

## Accepted Manuscript

Groundwater Variability across Temporal and Spatial Scales in the Central and Northeastern U.S.

Bailing Li, Matthew Rodell, James S. Famiglietti

PII: S0022-1694(15)00292-9

DOI: <http://dx.doi.org/10.1016/j.jhydrol.2015.04.033>

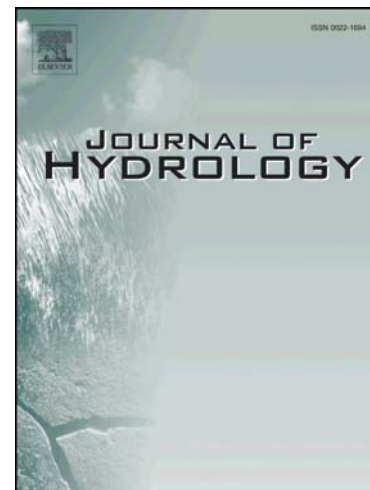
Reference: HYDROL 20392

To appear in: *Journal of Hydrology*

Received Date: 13 November 2014

Revised Date: 16 April 2015

Accepted Date: 17 April 2015



Please cite this article as: Li, B., Rodell, M., Famiglietti, J.S., Groundwater Variability across Temporal and Spatial Scales in the Central and Northeastern U.S., *Journal of Hydrology* (2015), doi: <http://dx.doi.org/10.1016/j.jhydrol.2015.04.033>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Groundwater Variability across Temporal and Spatial Scales in the Central and Northeastern  
U.S.

Bailing Li<sup>1,2</sup>, Matthew Rodell<sup>2</sup>, and James S. Famiglietti<sup>3,4,5</sup>

[1] Earth System Science Interdisciplinary Center, University of Maryland, College Park,  
Maryland

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

[3] NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA

[4] Department of Earth System Science, University of California, Irvine, CA

[5] Department of Civil and Environmental Engineering, University of California, Irvine, CA

Correspondence to:

Bailing Li ([bailing.li@nasa.gov](mailto:bailing.li@nasa.gov))

Key Points

- spatial variability of groundwater storage anomalies increases as a power function of extent
- spatial variability of groundwater storage anomalies depends on mean groundwater storage
- Seasonality of groundwater storage is similar to that of modeled recharge

## Abstract

Depth-to-water measurements from 181 monitoring wells in unconfined or semi-confined aquifers in nine regions of the central and northeastern U.S. were analyzed. Groundwater storage exhibited strong seasonal variations in all regions, with peaks in spring and lows in autumn, and its interannual variability was nearly unbounded, such that the impacts of droughts, floods, and excessive pumping could persist for many years. We found that the spatial variability of groundwater storage anomalies (deviations from the long term mean) increases as a power function of extent scale (square root of area). That relationship, which is linear on a log-log graph, is common to other hydrological variables but had never before been shown with groundwater data. We describe how the derived power function can be used to determine the number of wells needed to estimate regional mean groundwater storage anomalies with a desired level of accuracy, or to assess uncertainty in regional mean estimates from a set number of observations. We found that the spatial variability of groundwater storage anomalies within a region often increases with the absolute value of the regional mean anomaly, the opposite of the relationship between soil moisture spatial variability and mean. Recharge (drainage from the lowest model soil layer) simulated by the Variable Infiltration Capacity (VIC) model was compatible with observed monthly groundwater storage anomalies and month-to-month changes in groundwater storage.

## 1. Introduction

Aquifers are a vital source of fresh water. The United Nations estimates that about 2.5 billion people rely exclusively on groundwater for drinking water (<http://unesdoc.unesco.org/images/0022/002207/220723E.pdf>), and aquifers also provide 43% of the water used for crop irrigation worldwide (Siebert et al., 2010). Recent studies have shown that withdrawals, mostly related to irrigation, in the last several decades have led to significant declines of groundwater in many regions (Rodell et al., 2009; Wada et al., 2010; Famiglietti et al., 2011; Feng et al., 2013; Voss et al., 2013). Such declines, if not reversed, would lead to local and later regional dewatering of aquifers, which would have severe consequences for agricultural productivity and human health. Further, ecosystems may be permanently altered by reduced baseflow to streams and wetlands (Stromberg et al., 1996). Climate change is likely to exacerbate the situation in areas where precipitation rates, timing, and fraction as snowfall shift, while demands on water resources are likely to increase in a warmer environment (Green et al., 2011; Taylor et al., 2012). Improving observation and understanding of groundwater storage and its natural variability is essential if we are to preserve and better manage this precious resource in the future (Famiglietti and Rodell; 2013).

Groundwater varies slowly relative to soil moisture, surface water, and non-permanent snow cover, but it is dynamic on seasonal to interannual timescales (Rodell and Famiglietti, 2001; Alley et al., 2002; Weider and Boutt, 2010). Indeed, variations in terrestrial water storage, particularly groundwater storage, contribute to observed interannual and long term sea level changes (Konikow, 2011; Boening et al., 2012). Studies have shown that seasonal variations of

shallow groundwater are strongly influenced by climatologic variables including precipitation and evapotranspiration (Eltahir and Yeh, 1999). Groundwater storage responds to atmospheric conditions integrated over weeks to years, and its variability is known to be correlated with climate signals such as the El Nino Southern Oscillation (ENSO) in certain regions (Barco et al., 2010; Perez-Valdivia et al., 2012). As a slow varying component of the water cycle, groundwater has long “memory” and can influence the long-term trends and inter-annual variability of runoff and evapotranspiration (Istanbulluoglu et al., 2012; Wang, 2012). Accurately representing its multi-scale variability in hydrological and climate prediction models is challenging, due in part to limited knowledge of how groundwater variability scales spatially and temporally. Improved understanding of the scales of groundwater variability would benefit the development of such models.

Depending on the application, groundwater monitoring network design may benefit from knowledge of spatial variability which determines the number of wells required to quantify the regional mean groundwater storage condition at a given time. Studies on soil moisture have revealed that its spatial variability increases as a power function of extent (Famiglietti et al., 2008; Brocca et al., 2012; Li and Rodell, 2013), thus providing a mathematical form to relate spatial variability to spatial scale. No such study has been conducted for groundwater. Further, knowledge of groundwater scaling relationships would help to bridge gaps not only between field observations, but also between point and remote-sensing measurements. For instance, terrestrial water storage anomalies obtained from the Gravity Recovery and Climate Experiment (GRACE) mission have shown great promises for estimating groundwater storage changes in various regions (Yeh et al., 2006; Rodell et al., 2007; Zaitchik et al., 2008; Rodell et al., 2009; Famiglietti et al., 2011; Voss et al., 2013;) but the application of GRACE is also limited by its

low spatial resolution, which is about 150,000 km<sup>2</sup> at mid-latitudes (Rowlands et al., 2005; Swenson and Wahr, 2006). Hence there is a significant need for information that would help to interpolate between sparsely distributed well observations and GRACE based groundwater storage change estimates.

We examined the temporal and spatial variability of groundwater storage in nine regions of the U.S. based on monitoring well data archived by the USGS and the Illinois State Water Survey. In this study, “scale-dependency” refers to the dependency of spatial variability on extent, which is one of the scale triplet defined by Western and Blösch (1999) and indicates the dimension length covering all measurements. In addition to in situ data, North America Land Data Assimilation System (NLDAS-2, Xia et al., 2012a) precipitation forcing data and simulated groundwater recharge from the NLDAS-embedded VIC land surface model were analyzed. These were used to investigate the interaction between groundwater and atmospheric forcing and to corroborate inferred groundwater behaviors.

