015). An important component of this approach is the skill of LSMs in simulating ET (Mueller et al., 2011; Jiménez et al., 2011) which, along with precipitation, is the dominant driver of groundwater variability (Eltahir and Yeh, 1999; Li et al., 2015). Just as importantly, LSM simulations can provide the long time series of groundwater storage changes needed for identifying and quantifying groundwater drought in the context of historical variability (Houborg et al., 2012). However, simulated groundwater time series are subject to uncertainties stemming from simplified model physics (Koster et al., 2000; Yeh and Eltahir, 2005a and 2005b; Niu et al., 2007; Koirala et al., 2014) and errors in the forcing data used to drive the model (Xia et al., 2017; Li et al., 2019a, 2019b).

3.1 GRACE data assimilation for groundwater drought monitoring

The GRACE and GRACE-FO satellite missions were designed to map Earth's gravity field on a monthly basis with sufficient precision to quantify variations in terrestrial water storage (TWS), among other mass changes (Tapley et al., 2004). TWS comprises soil moisture, groundwater, snow, surface waters, and ice (Rodell and Famiglietti, 2001). GRACE and GRACE-FO, which span from 2002 to present with an 11-month gap in 2017-2018 (https://grace.jpl.nasa.gov/), provide the only satellite-based observations that can be used to infer changes in water stored below the surface soil layer, including groundwater. GRACE and GRACE-FO detect water storage changes throughout all seasons (Li et al., 2012), which is critical for continuous drought monitoring. Low levels of non-seasonal GRACE derived TWS have been shown to correlate strongly with drought (Andersen et al., 2005; Yirdaw et al., 2008; Chen et al., 2009; Leblanc et al., 2009; Li et al., 2012; Long et al., 2013; Thomas et al., 2017; Zhao et al. 2017).

One of the challenges in applying GRACE/GRACE-FO data for groundwater drought monitoring lies in isolating groundwater from the vertically integrated TWS observations. In addition, GRACE/GRACE-FO data are provided monthly, often with 2-5 months of latency, with an effective spatial resolution around 150,000 km^2 at mid-latitudes (Rowlands et al., 2005; Swenson et al., 2006). These characteristics complicate the use of GRACE and GRACE-FO data for operational applications such as drought monitoring, which requires near-real time, fine-scale information.

To overcome these challenges, methods have been developed, tested, and refined for assimilating GRACE/GRACE-FO data into LSMs using ensemble Kalman smoothers and similar techniques (Zaitchik et al., 2008; Girotto et al., 2016; Kumar et al., 2016). The assimilated results have significantly higher spatial and temporal resolutions than the GRACE and GRACE-FO data with

sub-weekly latency. The output states and fluxes, including soil moisture, groundwater, snow water equivalent, and streamflow, are in most cases superior to unassimilated (open loop) model output (Zaitchik et al., 2008; Su et al., 2010; Houborg et al., 2012; Li et al., 2012; Girotto et al., 2017; Kumar et al., 2017; Li et al., 2019b). An ensemble smoother can be represented by the update equation (Zaitchik et al., 2008):

$$\boldsymbol{X^a} = \boldsymbol{X^f} + \underline{\boldsymbol{K}}(\boldsymbol{Y_0} - \boldsymbol{Y_M}) \qquad (1)$$

where $\boldsymbol{X}^{a}$ represents the analysis of modeled states at the current time step; $\underline{\boldsymbol{K}}$ is the Kalman gain matrix which is calculated from observation errors and model uncertainty represented in the ensemble; $\boldsymbol{Y_o}$represents GRACE/GRACE-FO-derived TWS; $\boldsymbol{Y}_M$ represents model simulated TWS. Because GRACE/GRACE-FO TWS are anomalies (relative to temporal mean), $\boldsymbol{Y_o}$ is created by adding the long-term temporal mean of modeled TWS to GRACE/GRACE-FO TWS. The states ($\boldsymbol{X}^f$ and $\boldsymbol{X}^a$) are computed at the model's finer scale and daily time step, while $\boldsymbol{Y_o}$and $\boldsymbol{Y}_M$ are evaluated monthly at the coarser scale of GRACE data, hence the smoother enables spatial and temporal downscaling.

Below we discuss results obtained from assimilating GRACE/GRACE-FO data into the Catchment land surface model (CLSM, Koster et al., 2000) for groundwater storage and drought monitoring at the global scale (Li et al., 2019b). CLSM simulates subsurface water storage changes in three subsurface water zones, a surface layer (0-2 cm below the surface), the root zone (0-100 cm) and the full soil profile whose thickness is determined by CLSM's bedrock depth parameter. It also simulates seasonal snowpack changes in three snow layers and canopy interception. CLSM does not model water table variations, hence groundwater storage is derived

by subtracting the water stored in the root zone from that in the full soil profile. Based on this configuration, the state vector in equation (1), $\boldsymbol{X}$, represents root zone soil moisture, groundwater, snow, and canopy interception and $\boldsymbol{Y}_M$, represents the sum of the aforementioned variables.

Fig. 5 shows that monthly TWS from the open loop simulation, i.e., without assimilating GRACE data (top panel), GRACE data assimilation (2nd panel from the top), and GRACE alone (3rd panel from the top) for August 2015. It is clear that GRACE data assimilation brought simulated TWS closer to the GRACE observation while retaining the spatial details of the open loop simulation, which in turn reflect spatial variability in the high-resolution model parameter and forcing data. The large negative values in the Canadian Archipelago and the Gulf of Alaska are associated with ice sheet and glacier mass losses (Gardner et al., 2011). CLSM does not simulate these processes and therefore, did not simulate decreasing trends in those areas.

The bottom panel of Fig.5 shows TWS time series for an area in Southern Germany which experienced severe drought in 2003, 2015, and 2018 and which has been studied by Van Loon et al. (2017) using *in situ* groundwater data. Open loop CLSM indicates more severe drought in 2003 than in 2015, which is consistent with 2015 being a wetter year than 2003 (Van Loon et al., 2017). However, the region had experienced precipitation deficits for several years leading up to the 2015 drought which caused a long-term deficit in groundwater storage (Van Loon et al., 2017). GRACE detected this long-term decline in groundwater storage and hence, signified more severe drought in 2015. Open loop CLSM TWS did not capture this multi-year water storage decline, likely due to deficiencies in model physics and/or errors in the forcing data. Regardless of the source of uncertainty, data assimilation adjusted simulated TWS towards GRACE observations, resulting in a more accurate depiction of drought in 2015. The same plot

also demonstrates that the assimilation model fills data gaps including that between GRACE and GRACE-FO (August 2017 to June 2018) by relying on forcing data alone as input, while the soil moisture and groundwater "memory" ensures continuity from GRACE to GRACE-FO periods.

Fig.6 presents an evaluation of the data assimilation results using the independent, *in situ* groundwater observations described in section 2. GRACE data assimilation decreased the root mean square error (RMSE) in all regions and increased the correlation with *in situ* groundwater in six of the eight regions. Fig.6 also shows that GRACE/GRACE-FO data assimilation does not guarantee agreement with the *in situ* observations. For instance, groundwater from both the open loop and the GRACE/GRACE-FO data assimilation identified 2006 as a drought year in Red-LM, more than a year earlier than did the *in situ* observations. It is unknown whether this discrepancy was caused by errors in the forcing and GRACE data or perhaps by under-sampling of the region by the thirteen wells used to construct the average (Fig.1). Despite the limitations of this type of comparison, Li et al. (2019b) reported generally positive impacts of GRACE data assimilation on simulated groundwater in regions that spanned five continents based on time series from nearly 4,000 wells. In particular, RMSE decreased by 36% and 10% and correlation improved by 16% and 22% at the regional and point scales, respectively (Li et al., 2019b). These improvements demonstrate the value of GRACE/GRACE-FO data assimilation for groundwater storage change assessment and drought monitoring.

Fig. 7 shows global groundwater wetness percentile maps for Sept. 22, 2003 and Sept. 21, 2015 derived from GRACE/GRACE-FO data assimilation output. As shown in Fig.5, the GRACE-based groundwater indictor shows more severe drought in 2015 than in 2003 in Southern Germany and the surrounding region. In addition to spatial downscaling, GRACE/GRACE-FO data assimilation enables temporal downscaling so that drought conditions can be provided daily

(as opposed to monthly with GRACE/GRACE-FO observations). The percentiles were created by ranking groundwater estimates against a climatology derived from a 1948-2014 CLSM simulation forced by the Princeton meteorological forcing dataset (Sheffield et al., 2006). The model simulation is used for climatology because the GRACE and GRACE-FO data period alone (2002-present) is too short to reveal the full range of variability in groundwater. Scale factors were derived, based on the overlapping period between GRACE/GRACE-FO data assimilation (which utilized a different meteorological forcing dataset, ECMWF, Dee et al., 2011) and the long term simulation, and employed to ensure the consistency between the long term model climatology and the GRACE/GRACE-FO data assimilation (Li and Rodell, 2015; Li et al. 2019b). The scale factors were derived for each pixel and each calendar month to avoid non-Gaussian behaviors in simulated subsurface states (Mo, 2008).

3.2 Other groundwater drought indicators

Groundwater drought indicators have also been derived by subtracting soil moisture and snow water equivalents (usually estimated by LSMs) from GRACE TWS, under the assumption that other TWS change components are negligible (e.g., Thomas et al., 2014; Van Loon et al., 2017). However, the resulting indicators have the same spatial and temporal (monthly) scales and latency as those of GRACE and GRACE-FO data, making them unsuitable for real-time operational drought monitoring.

Van Loon et al. (2017) developed a groundwater drought indicator based on the correlation scale between *in situ* groundwater and the Standard Precipitation Index (SPI). While allowing detection of groundwater drought using readily available precipitation data, the approach requires long records of groundwater data and thus, is limited to regions where such records are

available. In addition, non-linear processes affecting groundwater such as groundwater discharge and recharge (Eltahir and Yeh,1999) are likely to be missed by the approach.

Standardized groundwater indices (SGIs) have been developed to identify groundwater drought (Broomfield et al., 2013) using the similar procedure as SPIs. However, as noted by Broomfield at al. (2013), such indicators do not show advantages over the percentile approach because groundwater is a continuous variable in time and therefore, there is no need to calculate SGIs at different accumulation periods as for SPIs.

Groundwater drought can also be identified using a threshold value approach (Peters et al., 2006). This results in binary drought or non-drought conditions being identified but does not provide information on drought severity (Broomfield et al., 2013). Further, variable thresholds for different months of the year may be necessary to account for groundwater seasonality.

4. Characteristics of groundwater drought at the global domain

Characteristics of groundwater drought are not well understood at the global domain due to limited spatial coverage of available *in situ* well records. Here we use CLSM simulated groundwater to examine groundwater drought and its relationship with climate and other factors. Specifically, we use the SPI accumulation period at which maximum correlation between SPIs and groundwater is reached to examine temporal scales of groundwater and groundwater drought (Li and Rodell, 2015; Broomfield et al., 2015; Li et al., 2019b). For this analysis, CLSM simulated groundwater storage from the climatology run (see section 3.1) which spans 1948 to 2014 was converted to standardized anomalies (with monthly mean removed). The Princeton precipitation data (Sheffield et al., 2006) used to force the climatology run were converted to

SPIs at 3 to 36-month accumulation periods. Use of the longer record of simulated groundwater (as opposed to that from GRACE/GRACE-FO data assimilation) provides more robust characterization of the relationship between groundwater and large-scale climate variability.

CLSM simulated groundwater correlates strongly with SPI3 in certain tropical and subtropical regions including the northern South America, the east-central Africa, Indonesia and the southeastern Australia, where the climate is both wet and strongly affected by the El-Nino Southern Oscillation (ENSO, Dai and Wigley, 2000). Abundant precipitation in these regions leads to wet soils and shallow water tables and hence rapid groundwater recharge.

As the SPI accumulation period increases from 3 to 12 months, the correlation between SPI and CLSM groundwater storage increases in most regions where the influence of ENSO is weak. Simulated groundwater is best correlated with SPI12 over a large portion of the global land (Figs.8d and 8e). Longer SPI integration periods are best correlated with groundwater storage in high latitude regions where groundwater recharge is dominated by spring snowmelt (Perez-Valdivia et al., 2012), while shorter integration periods are best correlated in the tropics. Russo and Lall (2017) and Asoka et al. (2017) also identified 12 months as the optimal correlation scale based on *in situ* groundwater data in the U.S. and India, respectively. Broomfield et al. (2015) reported that regional groundwater showed the strongest correlation with SPI12, even though the sub-regional groundwater correlated strongly with SPIs at 4 to 17-month periods depending on vadose thickness. The impact of seasonal variation of ET on groundwater (Eltahir and Yeh, 1999) is likely to be a factor in most cases excepting cold lands. In the Sahara Desert, where groundwater recharge is nearly zero, a low correlation is observed at all time scales.

As discussed earlier, temporal variability of groundwater may be affected by hydrogeological conditions which are represented in a simplified way in CLSM by a bedrock depth parameter. In

general, the deeper the bedrock, the deeper the water table and the slower the groundwater response. The impact of the bedrock depth can be observed in the Great Plains of the U.S. and in Northwestern Europe (Figs. 8e and 8f), where CLSM bedrock is particularly deep and groundwater correlates most strongly with SPI periods that often exceed 24 months. However, the effect of bedrock depth may be overridden by a wet climate, in which abundant precipitation keeps the water table shallow, thus enabling rapid recharge. This is the case in Central Africa, Southeast China, and Indonesia, where the bedrock is deep but CLSM groundwater correlates most strongly with short-period SPIs. In these regions, groundwater drought can develop during short-term precipitation deficits but can recover quickly once precipitation increases.

5. Discussions and future research

Groundwater drought is caused by an extended precipitation deficit and may be exacerbated by evaporative demand. Because it is driven by atmospheric conditions, it only occurs in aquifers that are in communication with the surface hydrology, i.e., unconfined and semi-confined aquifers. Due to its lagged relationship with low frequency hydrometeorological phenomena (Eltahir and Yeh, 1999), groundwater drought differs from agricultural and hydrological droughts and must be evaluated separately.

Analyses of *in situ* groundwater data reveal the strong influence of precipitation variations. In the northeastern U.S. and the Mississippi River basin, regional groundwater storage changes lag precipitation by 2-4 months with maximum correlation generally exceeding 0.5, suggesting that groundwater drought occurs often even in the humid and semi-humid climates. However, individual groundwater drought events may persist longer, well after precipitation has returned to normal, as observed in well time series from the Missouri sub-basin and in Illinois (see Eltahir and Yeh, 1999). Subsurface hydrogeological conditions also affect groundwater drought onset

and evolution through their partial control on recharge rates, but they can also be nullified in wet climates.

Considering the scarcity of groundwater observations around the globe, simulated or assimilated groundwater time series from hydrological models are a valuable alternative that has proven useful for global-scale groundwater drought monitoring. Forced by observation-based meteorological fields and constrained by water and energy balance equations, hydrological models are capable of simulating groundwater response to near-surface processes including precipitation and evapotranspiration, with reasonable skill (Li et al., 2019a). In particular, LSMs, which are both physically based and computationally efficient, can run with high spatial and temporal resolution at the continental to global scales. Assimilation of GRACE and GRACE-FO data into LSMs provides much needed constraints on global groundwater simulation, especially in data scarce regions where uncertainty in forcing and parameter data are likely high. Data assimilation also enables disaggregation of GRACE/GRACE-FO derived TWS so that groundwater and the other components of TWS can be evaluated independently. Recent studies have demonstrated that GRACE/GRACE-FO data assimilation into CLSM reduces uncertainty in simulated groundwater time series at both regional and point scales (Li et al., 2019b), while also filling observational gaps and extending the information to near-real time. These improvements enable more reliable and comprehensive groundwater drought monitoring.

One of the significant limitations of LSMs such as CLSM is that they typically do not simulate relevant anthropogenic activities such as groundwater pumping and irrigation, though such capabilities are now under development (e.g., Nie et al., 2019). This has limited the ability of data assimilation to properly downscale GRACE observations in regions with intensive withdrawals such as the North China Plain (Li et al., 2019b). Consequently, a resulting

GRACE/GRACE-FO-based groundwater drought indicator cannot separate drought conditions associated with climate variability from water depletion caused by groundwater extraction. While the need to account for anthropogenic activities in drought and water management has drawn considerable attention (e.g., Van Loon et al., 2017; Nie et al., 2018), doing so would present other challenges to drought identification. For instance, a percentile-based groundwater drought indicator would always yield exceptional drought in areas where intensive groundwater withdrawals have caused a long-term decline in the water table, such as California's Central Valley. This is due to the fact that most statistical methods for quantifying drought severity assume stationarity, while groundwater depletion is a non-stationary behavior. Under these circumstances, a separate system for quantifying anthropogenic impacts on groundwater (Van Loon & Van Lanen, 2013) may be an alternative for understanding and communicating groundwater depletion. However, non-stationary statistical methods will continue to be needed to analyze groundwater drought, as climate change may also lead to non-stationary behaviors in groundwater.