## 2. Data and Methods

Fig. 1 shows the locations of observation wells in Long Island (New York), New Jersey, Massachusetts, Pennsylvania and four sub-basins of the Mississippi River basin: the Upper Mississippi, Ohio-Tennessee, combined Red River and Lower Mississippi (hereafter referred to as “Red-LM”), and Missouri basins. The area and the number of wells within the boundary of each region are provided in Table 1. This data set has been previously used, in part or in whole, for validating GRACE derived or model estimated groundwater storage anomalies (Rodell et al., 2007; Zaitchik et al., 2008; Li and Rodell, 2014). The wells were culled from a much larger

archive through examination of the data and available metadata. Each well was determined to be open to an unconfined or semi-confined aquifer and representative of the local water table, i.e., exhibiting minimal direct effects of pumping or injections. Records from many locations were discarded due to brevity or large data gaps.

The majority of the data records were obtained from the USGS Groundwater Watch website (<http://groundwaterwatch.usgs.gov/>) and the rest from the larger USGS National Water Information System and the Illinois State Water Survey. Most of the sites logged one measurement per month; when multiple measurements were available per month, an average monthly value was used. The lengths of the regional data records range from 10 to over 30 years (Table 1) which is sufficiently long to study the seasonality of groundwater. The wells in Long Island, New Jersey, and Massachusetts are generally located in shallow sandy aquifers formed during the last glacial maximum. Most wells in Pennsylvania are located in fractured rock formations that are likely semi-confined. Wells in the Mississippi basin are installed in a diverse range of aquifer types, and their depths vary significantly. The region-averaged well depth ranges from 9 m below the surface in Massachusetts to 86 m in Red-LM, and the average depth to water varies from 4 m to 17 m (Table 1).

Because this study was only concerned with the variability of groundwater storage anomalies (departures from the long term mean) and not absolute quantity, we set the mean depth to water at each well to zero by subtracting the time series-mean from each measurement. Specific yield ( $S_y$ ) estimates are needed to convert depth-to-water levels to water storage anomalies as equivalent heights of water, but  $S_y$  was not provided in the metadata. The  $S_y$  values (see Table 1 for regional averages) were determined individually for each well based on published studies on the aquifer formation or, as a last resort, published  $S_y$  estimates for the

aquifer type. When multiple possible  $S_y$  values were found for a given well, a  $S_y$  within that range was selected based on the well depth and comparison of the dynamic range of water depths with those of neighboring wells. For each well, groundwater storage anomalies relative to the series mean were computed by multiplying the monthly depth-to-water measurements (mean removed) by the specific yield, and then taking the additive inverse (negative) of each value (because storage increases as depth-to-water decreases).

NLDAS-2 (Xia et al., 2012a) precipitation forcing data are based on daily precipitation measurements from over 10,000 gauges which are temporally disaggregated using Doppler radar images and spatially interpolated to a  $0.125^\circ$  grid that encompasses the conterminous U.S. and parts of Mexico and Canada (Cosgrove et al., 2003). NLDAS forcing spans 1979 to present and thus covers the entire period of our groundwater dataset.

Due to lack of recharge observations, drainage from the bottom of the lowest soil layer simulated by the Variable Infiltration Capacity (VIC) land surface model, driven by NLDAS, was used as an approximation of groundwater recharge. This drainage variable is often named “baseflow” or “subsurface runoff” in land surface models, which are misnomers that imply the water somehow circumvents aquifer storage and immediately enters the stream system. To avoid confusion, we eschew the model terminology and instead use the terms “drainage” and “recharge” through the rest of the text. VIC (Liang et al., 1994) simulates water and energy states (e.g., soil moisture and temperature) and fluxes (e.g., evapotranspiration and runoff) based on physical equations of the relevant processes on and within the land surface, using atmospheric forcing data (e.g., precipitation and solar radiation) to drive the model forward in time. The evolution of soil moisture is simulated in three soil layers, with 10 cm for the top layer and spatially varying depths for the lower two layers which may reach a total depth of 2.5 m at some



locations. Drainage is derived using an empirical function that depends on the wetness of the lowest soil layer and a shape parameter. VIC has been applied in a wide range of hydrological basins for modeling streamflow and other land surface processes (see Xia et al., 2012a). Within NLDAS-2 settings (forcing and spatial resolution), Xia et al. (2012b) showed that VIC produced more accurate streamflow and evapotranspiration estimates than other NLDAS-2 models.

Although VIC does not contain a groundwater component, its recharge estimates synthesize the combined effect of precipitation, evapotranspiration, surface runoff, soil wetness dynamics, and vertical flow through the soil, and thus are a reasonable substitute for groundwater recharge at model pixel and larger spatial scales and monthly to interannual temporal scales. This is not unprecedented, as Crosbie et al. (2013) used soil drainage from a coupled water, energy, and carbon model as an approximation of groundwater recharge in studying the potential impact of climate change on groundwater. Monthly NLDAS precipitation and VIC simulated recharge time series were extracted for the model pixels containing the well locations for use in the statistical calculations.

In situ groundwater at a large number of sites exhibited small but significant trends (based on the Mann-Kendall test at the 5% significance level), with the regional average trend ranging from -0.018 cm/month in Red-LM to 0.077 cm/month in New Jersey. NLDAS precipitation and VIC recharge did not have significant trends in any regions except Massachusetts, where both exhibited significant trends at some sites, but the average trend was still small (around 0.004 cm/month). Because long term linear trends can affect the correlations between groundwater storage anomalies and other variables at shorter time scales, which is the focus of this study, the best-fit linear trend was removed from groundwater storage anomalies

and recharge as a first step. We determined that removing the trend did not have a large impact on the results.

Monthly spatial means and standard deviations (representing spatial variability) of groundwater storage anomalies were calculated based on data from the observation wells in each of the nine study regions. Means and standard deviations of precipitation and recharge for each region were similarly calculated using data extracted from the pixels containing well sites. Temporal variability, including seasonality, and temporal correlations were calculated based on the regional mean time series. Unless otherwise stated, mean groundwater storage anomalies, mean recharge, and mean precipitation refer to the regional means of those fields.

Groundwater discharge (baseflow to streams, spring flow, submarine groundwater discharge, uptake by phreatophyte roots, and withdrawals from wells) was estimated for each region and month as the residual of a simple mass balance equation,

$$\text{discharge} = \text{recharge} - \Delta GW \quad (1)$$

where  $\Delta GW$  is the monthly change in groundwater storage, which was approximated as the difference between the groundwater storage anomaly at the current month and that of the previous month. Ideally, if daily groundwater observations were available,  $\Delta GW$  would have been calculated as the difference between groundwater storage anomalies on the first and last days of the month, but averaged over longer periods our approximation is reasonable. As previously described, NLDAS/VIC provided the recharge estimates.

### 3. Results

### 3.1 Spatial mean

Fig. 2 plots the time series of monthly regional mean groundwater storage anomalies, spatial standard deviations, and NLDAS precipitation. Groundwater storage anomalies from individual wells are also plotted in order to illustrate spatial variability. Correlation between monthly groundwater storage anomalies and precipitation is apparent during prolonged wet and dry events. For instance, the well-known drought that afflicted the Midwest and Northern Great Plains during 1987-1989 produced large negative storage anomalies in the Upper Mississippi and Missouri basins. The relationship between mean groundwater storage anomaly and spatial variability will be discussed in more detail later, but upon quick inspection it can be seen that local maxima of spatial variability are coincident with both high and low mean anomalies.

Table 2 shows that the unlagged correlation between monthly precipitation and monthly mean groundwater storage anomaly is generally low (less than 0.4). One reason is that, as seen in Fig. 3, groundwater storage anomalies exhibit strong seasonality in all basins while the seasonal cycle of precipitation can be weak or out of phase with that of groundwater storage. For example, in the Upper Mississippi and Missouri basins, annual minimum precipitation occurs in early spring, much earlier than that of groundwater storage anomalies. It is possible that increased water demand in the summer and consequent withdrawals have some bearing on the observed seasonal cycle of groundwater storage, especially in groundwater-dependent Long Island. However, groundwater recharge from VIC, which does not simulate human impacts on the water cycle, also exhibits a strong seasonal cycle that is consistent with that of groundwater storage (Figure 3), suggesting that precipitation type and evapotranspiration govern the seasonal variability of recharge and hence groundwater storage. Supporting that hypothesis is the study of Steenhuis et al. (1985), who estimated recharge in Long Island using two different techniques

and found that a high percentage of precipitation becomes recharge from late fall to early spring, while precipitation during the summer months does not contribute meaningfully to total annual recharge.