Another challenge to using LSMs and other global scale hydrological models for groundwater drought monitoring is that they do not simulate lateral flows of groundwater. Therefore, they may misrepresent groundwater dynamics in aquifers such as those located at the base of mountains, which are mainly recharged by lateral flow from the adjacent mountain blocks (Wilson and Guan, 2004; Markovich et al., 2019). GRACE/GRACE-FO generally does not resolve water storage changes at these small scales. Dynamic flow models are needed to simulate three-dimensional flows and to account for complex geological and hydrogeological conditions and processes such as preferential flow and focused recharge (Wilson and Guan, 2004). While significant improvements have been made, including coupling sophisticated

groundwater flow models with LSMs (Maxwell et al., 2014), it remains a challenge to calibrate such models for global scale simulation due to the need for more hydrogeological information and *in situ* well data (de Graaf et al., 2015, 2017). Future studies should focus on combining successful modeling and data assimilation approaches and using all available data for parameterization, calibration, and evaluation, in order to achieve optimal global groundwater drought monitoring.

Figures

Figure 1. Locations of groundwater wells in Long Island (LI), New Jersey (NJ), Massachusetts (MA), Pennsylvania (PA) and the four sub-basins of the Mississippi River: the Upper Mississippi (Up-Mis), the Ohio-Tennessee (Oh-Tn), the combined Red River and Lower Mississippi (Red-LM) and the Missouri. Numbers in parentheses indicate the number of wells in each region.

Figure 2. Time series of monthly groundwater storage anomalies at individual wells (gray lines), their regional mean (black line) in the eight regions. The top bars represent averaged monthly NLDAS-2 precipitation. **r** represents correlation between regional mean groundwater and NLDAS-2 precipitation. Average groundwater depth in each region is also provided.

Figure 3. Monthly seasonal cycles of regional mean groundwater storage anomalies and precipitation (gray bars) for the eight study regions.

Figure 4. Time series of non-seasonal groundwater storage (black lines) and precipitation (blue lines, smoothed with 6-month running average) anomalies in the eight regions. Numbers represent maximum correlation and the lag of maximum correlation.

Figure.5 Monthly non-seasonal TWS from CLSM open loop simulation, GRACE data assimilation, GRACE for August, 2015 (top three panels) and region averaged monthly non-seasonal TWS time series in Southern Germany (identified by the black square box in the top three panels). GRACE/GRACE FO data are based on the CSR RL06 product.

Figure.6 Monthly non-seasonal groundwater storage anomalies from the open loop, GRACE data assimilation (DA) and in situ data in the four Mississippi sub-basins and four northeast U.S. regions. RMSE and correlation between in situ and simulated (DA and open loop) groundwater time series (in parentheses) are provided.

457 Figure.7 GRACE-based global groundwater wetness percentile maps where higher percentiles
458 represent wetter conditions. Based on the U.S Drought Monitor, drought severity can be
459 categorized using the following thresholds: exceptional drought: <=2%, extreme drought: 2-5%,
460 severe drought: 5-10%, moderate drought: 10-20%, abnormally dry: 20-30%.

461 Figure.8 Correlation between the CLSM simulated groundwater and SPI3, SPI6 and SPI12 (a, b
462 and c), maximum correlation between groundwater and SPIs, month of maximum correlation and
463 CLSM bedrock depths (d, e and f).

464

465

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. Geophy. Res. Lett., 32, L18405.

Anderson, W. P., Jr., & Emanuel, R. E. (2008). Effect of interannual and interdecadal climate oscillations on groundwater in North Carolina. Geophys. Res. Lett. 35, L23402, doi:10.1029/2008GL036054.

Asoka, A., Gleeson, T., Wada, Y. and Mishra, V. (2017). Relative contribution of monsoon precipitation and pumping to changes in groundwater storage in India. Nature Geoscience, 10(2), pp.109-117.

Asong, Z. E., Wheater, H. S., Bonsal, B., Razavi, S., and Kurkute, S. (2018). Historical drought patterns over Canada and their teleconnections with large-scale climate signals, Hydrol. Earth Syst. Sci., 22, 3105–3124, https://doi.org/10.5194/hess-22-3105-2018.

Barco, J., Hogue T. S., Girotto, M., Kendall, D. R. & Putti, M. (2010). Climate signal propagation in southern California aquifers. Water Resour. Res. 46, W00F05(2010), doi:10.1029/2009WR008376.

Bloomfield, J. P. and Marchant, B. P. (2013). Analysis of ground-water drought building on the standardised precipitation index approach, Hydrol. Earth Syst. Sci., 17, 4769–4787, doi:10.5194/hess-17-4769-2013.

Bloomfield, J. P., Marchant, B. P., Bricker, S. H., & Morgan, R. B. (2015). Regional analysis of groundwater droughts using hydro-graph classification, Hydrol. Earth Syst. Sci., 19, 4327–4344, doi:10.5194/hess-19-4327-2015.

Changnon, S. A., Jr. (1987). Detecting drought conditions in Illinois, ISWS/CIR-169-87, 36 pp., Ill. State Water Surv., Champaign.

Chen, J. L., Wilson, C. R., Tapley, B. D., Yang, Z. L., & Niu, G. Y. (2009). 2005 drought event in the Amazon River basin as measured by GRACE and estimated by climate models. Journal of Geophysical Research: Solid Earth, 114(B5).

Condon, L. E., Atchley, A. L., & Maxwell, R. M. (2020). Evapotranspiration depletes groundwater under warming over the contiguous United States. Nature Communications, 11(1), 1-8.

Cosgrove, B. A., Lohmann, D., Mitchell, K. E., Houser, P., Wood. E.F. & et al. (2003). Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project, J. Geophys. Res.,108(D22), 8842, doi:10.1029/2002JD003118, 2003.

Cunningham, S.C., Thomson, J.R., Mac Nally, R., Read, J., & Baker, P.J. (2011). Groundwater change forecasts widespread forest dieback across an extensive floodplain system, Freshwater Biol., 56 (2011), 1494-1508.

Dee, D.P, Uppala S.M., Simmons A.J., Berrisford P., Poli P., Kobayashi S., et al. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc, 137: 553 – 597. DOI:10.1002/qj.828.

de Graaf, I. D., Sutanudjaja, E. H., Van Beek, L. P. H., & Bierkens, M. F. P. (2015). A high-resolution global-scale groundwater model. Hydrology and Earth System Sciences, 19(2), 823-837.

de Graaf, I.E.M., van Beek, R.L.P.H., Gleeson, T., Moosdorf, N., Schmitz, O., Sutanudjaja, E.H. & Bierkens, M.F.P. (2017). A global-scale two-layer transient groundwater model: Development and application to groundwater depletion. Advances in Water Resources, (102), 53-67.

Döll, P., Schmied, H. M., Schuh, C., Portmann, F. T. & Eicker, A. (2014). Global-scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites, Water Resour. Res., 50, 5698–5720, doi:10.1002/2014WR015595.

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., et al. (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.

Famiglietti, J.S., & Rodell, M. (2013). Water in the Balance, Science, 340 (6138), 1300-1301, doi: 10.1126/science.1236460.

Famiglietti, J.S. (2014). The global groundwater crisis. Nature Climate Change, 4(11), 945-948, doi:10.1038/nclimate2425.

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.

Gardner, A. S., Moholdt, G., Wouters, B., Wolken, G. J., Burgess, D. O., Sharp, M. J., Cogley, J. G., Braun, C. & Labine, C. (2011). Sharply increased mass loss from glaciers and ice caps in the Canadian Arctic Archipelago, Nature, 473, 357-360.

Girotto, M., de Lannoy, G.J.M., Reichle, R. & Rodell, M. (2016). Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model, Wat. Resour. Res., 52, 4164–4183, doi:10.1002/2015WR018417, 2016.

Girotto, M., De Lannoy, G.J.M., Reichle, R.H., Rodell, M., Draper, C., Bhanja, S. N. & Mukherjee, A. (2017). Benefits and pitfalls of GRACE data assimilation: a case study of terrestrial water storage depletion in India, Geophys. Res. Lett., 44, 4107-4115 doi:10.1002/2017gl072994, 2017.

Gleeson, T., Moosdorf, N., Hartmann,J. and van Beek, L.P.H. (2014). A glimpse beneath earth's surface: GLobal HYdrogeology MaPS (GLHYMPS) of permeability and porosity. Geophysical Research Letters, 41: 2014GL059856 doi: 10.1002/2014gl059856.

Green, T.R., Taniguchi, M., Kooi, H., Gurdak, J. J., Allen, D.M., Hiscock, K.M., Treidel H. and Aureli, A. (2011). Beneath the surface of global change, Impacts of climate change on groundwater. Journal of Hydrology 405, 3, 532-560.