Table 2 shows that groundwater storage anomalies are more strongly correlated with monthly groundwater recharge than with monthly precipitation. That is not surprising considering the processes that occur between the incidence of precipitation on the land surface and groundwater recharge, including snowmelt (when applicable), infiltration, and drainage through the soil layers, all of which modulate the timing and quantity of recharge. Those processes are incorporated into the simulation of recharge by the VIC model. Annual maximum recharge often occurs in March to May while the annual minimum occurs in summer to early fall (Fig. 3). That recharge peaks before precipitation in most of these regions may seem counterintuitive. However, two factors contribute to the strong seasonality of groundwater recharge and its springtime peak. One is that, in the northern and high altitude parts of these regions, seasonal snowpack can store and, during spring, release a significant portion of annual precipitation. Snowmelt provides a slow, steady source of water that is ideal for diffuse recharge. Second, by the time precipitation peaks in late spring to early summer, evaporative demand and plant root uptake have also increased, reducing the water available for groundwater recharge.

The seasonal cycle of simulated groundwater recharge led that of observed groundwater storage by only about a month, resulting in their high correlation (3<sup>rd</sup> row of Table 2). Recharge was even more highly correlated with groundwater storage change (calculated as the difference between the current and previous month's groundwater storage anomalies), which is logical because recharge is one of two components directly contributing storage change (equation (1)),

while storage anomalies represent an accumulation of changes. Groundwater storage exhibits larger seasonal fluctuations than recharge. Seasonal variations in groundwater discharge appear to contribute appreciably to the annual cycle of groundwater storage (Fig. 3). It is also possible that VIC underestimates recharge seasonality, which would affect the discharge estimates computed as a residual, or that the impacts of withdrawals for agriculture accentuated the groundwater storage decline during the growing season despite our attempts to eliminate wells directly influenced by pumping. Nevertheless, the phase of discharge makes sense intuitively, as in most regions its maximum occurs sometime between the spring maximum groundwater levels (when seepage areas are enlarged and flow gradients increased, enhancing stream, spring, and submarine groundwater discharge) and summer (when withdrawals for irrigation and by plant roots peak).

Fig. 4 plots time series of monthly mean groundwater storage anomalies, mean precipitation, and mean groundwater recharge with their seasonal cycles removed in order to highlight interannual variability. A backward 6-month moving average was applied to precipitation and recharge to smooth out high frequency variability. It can be seen that the non-seasonal changes in the three variables are often well correlated. Groundwater storage changes in the Red-LM basin often lagged precipitation anomalies by several months (best seen in the late 1990s and mid 2000s), reflecting the deeper aquifers in that region. In Long Island, Upper Mississippi, and Missouri, the long term dynamic range of groundwater storage is much larger than the average seasonal variability (Fig. 3), as the former is stretched by prolonged wet or dry periods. This result supports the finding by Wang (2012) that groundwater storage changes under persistent dry and wet conditions cannot be ignored in annual water budget analyses.

The large inter-annual variability of groundwater storage anomalies is facilitated by the nearly unbounded range of groundwater storage. That is to say, the water table only approaches an upper bound near surface water features or during exceptional floods; and aquifers rarely go dry in non-arid regions due to natural forces alone. In contrast, soil moisture's range is narrowly limited by saturation and wilting point, and thus its inter-annual variability is relatively small.

### 3.2 Spatial Variability

Fig. 2 shows that spatial variability of groundwater storage anomalies, as manifested in the standard deviation, varies in time. In general, it is larger when the mean groundwater storage anomaly is relatively low or high, such as during the severe drought in the Upper Mississippi basin in 1987-89 and the flooding in the same basin in 1993.

Fig. 5 further explores the dependency of spatial variability on the spatial mean groundwater storage anomaly. Spatial variability increases with increasing magnitude of the groundwater storage anomaly (seen as an upward concave shape in Figure 5) in several regions including the four northeastern regions, the Ohio-Tennessee, and the Missouri basins. This is the reverse of the upward convexity of the relationship between soil moisture spatial variability and mean (Owe et al., 1982; Famiglietti et al., 2008; Rosenbaum et al. 2012). As discussed in Li and Rodell (2013), while physical processes contribute to the upward convexity of the standard

deviation versus mean soil moisture curve, it is ultimately controlled by the inflexible lower and upper bounds of a soil's storage capacity (wilting point and saturation), which minimize variability when approached. Unconfined groundwater storage, on the other hand, rarely confronts a hard limit, as discussed earlier. As a result, the upper and lower bounds of storage are unlikely to restrict spatial variability of groundwater anomalies. In fact, spatial variability tends to increase as groundwater storage approaches extremes. This is because the dynamic range of groundwater storage is heterogeneous in space, and thus spatial differences in high and low groundwater anomalies during wet and dry periods are likely to be enhanced, leading to increased variability during these periods.

In the Upper Mississippi, Red-LM, and Mississippi an upward concave relationship is not observed, but neither is an upward convexity. Part of the reason is that groundwater storage anomalies from certain wells may exhibit longer scale temporal variability with extremes misaligned with those at other wells (e.g., Fig. 2, the Red-LM panel). As a result, spatial variability did not always peak at the time of maximum or minimum spatial means. As indicated earlier, the longer term variability is likely associated with the deeper and wider range of well depths and low rates of groundwater recharge in Red-LM (Table 1). This result suggests that the dependence of spatial variability on spatial mean may be obscured if the scale of groundwater temporal variability differs significantly among the wells in a region, which is likely when a wide range of water table depths exists.

### 3.3 Scale dependency of groundwater storage anomalies

Understanding the scale dependency of groundwater storage anomalies may be valuable for groundwater monitoring network design, interpreting remotely sensed observations related to groundwater, interpolating between sparse well observations, and identifying environmental controls on groundwater variability. Past studies have shown that soil moisture spatial variability increases as a power function of extents (e.g., Famiglietti et al., 2008, Li and Rodell, 2013) which can be linearized through log-transformation (Hu et al., 1998):

$$\log(\sigma_{\lambda}) = H \log(\lambda) + C \quad (2)$$

where  $\sigma_{\lambda}$  is the spatial variability at extent  $\lambda$ ,  $H$  is the slope of the linear relationship, indicating the strength of scale dependency, and  $C$  is the intercept at the y-axis.

The scale dependency of groundwater storage anomalies is illustrated in Fig. 6(a), which plots the logarithm of spatial variability of groundwater storage anomalies as a function of log-extent for each of the nine regions. Due to the irregular shapes of the regions, spatial extents were calculated as the square root of the area of each region, which is consistent with the soil moisture scaling approach of Famiglietti et al. (2008). Spatial variability was determined by averaging the spatial standard deviations over all months of data. The log of spatial variability of groundwater storage anomalies increases more or less linearly as the log of extent increases. Deviations from the fitted line may be caused by differences in the dynamic ranges of groundwater storage among the regions, related to either climate or aquifer properties, or by well coverage heterogeneities. To eliminate the influence of dynamic range differences on the scaling relationship, spatial variability was normalized by the temporal standard deviation of groundwater storage anomalies (averaged over all wells in each region), and the results are also plotted against spatial extent in Fig. 6(a). The normalized data points form a nearly perfect straight line on a log-log graph. As shown in Table 3, normalization slightly increased the slope



of the linear relationship and, by reining in the outliers, caused it to be significant at the 5% significant level. An additional piece of evidence that the relationship theorized here is a real phenomenon is that normalized standard deviation in the Mississippi basin (the right-most point in Fig. 6(a)), where the well set comprises only those of the four sub-basins, is larger than that in any of the sub-basins (the next four points from right to left).

The observed scale dependency of groundwater storage variability logically leads to the question of whether the properties and processes influencing groundwater storage changes also exhibit scale dependency. Fig. 6 (b) shows that spatial variability of modeled recharge (also averaged over all months) is strongly and significantly dependent on spatial scale. Variability of specific yield is only slightly dependent on spatial scale with an insignificant increasing slope (Table 3). It is clear that normalization removed much of the impact of specific yield (note the change in correlation between the points in Fig. 6(c) and those in Fig. 6(a) before and after normalization). Based on this limited information one might conclude that the scale-dependency of groundwater storage anomaly spatial variability is influenced more by climate than by aquifer properties. However, other aquifer properties such as transmissivity, whose estimation is beyond the scope of this study, may be important.

The stronger scale-dependency of the spatial variability of recharge (Table 3) compared to that of groundwater storage anomalies is probably related to both the spatial heterogeneity of precipitation and the tendency of groundwater storage to revert to a mean state over time via lateral flows and discharge. It also may be due in part to the larger support (the area represented by each of the values that contribute to the mean, in this case  $0.125^\circ$  grid pixels) of the model compared with that of depth-to-water measurements. Li and Rodell (2013) showed that soil moisture data with larger support exhibit stronger scale dependency than data with smaller

support. Further, Fig. 6(b) shows that normalization increased the scale dependency of recharge, probably because the dynamic ranges of recharge were smaller in the larger regions (see Fig. 3).

To further investigate the environmental control on groundwater spatial variability, Fig. 7 plots the temporal correlation between groundwater storage anomalies of any two wells (located in the same region) as a function of their separation distance. A decreasing trend (as indicated by the solid line which is significant at the 5% level) in correlation is observed with increasing distances. This is consistent with the previously demonstrated increasing spatial variability of groundwater storage anomalies with increasing region size. Many low and even negative correlations are also observed at short separation distances (less than 200 km), which is not unexpected considering that hydrogeological conditions (aquifer formation and depth to groundwater) can change substantially over short distances. For example, it was determined that a few very shallow wells in Massachusetts caused many negative correlations at short separation distances, as groundwater variations in those wells was much more strongly controlled by changing meteorological conditions than those in deeper wells.

The derived log-log relationship between spatial variability of groundwater storage anomalies and regional extent may be useful for determining sampling sizes that are able to characterize variability of groundwater storage anomalies with a desired level of accuracy and enough precision for the particular application (e.g., water resources monitoring; validation of numerical models or remote sensing retrievals). As we know, the number of point samples required to capture the spatial mean depends on the spatial variability of the field (Wang et al., 2008):

$$N = t_{1-\alpha/2, N-1}^2 \frac{\sigma^2}{d^2} \quad (3)$$

where  $N$  is the number of samples required;  $\sigma$  is the standard deviation of the sample data (spatial variability);  $d$  is the desired accuracy (absolute error) between the true (population) mean and the sample mean;  $t_{1-\alpha, N-1}^2$  is the Student's t-distribution at the given significance level,  $\alpha$  (5% was used below), and the degree of freedom (number of samples),  $N$ . Since the  $N$  is unknown initially, an iterative method was used to estimate  $N$  based on given  $\sigma$  and  $d$  (Wang et al., 2008). Equation (3) assumes the data are normally distributed, which is difficult to test given the small number of wells in some regions. However, Fig.2 shows that groundwater storage anomalies are generally close to their mean with only a few outliers, which is a behavior of normal distributions.

Equation (3) demands that more samples are needed if the field is highly heterogeneous or if the desired accuracy is higher. We substituted the slope and y-intercept for the non-normalized groundwater storage spatial variability from Table 3 into equation (2), and then combined equations (2) and (3) to determine the number of wells needed to estimate the mean groundwater anomaly at a given error level:

$$N = t_{1-\alpha/2, N-1}^2 \frac{9.58\lambda^{0.26}}{d^2} \quad (4)$$

Here we define the extent,  $\lambda$ , as the square root of the study area. Equation (4) can also be rearranged to estimate the accuracy of regional mean groundwater anomalies based on the number of wells:

$$d = t_{1-\alpha/2, N-1} \sqrt{\frac{9.58\lambda^{0.26}}{N}} \quad (5)$$

Fig. 8 plots the solutions of equation (4) for various levels of desired accuracy (absolute error). The number of wells required increases as the extent increases but the rate of increase gradually subsides due to the power function. Also plotted in Fig. 8 are the numbers of wells vs. the spatial extent for each of the nine regions in this study, illustrating the estimated accuracy of mean groundwater storage anomalies based on the sets of wells used here. We estimate that the absolute error is the largest (around 5 cm) in the sub-basins of the Mississippi river and smallest (less than 2 cm) in Massachusetts where more wells are available.

Ideally, the parameters estimated for the log-log linear relationship between spatial variability of groundwater storage anomalies and extent would be tested using independent data. Because other data were not available, we constructed an alternative set of regions by splitting, longitudinally, each of the four northeastern regions into two equal sub-regions and merging adjacent Mississippi sub-basins, Upper Mississippi and Ohio-Tennessee, Red-LM and Ohio-Tennessee, Upper Mississippi and Missouri, and Red-LM and Missouri, into four larger basins. In each of these 12 new regions, spatial variability of groundwater storage anomalies was calculated using well data (as the truth) and using equation (2) with parameters from Table 3. Fig. 9 shows that, in general, the derived log-log linear relationship yielded reasonable spatial variability estimates based on extent, especially when spatial variability was normalized as before (right panel). Note that each of the four northeastern regions was split into two equally sized halves, thus the predicted spatial variability was identical for each pair.

#### 4. Summary and Discussion

We used groundwater well observations from the USGS archive to assess spatial and temporal variability of groundwater storage anomalies in nine regions, how they vary with spatial scale, and their relationship with other hydro-meteorological variables. Key conclusions were (1) the spatial variability of groundwater storage anomalies increases linearly with spatial scale when plotted on a log-log graph; (2) spatial variability of groundwater storage anomalies increases under relatively dry and wet conditions in most regions, in contrast to the upward convexity of analogous plots of soil moisture variability vs. mean; and (3) the variability of land surface model simulated recharge is consistent with observed groundwater dynamics. While only the second of these conclusions might be considered unexpected, the significance of this study is that it is the first to evaluate these matters at large scales with real data. Of particular importance, the log-log linear relationship (conclusion 1) has been identified in other hydrological variables but it had never been demonstrated with groundwater data. Further, based on the power function and parameters stemming from our analysis, we derived relationships that can be used to estimate the number of wells required to compute spatial mean groundwater storage anomalies at a specified level of accuracy for a given region size, and for assessing the accuracy of area mean estimates given a certain number of wells, under the assumption of normally distributed data.

Overall, our results are consistent with the expectation that relationships between groundwater storage variability and other processes (precipitation recharge, and discharge) are influenced by climatic and hydrogeological conditions. Groundwater storage anomalies derived from in situ measurements exhibit seasonal and inter-annual variability that is positively correlated with hydro-climatic variability but also influenced by other factors. In particular, due

to the filtering effects of the overlying geology, monthly regional-mean groundwater storage anomalies often lag monthly precipitation, more so in regions with drier climate or deeper wells.