Gurdak, J. J., Hanson, R.T., McMahon, P.B., Bruce, B.W., McCray, J.E., Thyne, G.D. & et al. (2007). Climate variability controls on unsaturated water and chemical movement, High Plains Aquifer, USA. Vadose Zone J. 6, 533–547.

Hanson, R. T., Dettinger, M. D. & Newhouse, M. W. (2006). Relations between climatic variability and hydrologic time series from four alluvial basins across the southwestern United States. Hydrogeol. J. 14, 1122–1146.

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.

Hughes, J. D., Petrone, K. C. & Silberstein, R. P. (2012). Drought, groundwater storage and stream flow decline in southwestern Australia, Geophys. Res. Lett., 39, L03408, doi:10.1029/2011GL050797.

Kløve, B., Ala-Aho, P., Bertrand, G., Gurdak, J. J., Kupfersberger, H., Kværner, J., & Uvo, C. B. (2014). Climate change impacts on groundwater and dependent ecosystems. Journal of Hydrology, 518, 250-266.

Jiménez Cisneros et al. Freshwater resources. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 229-269, 2014.

Kath, J, Reardon-Smith, K., Le Brocque, A. F., Dyer, F. J., Dafny, E., Fritz, L. & Batterham, M. (2014). Groundwater decline and tree change in floodplain landscapes: Identifying non-linear threshold responses in canopy condition. Global Ecology and Conservation, 2, 148–16.

572 Koirala, S., Yeh, P. J.-F., Hirabayashi, Y., Kanae, S. and Oki, T. (2014). Global-scale land
573 surface hydrologic modeling with the representation of water table dynamics, Journal of
574 Geophysical Research – Atmospheres, 118, doi: 10.1002/2013JD020398.

575 Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar. (2000). A catchment-
576 based approach to modeling land surface processes in a general circulation model 1. Model
577 structure. Journal of Geophysical Research-Atmospheres, 105: 24809-24822 [DOI:
578 10.1029/2000JD900327].

579 Koster, R. D., Guo, Z., Dirmeyer, P., Bonan,G., Chan,E., Cox,P. & et al. (2006). GLACE: The
580 Global Land-Atmosphere Coupling Experiment. Part I: Overview. J. Hydrometeor., 7 (4): 590-
581 610 [10.1175/JHM510.1].

582 Kumar, S. V., Zaitchik, B. F., Peters-Lidard, C., Rodell, M., Reichle, R., Li, B., et al. (2016).
583 Assimilation of Gridded GRACE Terrestrial Water Storage Estimates in the North American
584 Land Data Assimilation System, 2016: J. Hydrometeoro., 17(1951-1972), DOI: 10.1175/JHM-D-
585 15-0157.1.

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

589 Li, B., M. Rodell, B. F. Zaitchik, et al. (2012). Assimilation of GRACE terrestrial water storage
590 into a land surface model: Evaluation and potential value for drought monitoring in western and
591 central Europe. Journal of Hydrology, 446-447: 103-115 [10.1016/j.jhydrol.2012.04.035]

592 Li, B., & Rodell, M. (2015). Evaluation of a model-based groundwater drought indicator in the
593 conterminous U.S. Journal of Hydrology, 526: 78-88 [10.1016/j.jhydrol.2014.09.027]

594 Li, B., Rodell, M. & Famiglietti, J.S. (2015). Groundwater variability across temporal and spatial
595 scales in the central and northeastern U.S. Journal of Hydrology, 525: 769-780
596 [10.1016/j.jhydrol.2015.04.033]

597 Li, B., M. Rodell, J. Sheffield, E. Wood, and E. Sutanudjaja. (2019a). Long-term, non-
598 anthropogenic groundwater storage changes simulated by three global-scale hydrological
599 models. Scientific Reports, 9 (1): 10746 [10.1038/s41598-019-47219-z]

600 Li, B., M. Rodell, S. Kumar, et al. (2019b). Global GRACE Data Assimilation for Groundwater
601 and Drought Monitoring: Advances and Challenges." Water Resources Research,
602 2018WR024618 [10.1029/2018wr024618].

603 Long, D., Scanlon, B. R., Longuevergne, L., Sun, A. Y., Fernando, D. N., & Save, H. (2013).
604 GRACE satellite monitoring of large depletion in water storage in response to the 2011 drought
605 in Texas. Geophysical Research Letters, 40(13), 3395-3401.

606 Markovich, K., H, Manning, A. H., Condon, L. E., & McIntosh, J. C. (2019). Mountain-block
607 recharge: A review of current understanding. Water Resources Research, 55, 8278–8304.
608 https://doi.org/10.1029/ 2019WR025676.

609 Maxwell, R., Kollet, S. (2008). Interdependence of groundwater dynamics and land-energy
610 feedbacks under climate change. Nature Geosci., 665–669, doi:10.1038/ngeo315.

611 Maxwell, R. M., et al. (2014). Surface-subsurface model intercomparison: A first set of
612 benchmark results to diagnose integrated hydrology and feedbacks, Water Resour. Res.,50,
613 1531–1549, doi:10.1002/2013WR013725.

614 Meixner, T., A.H. Manning, D.A. Stonestrom, D.M. Allen, H. Ajami, K.W. Blasch, A.E.
615 Brookfield, C.L. Castro, J.F. Clark, D.J. Gochis, A.L Flint, K.L. Neff, R. Niraula, M. Rodell,
616 B.R. Scanlon, K. Singha, and M.A. Walvoord (2016). Implications of projected climate change
617 for groundwater recharge in the western United States, J. Hydrology, 534, 124-138,
618 doi:10.1016/j.jhydrol.2015.12.027.

619 Mishra, A. K. and Singh, V. P. (2010). A review of drought concepts, J. Hydrology, 391(2010),
620 202-216, http://dx.doi.org/10.1016/j.jhydrol.2010.07.012.

621 Mo, K. C. (2008) Model-based drought indices over the United States, J. Hydrometeor., 9, 1212-
622 1230, DOI: 10.1175/2008JHM1002.1.

623 Mueller et al. (2011). Evaluation of global observations-based evapotranspiration datasets and
624 IPCC AR4 simulations. Geophys. Res. Lett. 38, L06402(2011), doi:10.1029/2010GL046230.

625 Mukherjee, A., Scanlon, B., Aureli, A., Langan, S., Guo, H., and McKenzie, A. (2020). Global
626 Groundwater: Source, Scarcity, Sustainability, Security and Solutions. Elsevier, 1st. Ed., ISBN.
627 9780128181720.

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

631 Nie, W., B. Zaitchik, M. Rodell, S.V. Kumar, M.C. Anderson, and C. Hain (2018). Groundwater
632 withdrawals under drought: reconciling GRACE and land surface models in the United States
633 High Plains aquifer, Wat. Resour. Res., 54, doi:10.1029/2017WR022178.

634 Nie, W., B.F. Zaitchik, M. Rodell, S.V. Kumar, K.R. Arsenault, B. Li, and A. Getirana (2019).
635 Assimilating GRACE into a land surface model in the presence of an irrigation-induced
636 groundwater trend, Wat. Resour. Res., 55 (12), 11274-11294, doi:10.1029/2019WR025363.

637 Niraula, R., T. Meixner, F. Dominguez, M. Rodell, H. Ajami, D. Gochis, and C. Castro (2017).
638 How might recharge change under projected climate change in western US? Geophys. Res. Lett.,
639 44, 10,407–10,418, doi:10.1002/2017GL075421.

640 Peters, E., Bier, G., van Lanen, H. A. J., and Torfs, P. J. J. F. (2006). Propagation and spatial
641 distribution of drought in a groundwater catchment, J. Hydrol., 321, 257–275.

642 Perez-Valdivia, C., D. Sauchyn, and J. Vanstone (2012). Groundwater levels and teleconnection
643 patterns in the Canadian Prairies, Water Resour. Res., 48, W07516,
644 doi:10.1029/2011WR010930.

Rodell, M., and J. S. Famiglietti (2011). An analysis of terrestrial water storage variations in Illinois with implications for the Gravity Recovery and Climate Experiment (GRACE), Wat. Resour. Res., 37, 1327-1340, doi:10.1029/2000WR900306.

Rodell, M., and J. S. Famiglietti (2002). The potential for satellite-based monitoring of groundwater storage changes using GRACE: The High Plains aquifer, central U.S., J. Hydrol., 263, 245-256.

Rodell, M., Chen, J., Kato, H., Famiglietti, J. S., Nigro J. & Wilson, C. R. (2007). Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE, Hydrologeology Journal, 15:159-166, doi: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., Famiglietti, J.S., Wiese, D.N., Reager, J.T., Beaudoing, H.K., Landerer, F.W. & Lo, M.-H. (2018). Emerging trends in global freshwater availability, Nature, 557, 651-659, doi:10.1038/s41586-018-0123-1, 2018.

Rowlands, D. D., Luthcke, S. B., Klosko, S. M., Lemoine, F. G. R., Chinn, D. S., McCarthy, et al. (2005). Resolving mass flux at high spatial and temporal resolution using GRACE intersatellite measurements. Geophys. Res. Lett., 32, L04310, doi:10.1029/2004GL021908.

Russo, T.A. and Lall, U. (2017). Depletion and response of deep groundwater to climate-induced pumping variability. Nature Geoscience, 10(2), p.105, DOI: 10.1038/NGEO288.

Sheffield, J., Goteti, G. & Wood, E. F. (2006). Development of a 50-yr high-resolution global dataset of meteorological forcings for land surface modeling. J. Clim. 19, 3088–3111.

Sheffield, J., B. Livneh, and E. Wood (2012). Representation of Terrestrial Hydrology and Large Scale Drought of the Continental US from the North American Regional Reanalysis. J. Hydrometeor. doi:10.1175/JHM-D-11-065.1.

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. Geophysical Research Letters, 33(16), https://doi.org/10.1029/2006GL026962.

Stromberg, J. C., Tress, J.A., Wilkins S.D. & Clark, S.D. (1992). Response of velvet mesquite to groundwater decline, J. Arid Environ., 23 (1992), pp. 45-58.

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, oi:10.1029/2009JD013035.

Sutanudjaja et al. (2018). PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model, Geosci. Model Dev., 11, 2429–2453, https://doi.org/10.5194/gmd-11-2429-2018.

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

Taylor, R.G., B. Scanlon, P. Döll, M. Rodell, R. van Beek, Y. Wada, L. Longuevergne, M. Leblanc, J.S. Famiglietti, M. Edmunds, L. Konikow, T.R. Green, J. Chen, M. Taniguchi, M.F.P Bierkens, A. MacDonald, Y. Fan, R.M. Maxwell, Y. Yechieli, J.J. Gurdak, D. Allen, M. Shamsudduha, K. Hiscock, P.J.-F. Yeh, I. Holman, and H. Treidel (2013). Ground water and climate change, Nature Climate Change, 3(4), 322-329, doi:10.1038/nclimate1744.

Thomas, B.F., Famiglietti, J.S., Landerer, F.W., Wiese, D.N., Molotch, N.P. and Argus, D.F., (2017). GRACE groundwater drought index: Evaluation of California Central Valley groundwater drought. Remote Sensing of Environment, 198, 384-392, https://doi.org/10.1016/j.rse.2017.06.026.

Van Loon, A. & Van Lanen, H. A. J. (2013). Making the distinction between water scarcity and drought using observation-modeling framework, 49, 1483-1502, doi:10.1002/wrcr.20147.

Van Loon, A. F. et al. (2016). Drought in the anthropocene. Nat. Geosci. 9, 89–91.

Van Loon, A., Kuman, R. & Mishra, V. (2017). Testing the use of standardized indices and GRACE satellite data to estimate the European 2015 groundwater drought in near-real time, https://www.hydrol-earth-syst-sci.net/21/1947/2017/, doi:10.5194/hess-21-1947-2017.

Velasco, E. M., Gurdak, J. J., Dickinson, J. E., Ferré, T. P. A. & Corona, C. R. (2015). Interannual to multidecadal climate forcings on groundwater resources of the U.S. West Coast. J. Hydrol. Reg. Stud. 11, 250-265(2015), doi:10.1016/j.ejrh.2015.11.018.

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.

Wang, D. (2012) Evaluating interannual water storage changes at watersheds in Illinois based on long-term soil moisture and groundwater level data, Water Resour. Res., 48, W03502, doi:10.1029/2011WR010759.

Wilson, J. & Guan H. (2004). Mountain-Block Hydrology and Mountain-Front Recharge, Groundwater Recharge in A Desert Environment: The Southwestern United States, edited by Fred M. Phillips, James Hogan, and Bridget Scanlon, 2004, A GU, Washington, DC.

Wu, W., Geller, M. A. & Dickinson, R. E. (2002). The response of soil moisture to long-term variability of precipitation, J. Hydrometeor., 3, 604-613.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E. & et al. (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, doi:10.1029/2011JD016048.

Xu, T. & Valocchi, A. (2015). Data-driven methods to improve baseflow prediction of a regional groundwater model, Computers & Geosciences, 85(Part B), https://doi.org/10.1016/j.cageo.2015.05.016.

Yeh, P. J.-F., and Eltahir, E. A. B. (2005a). Representation of water table dynamics in a land surface scheme: 1. model development. Journal of Climate, Vol. 18, No. 12, pages 1861-1880.

Yeh, P. J.-F., and Eltahir, E. A. B. (2005b). Representation of water table dynamics in a land surface scheme: 2. subgrid heterogeneity. Journal of Climate, Vol. 18, No. 12, pages 1881-1901.

Yirdaw, S. Z., Snelgrove, K. R., & Agboma, C. O. (2008). GRACE satellite observations of terrestrial moisture changes for drought characterization in the Canadian Prairie. Journal of Hydrology, 356(1-2), 84-92.

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

Zhao, M., Velicogna, I., & Kimball, J. S. (2017). A global gridded dataset of grace drought severity index for 2002–14: Comparison with pdsi and spei and a case study of the australia millennium drought. Journal of Hydrometeorology, 18(8), 2117-2129.

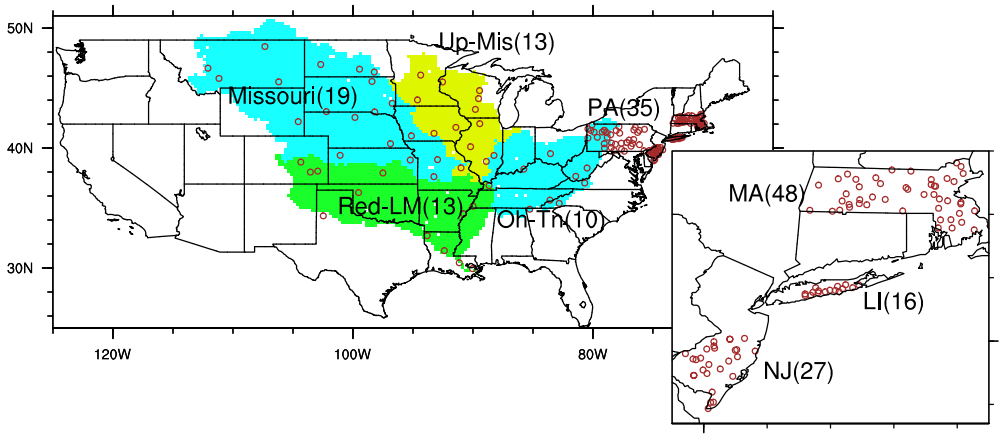

Figure 1. Locations of groundwater wells in Long Island (LI), New Jersey (NJ), Massachusetts (MA), Pennsylvania (PA) and the four sub-basins of the Mississippi river: the Upper Mississippi (Up-Mis), the Ohio-Tennessee (Oh-Tn), the combined Red River and Lower Mississippi (Red-LM) and the Missouri. Numbers in parentheses indicate the number of wells in each region.