Observed changes in groundwater storage and recharge simulated by the NLDAS/VIC land surface model were well correlated (Table 2). That bodes well for the use of VIC simulated recharge as a substitute for observation-based recharge estimates, which are typically only available from field scale studies or depend on water budget approaches that carry with them a significant degree of uncertainty (Scanlon et al., 2002). Conversely, it bolsters our underlying assumption that data from scattered groundwater wells are sufficient for regional groundwater storage analysis. However, we caution that the suitability of VIC recharge appears to be region-specific and that recharge estimates from other land surface models should be independently verified. Xia et al. (2012a) showed that recharge simulated by the four NLDAS-2 models exhibited the largest disparities and lowest inter-correlation compared with other modeled states and fluxes, especially in the drier U.S. interior, indicating that it is highly sensitive to model physics and parameters.

The large inter-annual variability of groundwater storage anomalies and its relationship with or impacts on other variables underscore the importance of accounting for groundwater storage changes in water budget analyses and of properly representing groundwater storage and its interannual variability in hydrological models. As shown in this study, temporal variability of groundwater storage is largely controlled by precipitation and evapo-transpiration, but its behavior also depends on other meteorological factors and soil and aquifer properties. Simulating groundwater storage in a model is complicated by the three dimensional heterogeneity of the subsurface and a lack of detailed hydrogeological information in most of the world. Further, modeled groundwater is sensitive to model parameters affecting groundwater depths (Li and

Rodell, 2014) and may be more vulnerable to uncertainties associated with model physics and static parameters than other near surface states such as soil moisture, due to groundwater's low replenishment rate (Li et al., 2012). Thus, understanding and properly representing large scale aspects of groundwater storage variability, such as those described here, may be a useful path to enhancing groundwater simulation in large scale numerical models. Such improved understanding also will aid in interpreting the low spatial resolution terrestrial water storage data provided by GRACE. Further, constraining a land surface model, whose parameterization has been informed by conclusions of studies based on in situ measurement such as this, with observations from GRACE via data assimilation, may be the best hope for global accounting of groundwater storage changes (Zaitchik et al., 2008).

Our study showed that spatial variability of groundwater storage anomalies exhibits dependence on spatial mean, with higher spatial variability at low and high mean storage anomalies in most regions. This relationship contrasts with that between soil moisture spatial variability and spatial mean. A possible implication of this difference is that the contribution of groundwater to local evapotranspiration and runoff may be enhanced relative to that of soil moisture under very wet and dry conditions. Similarly, groundwater levels may be a more important consideration for flood prediction than soil moisture. This is further justification for incorporating groundwater into various types of Earth system and water management models. The dependence of spatial variability on mean seems to be more distinct in humid regions with more uniformly shallow water tables. Deeper water tables and/or low rates of recharge tend to produce a wider range of temporal variability among wells that may obscure the relationship between spatial variability of groundwater storage anomalies and spatial mean.

The linear relationship between log-spatial-variability of groundwater storage anomalies and log-spatial-extent provides the theoretical basis for scaling groundwater measurements and a mathematical form to obtain spatial variability of groundwater storage anomalies at any given scale. We applied that relationship to derive an equation for estimating the number of wells required to determine area-mean groundwater storage anomalies with a given level of accuracy, and, to estimate accuracy based on the region size and number of wells available. We expect these equations will be valuable for assessing uncertainty in groundwater resource studies and for designing groundwater monitoring networks. The current scarcity of publicly available groundwater data in much of the world compels scientists and water resources managers to extrapolate data from a few in situ observations to their scale of interest. Knowledge of the spatial behavior of groundwater storage variability will benefit these endeavors. However, because the scale dependency likely changes with support, studies on the scale dependency of anomalies of GRACE derived TWS and model estimated groundwater would help to optimize downscaling and disaggregation (into groundwater and other components) of GRACE TWS via data assimilation (Zaitchik et al., 2008).

The scale dependency of groundwater storage anomalies reflects the combined influences of various controls on groundwater at different spatial scales, including meteorological inputs, topography, geology, and vegetation. It is this combination of influences, and not necessarily the extent itself, that induces the observed log-log linear relationship between spatial variability of groundwater storage anomalies and extent. Note that, since wells that are directly influenced by human activities were excluded from this study, the reported log-log linear relationship may not be appropriate for estimating spatial variability of groundwater storage anomalies in regions where groundwater withdrawals exert substantial control over regional mean water levels.



Uncertainty in our results and conclusions mainly arises from reliance on data from a finite number of wells to characterize spatial variability in our nine study regions, which was beyond our control. There are many ways in which the study could have been improved given a greater density of high quality groundwater observation time series, such as investigating smaller scale variability and understanding areas of influence of individual wells, but we did our best with what was available. Further, the degree to which these sets of wells represent heterogeneity of groundwater storage variations, which results from both geology and meteorology, is unknown. If the wells were preferentially installed in certain types of aquifer, our results could be biased. The normalization method used in developing Fig. 6 was one attempt to remove the influence of climate and geology on what we inferred about scaling of groundwater variability. We imposed strict criteria for well selection from among those archived by the USGS, specifically to eliminate wells that could be directly impacted by pumping or injections, wells installed in confined aquifers, and wells that were not monitored frequently enough to capture the seasonal cycle. Nevertheless, conformity with these criteria was judged based on incomplete metadata and our own assessments of the time series, and it is very possible that flawed candidate wells slipped through. We relied on the same incomplete metadata and literature review to estimate specific yield values, and the results are surely imperfect. Finally, we did not attempt to account for atmospheric pressure effects on measured groundwater levels. We believe that accounting for these effects would be inconsequential to our major conclusions, which were based on either multi-annual temporal means (over which periods atmospheric pressure effects average out) or a preponderance of data.

### Figure Captions

Fig. 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 basin. The area of each region given in Table 1 corresponds to the parts of the map that are shaded or encompassed by rectangles.

Fig. 2. Time series of monthly groundwater storage anomalies at individual wells (grey lines), the regional mean (black lines), and spatial variability (standard deviation, marked lines) for each of the nine study regions. Monthly NLDAS precipitation (top black bars) is also plotted.

Fig. 3. Monthly seasonal cycles of regional mean groundwater storage anomalies, recharge anomalies, discharge anomalies, and changes in groundwater storage, and precipitation (grey bars), for each of the nine study regions. Anomalies, computed by removing the long term means, are plotted in order to facilitate comparison among variables.

Fig. 4. Time series of deseasonalized groundwater storage anomalies, deseasonalized precipitation and deseasonalized recharge anomalies in the nine regions. Precipitation and recharge are plotted as the running average of the prior 6-months in order to accentuate significant wet and dry periods.

Fig. 5. Spatial variability (standard deviation) of groundwater storage anomalies as a function of mean anomaly in each region.

Fig. 6. Spatial variability (standard deviation) of groundwater storage anomalies (a), groundwater recharge (b), and specific yield (c) as a function of spatial extent. Log scales were used in both x and y axes. Spatial standard deviations of groundwater storage anomalies and recharge were averaged over all months. Spatial extent was calculated as the square root of the area in each region. Normalized standard deviation in a region was computed by dividing the spatial standard deviation by the average (over all wells) of the long term temporal standard deviation.

Fig. 7. Coefficients of temporal correlation between groundwater storage anomalies of any two wells (located in the same region) as a function of their separation distance.

Fig. 8. Number of wells required to represent the spatial mean at five different absolute error levels (“d”) as a function of spatial extent. The x and y values of each circle (with region id enclosed) correspond to the extent and number of wells at each of the nine regions, respectively.

Fig.9. Predicted spatial variability (StD) of groundwater storage anomalies (left panel) based on equation (2) versus estimated spatial variability using in situ groundwater data in equally split regions of Long Island, New Jersey, Massachusetts, Pennsylvania and combined regions of Upper Mississippi and Ohio-Tennessee, Red-LM and Ohio-Tennessee, Upper Mississippi and Missouri, and Red-LM and Missouri. The right panel shows the similar results for normalized spatial variability.

**Acknowledgements**

This research was funded by NASA's Terrestrial Hydrology Program. We thank the USGS and the Illinois State Water Survey for providing the groundwater data, David Mocko for providing monthly NLDAS-2 VIC model output and anonymous reviewers for their comments that have helped improve the quality of this paper.