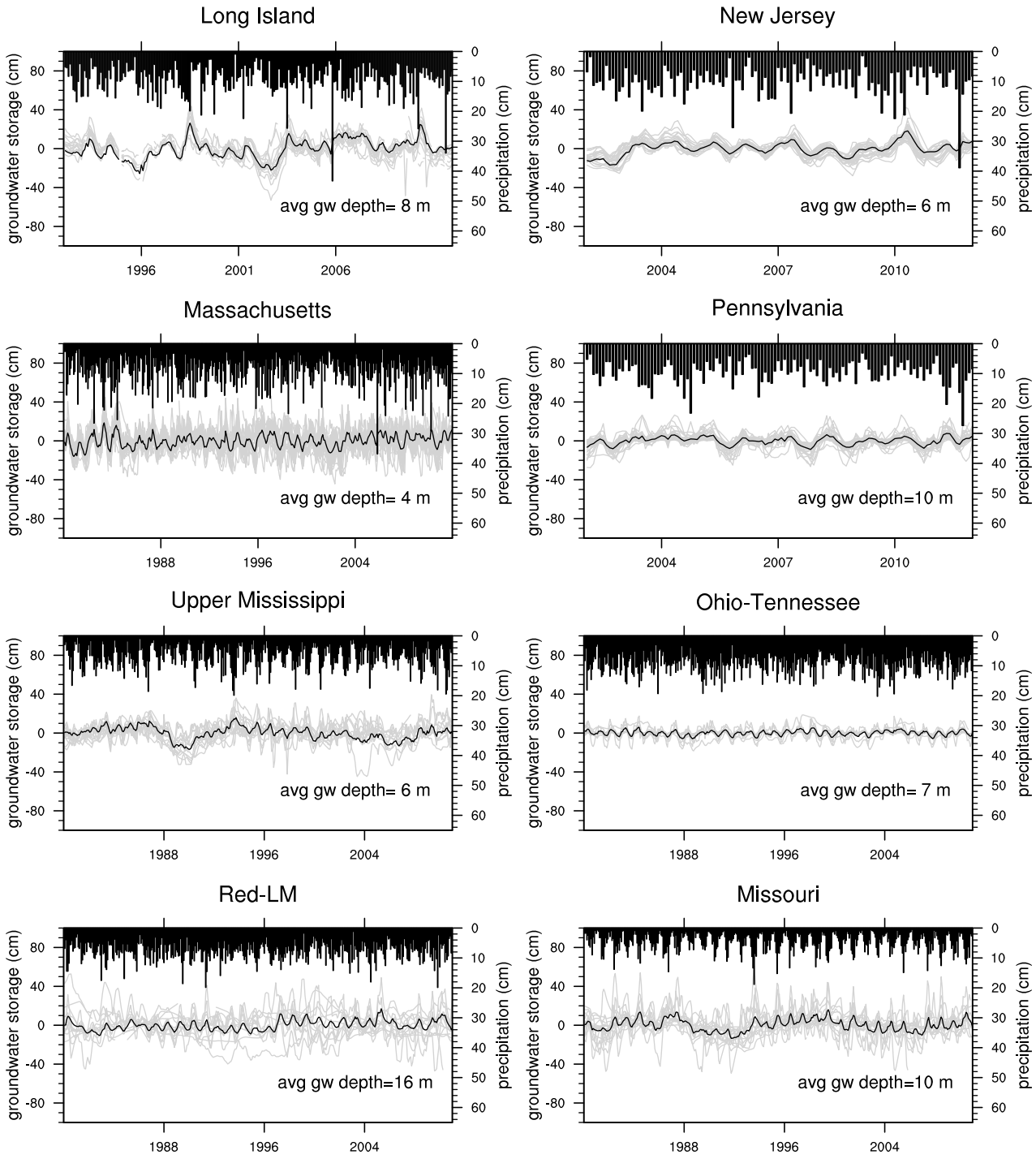

Figure 2. Time series of monthly groundwater storage anomalies at individual wells (gray lines), their regional mean (black line) in the eight regions. The top bars represent averaged monthly NLDAS-2 precipitation. r represents correlation between regional mean groundwater and NLDAS-2 precipitation. Average groundwater depth in each region is also provided.

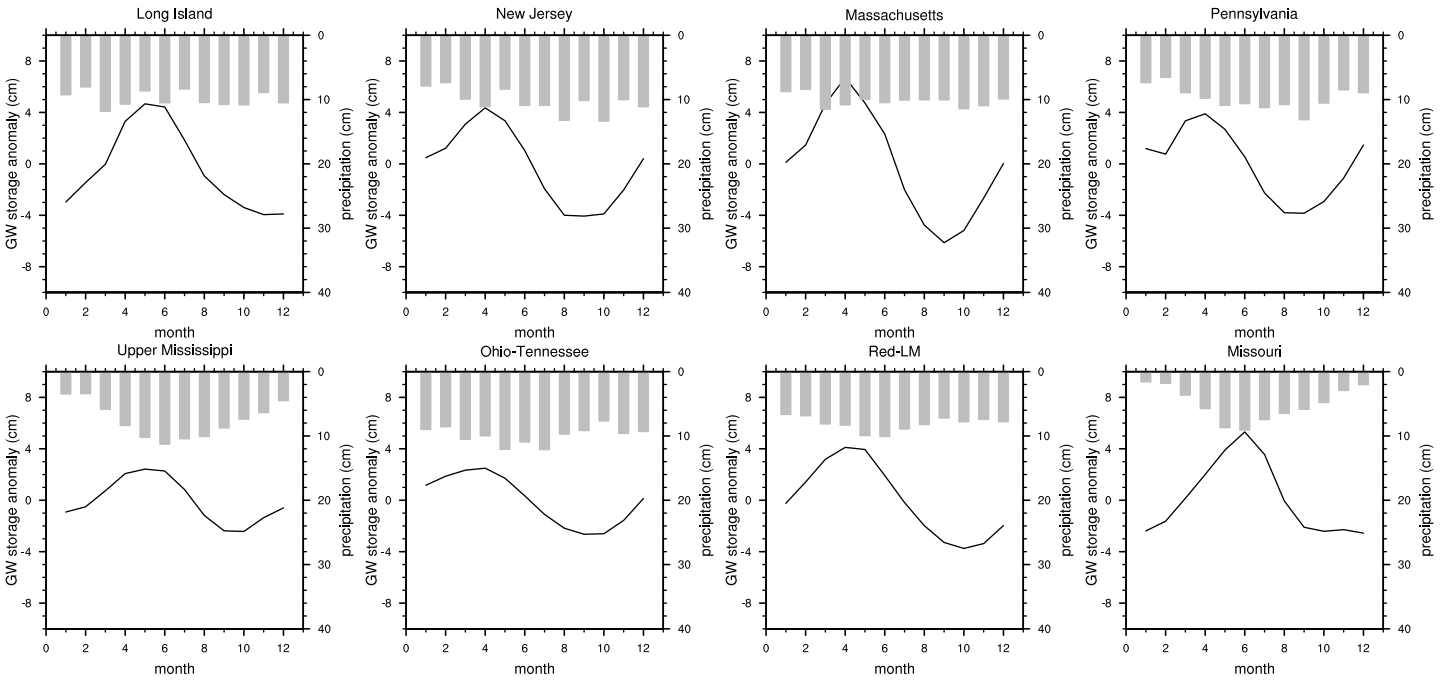

Figure 3. Monthly seasonal cycles of regional mean groundwater storage anomalies and precipitation (gray bars) for the eight study regions.

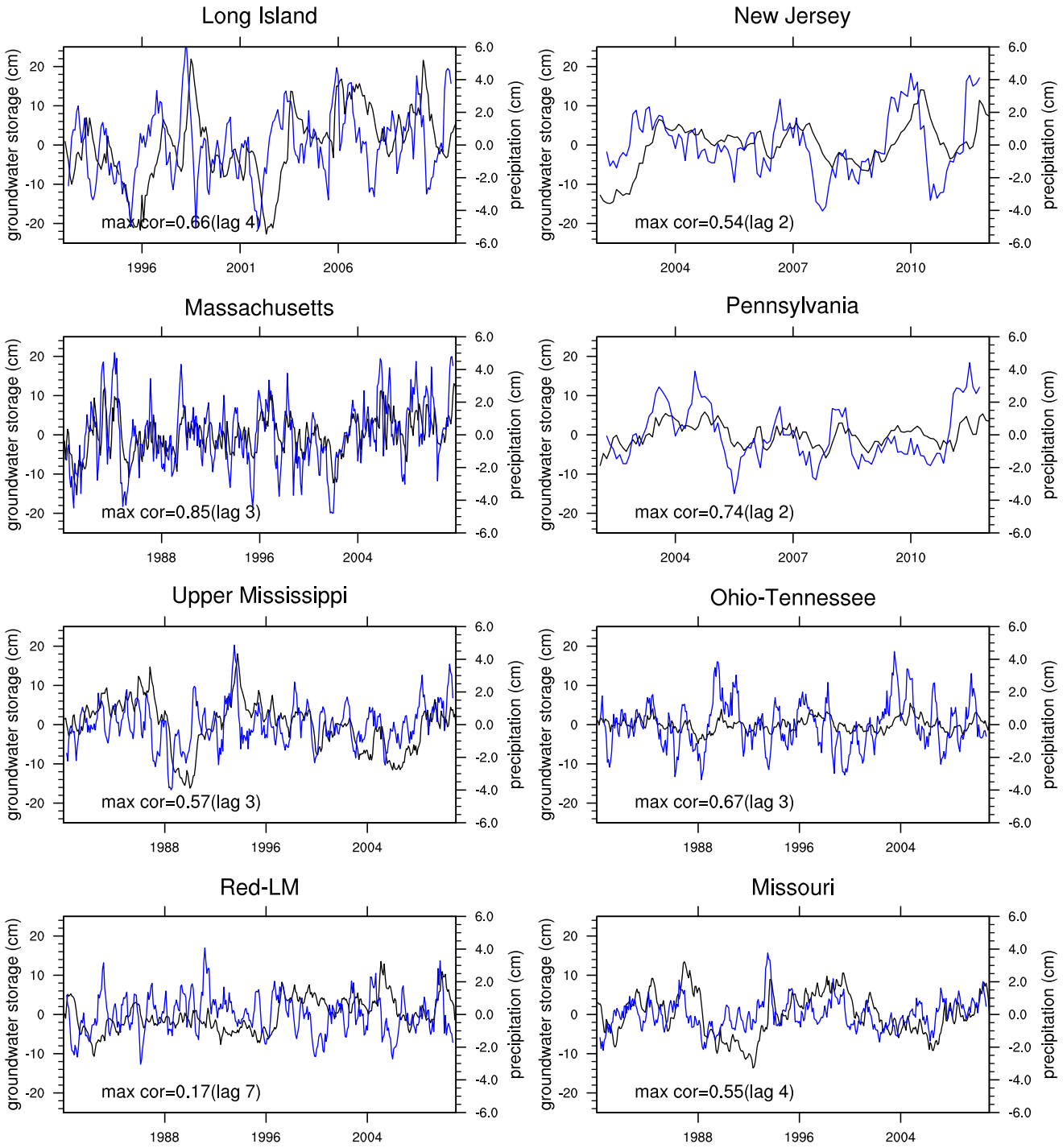

Figure 4. Time series of non-seasonal groundwater storage (black lines) and precipitation (blue lines, smoothed with 6-month running average) anomalies in the eight regions. Numbers represent maximum correlation and the lag of maximum correlation.