## References

- Alley, W. M., R.W. Healy, J.W. LaBaugh and T.E. Reilly, 2002. Flow and storage in groundwater systems. *Science*, 296(5575), 1985-1990.
- Barco, J., T. S. Hogue, M. Girotto, D. R. Kendall and M. Putti, 2010. Climate signal propagation in southern California aquifers, *Water Resour. Res.*, 46, W00F05, doi:10.1029/2009WR008376.
- Boening, C., J. K. Willis, F. W. Landerer, R.S. Nerem and J. Fasullo, 2012. The 2011 La Niña: So strong, the oceans fell. *Geophysical Research Letters*, 39(19).
- Brocca, L., T. Tullo, F. Melone, T. Moramarco and R. Morbidelli, 2012. Catchment scale soil moisture spatial-temporal variability, *J. Hydrology.*, 422–423 (2012) 63–75.
- Cosgrove, B., D. Lohmann, K. E. Mitchell, P. R. Houser, E. F. Wood, J. C. Schaake, A. Robock, C.R. Marshall, J. Sheffield, Q. Duan, L. Luo, R. W. Higgins, R. T. Pinker, J. D. Tarpley and J. Meng, 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.
- Crosbie, R. S., B. R. Scanlon, F. S. Mpelasoka, R. C. Reedy, J. B. Gates and L. Zhang, 2013. Potential climate change effects on groundwater recharge in the High Plains Aquifer, USA, *Water Resour. Res.*, doi: 10.1002/wrcr.20292.
- Eltahir, E.A.B. and P.J.-F. Yeh, 1999. On the asymmetric response of aquifer water level to floods and droughts in Illinois. *Water Resour. Res.* **35** (4), 1199–1217.

Famiglietti, J. S., D. Ryu, A. A. Berg, M. Rodell and T. J. Jackson, 2008. Field observations of soil moisture variability across scales, *Water Resour. Res.*, **44**, W01423, doi:10.1029/2006WR005804.

Famiglietti, J. S., M. Lo, S. L. Ho, J. Bethune, K. J. Anderson, T. H. Syed, S. C. Swenson, C. R. deLineage and M. Rodell, 2011. Satellites measure recent rates of groundwater depletion in California's Central Valley, *Geophys. Res. Lett.*, **38**, L03403, doi:10.1029/2010GL046442.

Famiglietti, J. S. and M. Rodell, 2013. Water in the Balance, *Science*, **340**, 1300-1301  
[10.1126/science.1236460](https://doi.org/10.1126/science.1236460)

Feng, W., M. Zhong, J.-M. Lemoine, R. Biancale, H.-T. Hsu and J. Xia, 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, doi:10.1002/wrcr.20192.

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

Hu, Z., Y. Chen and S. Islam, 1998. Multiscaling properties of soil moisture images and decomposition of large- and small-scale features using wavelet transforms, *Int. J. Remote Sens.*, **19**, 2451-2467.

Istanbulluoglu, E., T. Wang, O. M. Wright and J. D. Lenters, 2012. Interpretation of hydrologic trends from a water balance perspective: The role of groundwater storage in the Budyko hypothesis, *Water Resour. Res.*, **48**, W00H16, doi:10.1029/2010WR010100.

Konikow, L. F., 2011. Contribution of global groundwater depletion since 1900 to sea-level rise. *Geophysical Research Letters*, 38(17).

Li, B., M. Rodell, B. F. Zaitchik, R. H. Reichle, R. D. Koster and T. M. van Dam, 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. Hydrology*, 446–447 (2012) 103–115, <http://dx.doi.org/10.1016/j.jhydrol.2012.04.035>.

Li, B. and M. Rodell, 2013. Spatial variability and its scale dependency of observed and modeled soil moisture over different climate regions, *Hydrol. Earth Syst. Sci.*, 17, 1177-1188, 2013, [www.hydrol-earth-syst-sci.net/17/1177/2013/](http://www.hydrol-earth-syst-sci.net/17/1177/2013/), doi:10.5194/hess-17-1177-2013.

Li, B. and M. Rodell, 2014. Evaluation of a model-based groundwater drought indicator, *J. Hydrology*, <http://dx.doi.org/10.1016/j.jhydrol.2014.09.027>.

Liang, X., D. P. Lettenmaier, E. F. Wood and S. J. Burges, 1994. A simple hydrologically based model of land surface water and energy fluxes for GCMs, *J. Geophys. Res.*, 99, 14,415–14,428, doi:10.1029/94JD00483.

Owe, J., E. B. Jones and T. J. Schmugge, 1982. Soil moisture variation patterns observed in Hand County, South Dakota, *Water Resour. Bull.*, 18, 949– 954.

Perez-Valdivia, C., D. Sauchyn and J. Vanstone, 2012. Groundwater levels and teleconnection patterns in the Canadian Prairies, *Water Resour. Res.*, 48, W07516, oi:10.1029/2011WR010930.

Rodell, M. and J. S. Famiglietti, 2001. An analysis of terrestrial water storage variations in Illinois with implications for the Gravity Recovery and Climate Experiment (GRACE), *Water Resour. Res.*, **37**(5), 1327-1339.

Rodell, M., J. Chen, H. Kato, J. S. Famiglietti, J. Nigro and C. R. Wilson, 2007. Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE, *Hydrogeology Journal*, 15:159-166, doi:10.1007/s10040-006-0103-7.

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



Rosenbaum, U., H. Boga, M. Herbst, J.A. Huisman, T.J. Peterson, A. Weuthen, and H. Vereecken, 2012. Seasonal and event dynamics of spatial soil moisture patterns at the small catchment scale. *Water Resour. Res.*, doi: 10.1029/2011WR011518.

Rowlands, D. D., S. B. Luthcke, S. M. Klosko, F. G. R. Lemoine, D. S. Chinn, J. J. McCarthy, C. M. Cox and O. B. Anderson, 2005. Resolving mass flux at high spatial and temporal resolution using GRACE intersatellite measurements. *Geophysical Research Letters*, 32, L04310, doi:10.1029/2004GL021908.

Scanlon, B. R., R. W. Healy and P. G. Cook, 2002. Choosing appropriate techniques for quantifying groundwater recharge. *Hydrogeology Journal*, 10(1), 18-39.

Siebert, S., J. Burke, J. M. Faures, K. Frenken, J. Hoogeveen, P. Döll and F. T. Portmann, 2010. Groundwater use for irrigation—a global inventory, *Hydrol. Earth Syst. Sci.*, 14, 1863–1880, 2010, [www.hydrol-earth-syst-sci.net/14/1863/2010/](http://www.hydrol-earth-syst-sci.net/14/1863/2010/)doi:10.5194/hess-14-1863-2010.

Steenhuis, T.S., C.D. Jackson, S.K.J Kung and W. Brutsaert, 1985. Measurement of groundwater recharge on eastern Long Island, New York, USA. *Journal of Hydrology* 79, no. 1: 145-169.

Stromberg, J. C., R. Tiller and B. Richter, 1996. Effects of groundwater decline on riparian vegetation of semiarid regions: the San Pedro, Arizona, *Ecological Applications*, 6(1), 113-131.

Swenson, S. C. and J. Wahr, 2006. Post-processing removal of correlated errors in GRACE data, *Geophys. Res. Lett.*, 33, L08402, doi:10.1029/2005GL025285.

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, 2012. Groundwater and climate change, *Nature Climate Change*, doi:10.1038/nclimate1744.

Voss, K. A., J. S. Famiglietti, M. Lo, C. de Linage, M. Rodell and S. C. Swenson, 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**, doi:10.1002/wrcr.20078.