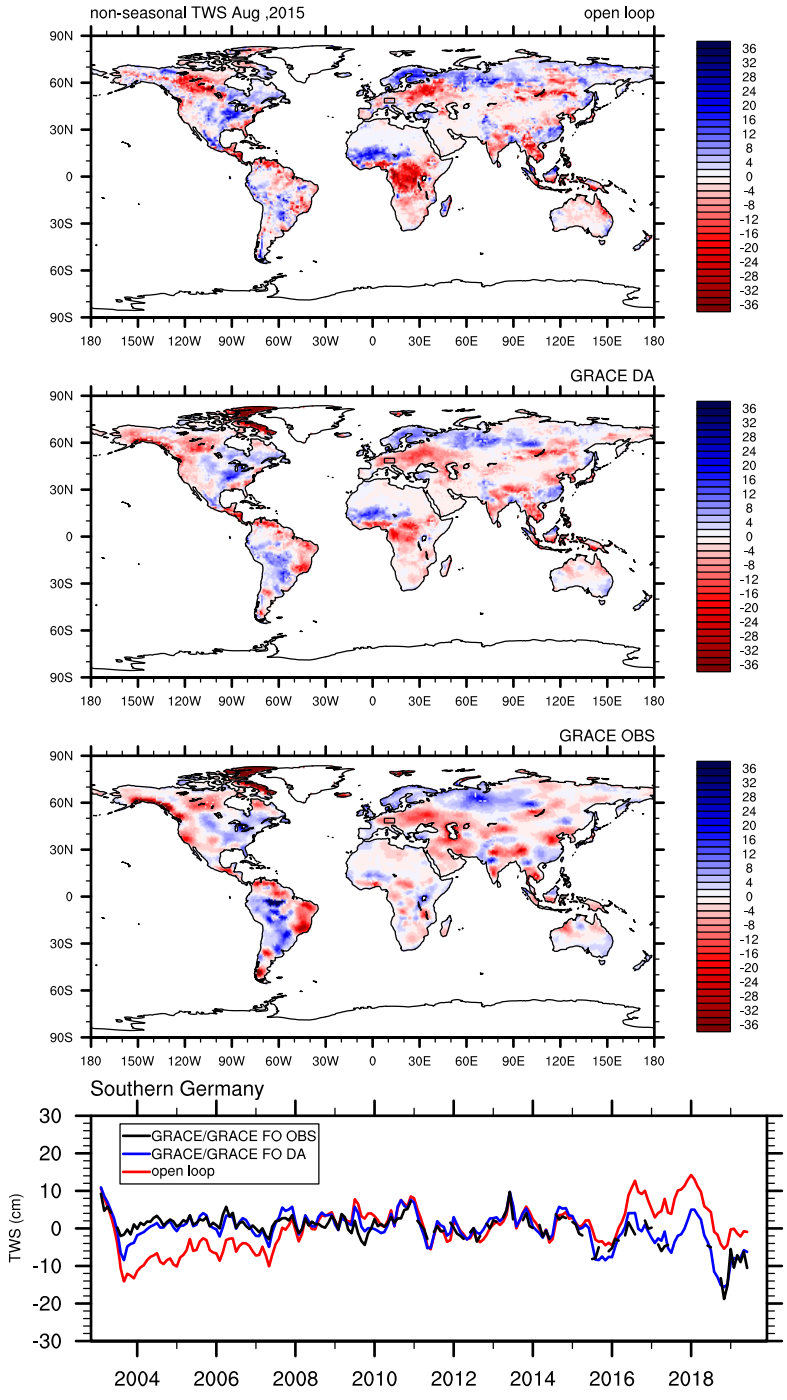

Figure.5 Monthly non-seasonal TWS from CLSM open loop simulation, GRACE data assimilation, GRACE for August, 2015 (top three panels) and region averaged monthly non-seasonal TWS time series in Southern Germany (identified by the black square box in the top three panels). GRACE/GRACE FO data are based on the CSR RL06 product.

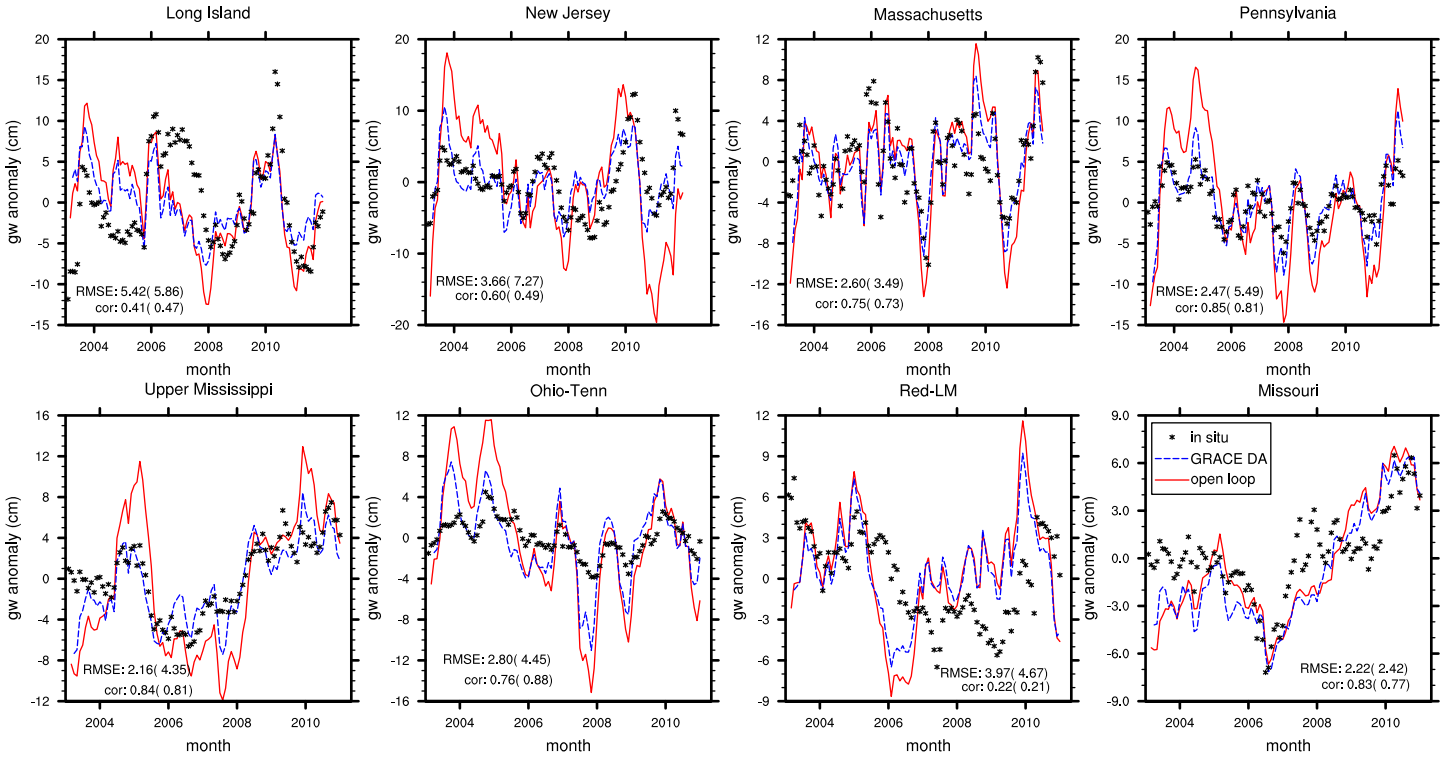

Figure.6 Monthly non-seasonal groundwater storage anomalies from the open loop, GRACE data assimilation (DA) and in situ data in the four northeast U.S. regions and the four Mississippi sub-basins. RMSE and correlation between in situ groundwater and that of DA and open loop (in parentheses) are provided.