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

Wang, C., Q. Zuo and R. Zhang, 2008. Estimating the necessary sampling size of surface soil moisture at different scales using a random combination method. *J. Hydrol.* **352** (3–4), 309–321. doi:10.1016/j.jhydrol.2008.01.011.

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.

Weider, K., and D. F. Boutt, 2010. Heterogeneous water table response to climate revealed by 60 years of ground water data, *Geophys. Res. Lett.*, 37, L24405, doi:10.1029/2010GL045561.

Western, A. W. and G. Blösch, 1999. On the spatial scaling of soil moisture, *J. Hydrol.*, **217**, 203–224.

Xia, Y., K. Mitchell, M. Ek, J. Sheffield, B. Cosgrove, E. Wood, L. Luo, C. Alonge, H. Wei, J. Meng, B. Livneh, D. Lettenmaier, V. Koren, Q. Duan, K. Mo, Y. Fan and David Mocko, 2012a. 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.

Xia, Y., K. Mitchell, M. Ek, B. Cosgrove, J. Sheffield, L. Luo, C. Alonge, H. Wei, J. Meng, B. Livneh, Q. Duan and D. Lohmann, 2012b. Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2.

Validation of model-simulated streamflow, *J. of Geophys. Res.*, 117, no. D03110,  
doi:10.1029/2011JD016051.

Yeh, P. J.-F., S. C. Swenson, J. S. Famiglietti and M. Rodell, 2006. Remote sensing of groundwater storage changes in Illinois using the Gravity Recovery and Climate Experiment (GRACE), *Water Resour. Res.*, 42, W12203, doi:10.1029/2006WR005374.

Zaitchik, B.F., M. Rodell and R. 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.

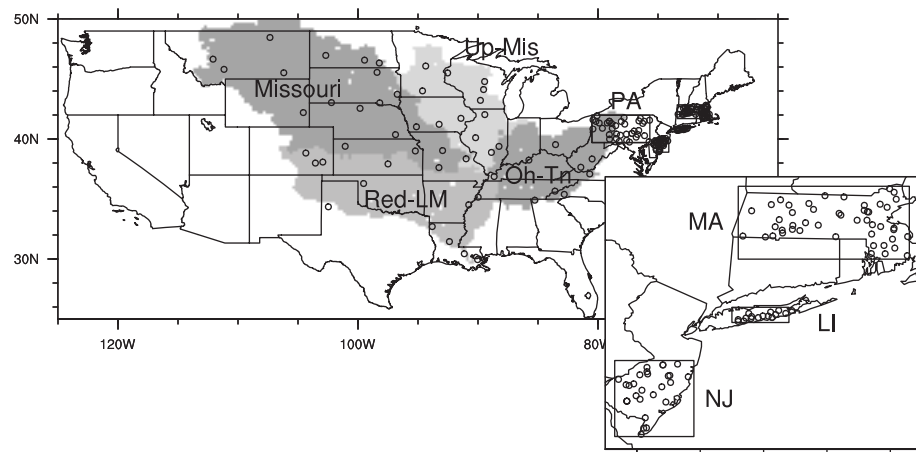


Figure 1. Locations of groundwater wells in Long Island ("LI"), New Jersey ("NJ"), Massachusetts ("MA"), Pennsylvania ("PA") and the four sub-basins of the Mississippi river: the Upper Mississippi ("Up-Mis"), the Ohio-Tennessee ("Oh-Tn"), the combined Red River and Lower Mississippi ("Red-LM") and the Missouri basin. The area of each region given in Table 1 corresponds to the parts of the map that are shaded or encompassed by rectangles.

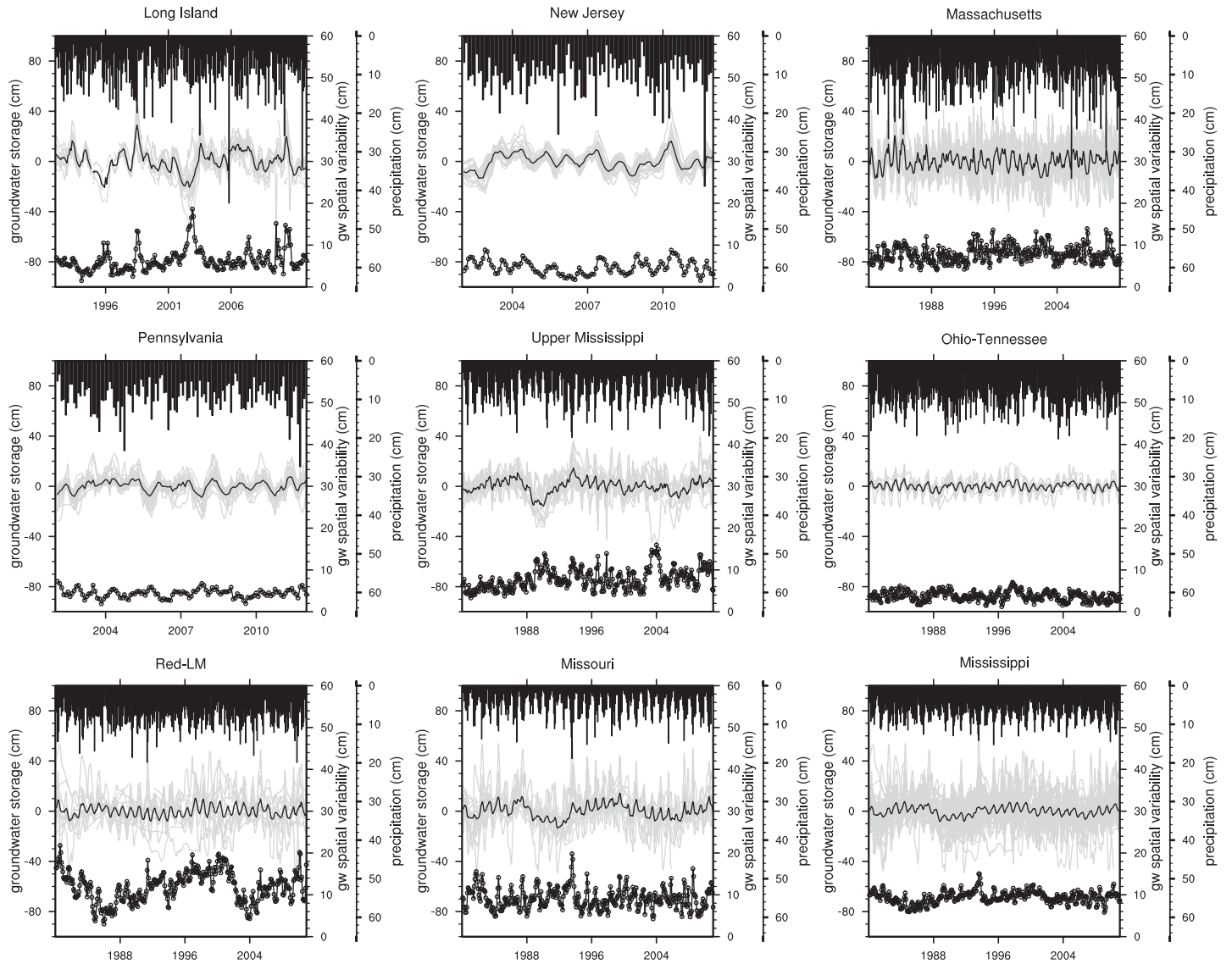


Figure 2. Time series of monthly groundwater storage anomalies at individual wells (grey lines), the regional mean (black lines), and spatial variability (standard deviation, marked lines) for each of the nine study regions. Monthly NLDAS precipitation (top black bars) is also plotted.