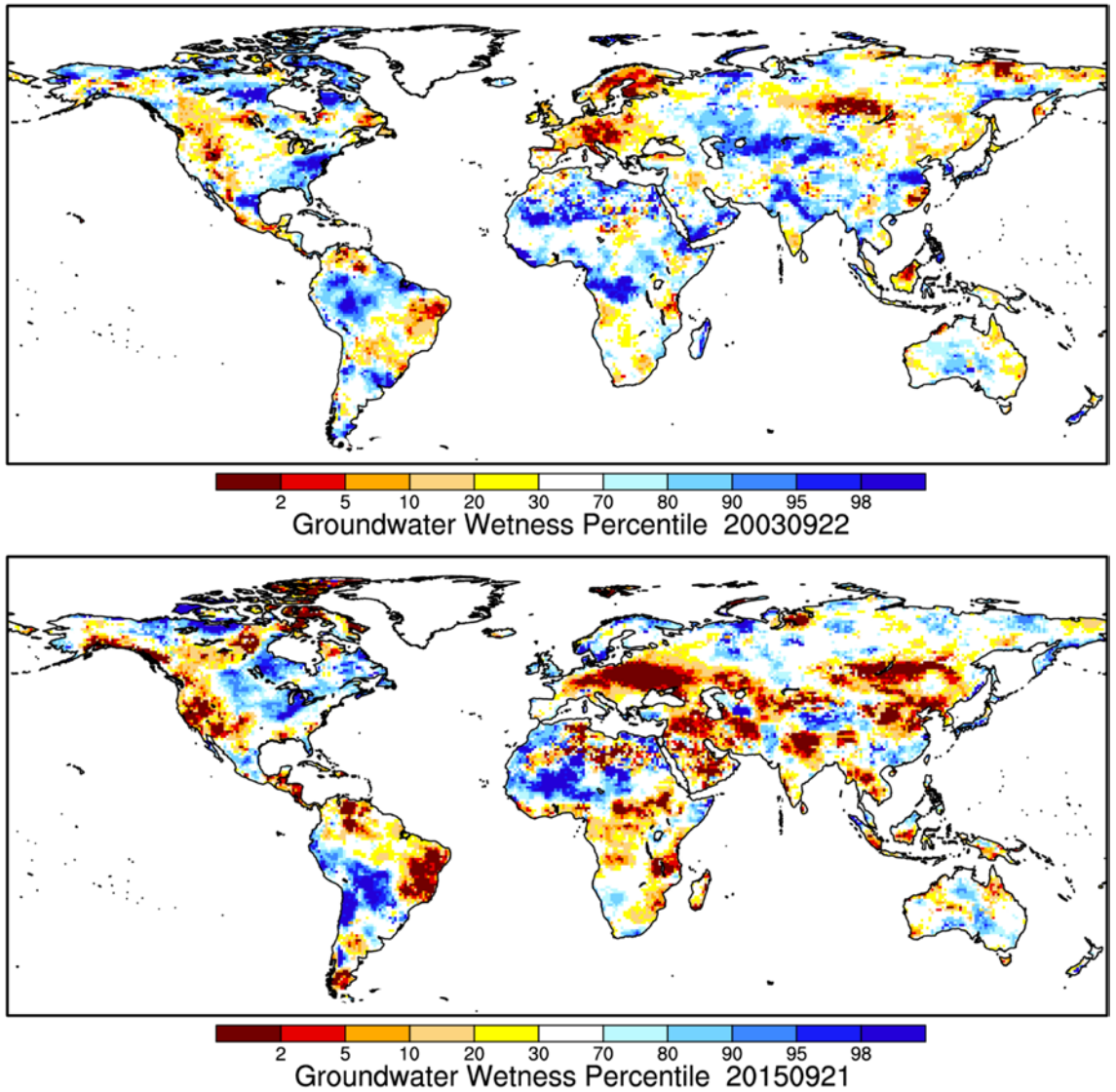

Figure.7 GRACE-based global groundwater wetness percentile maps where higher percentiles represent wetter conditions. Based on the U.S Drought Monitor, drought severity can be categorized using the following thresholds: exceptional drought: <=2%, extreme drought: 2-5%, severe drought: 5-10%, moderate drought: 10-20%, abnormally dry: 20-30%.

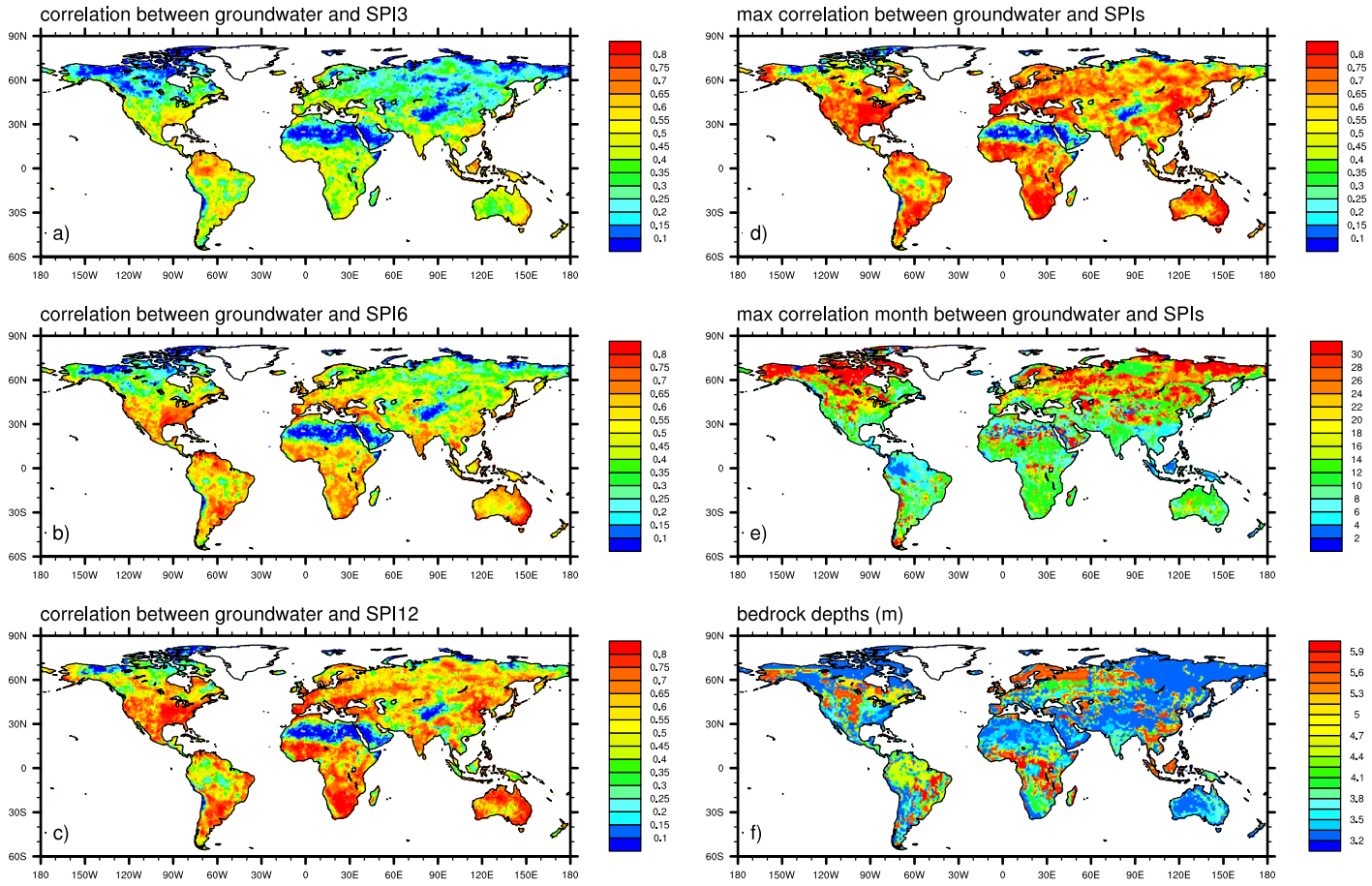

Figure.8 Correlation between the CLSM simulated groundwater and SPI3, SPI6 and SPI12 (a,b and c), maximum correlation between groundwater and SPIs, month of maximum correlation and CLSM bedrock depths (d,e and f).