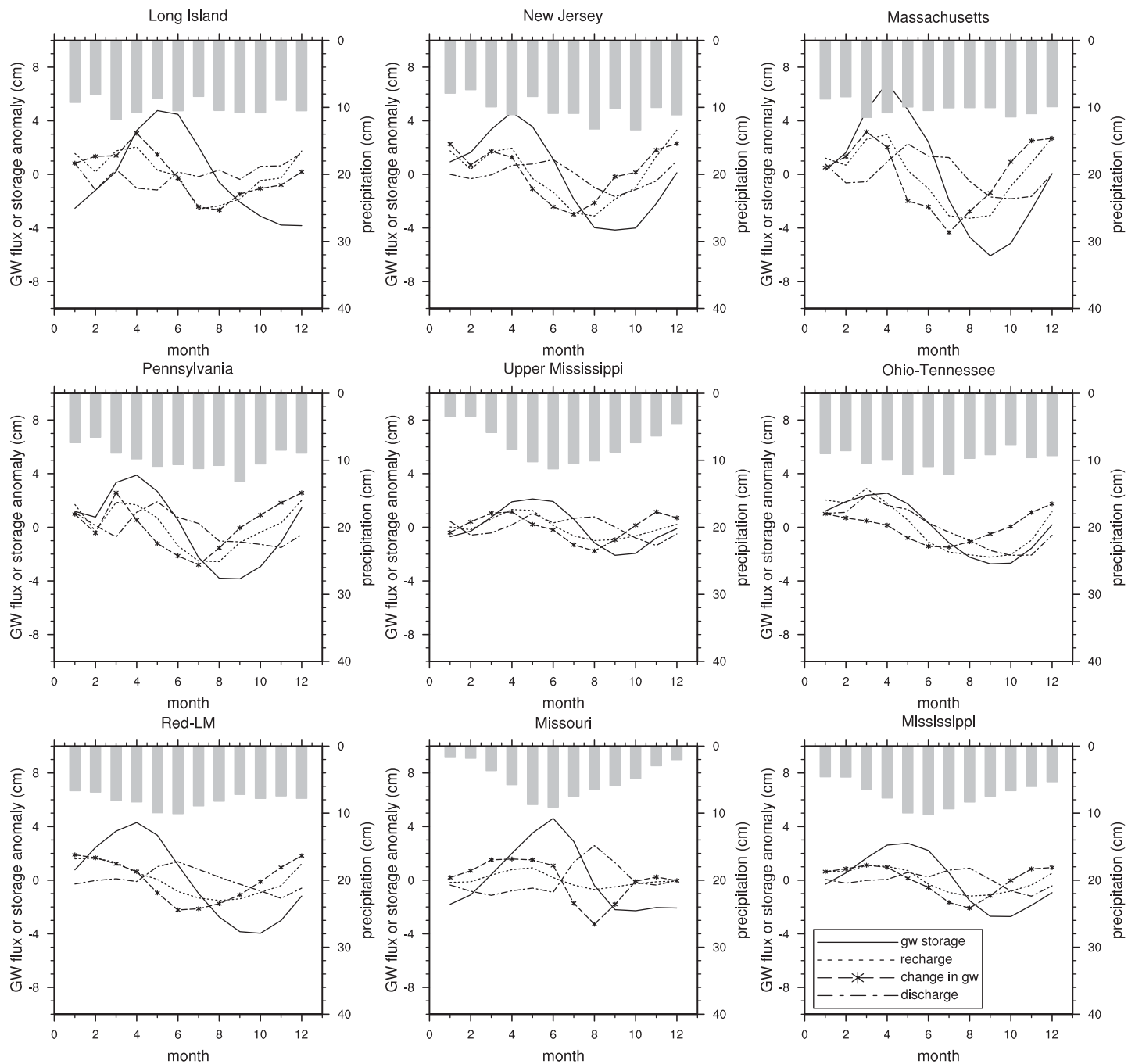


Figure 3. Monthly seasonal cycles of regional mean groundwater storage anomalies, recharge anomalies, changes in groundwater storage, and precipitation (grey bars), for each of the nine study regions. Anomalies, computed by removing the long term means, are plotted in order to facilitate comparison among variables.

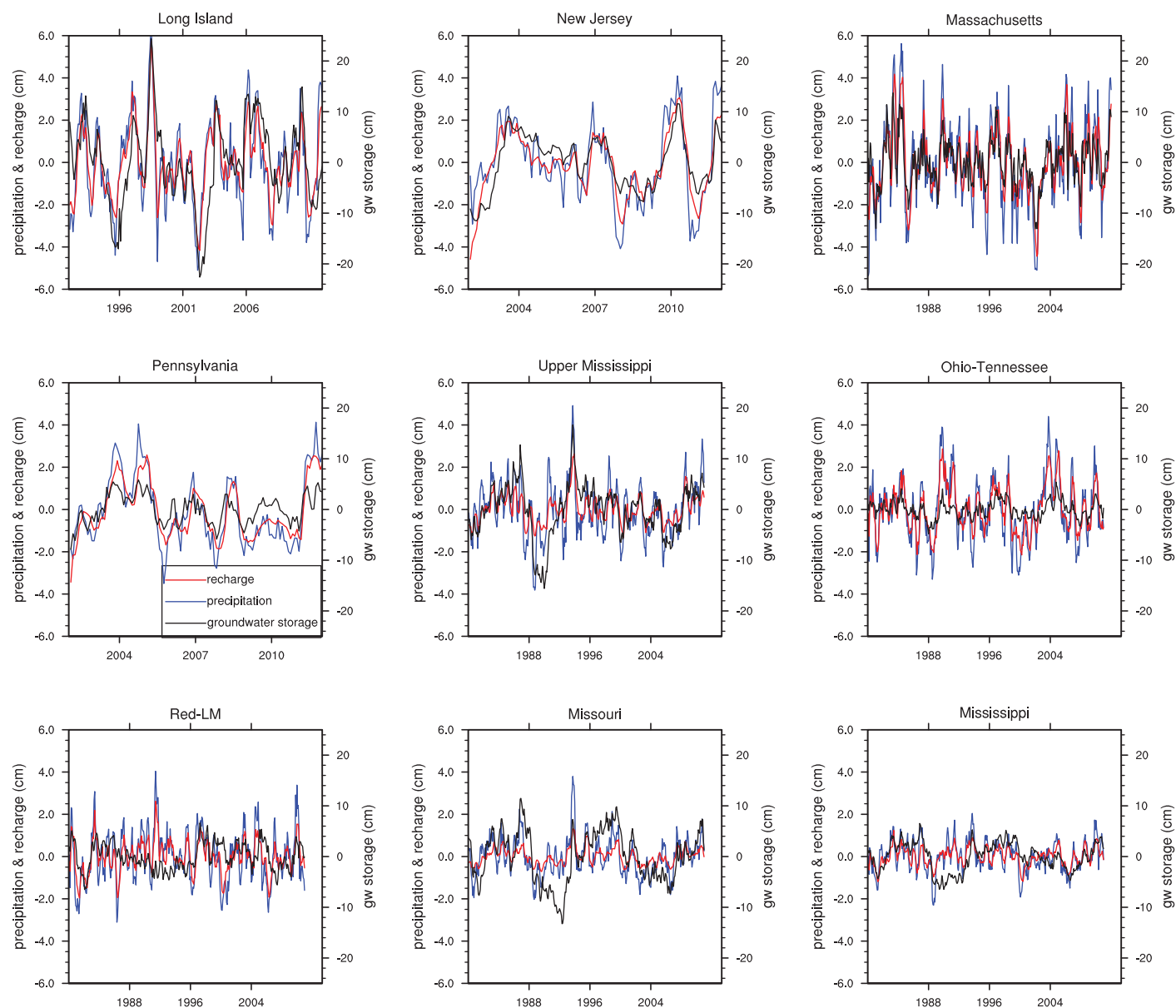


Figure 4. Time series of deseasonalized groundwater storage, deseasonalized precipitation and deseasonalized recharge anomalies in the nine regions. Precipitation and recharge are plotted as the running average of the prior 6-months in order to accentuate significant wet and dry periods.



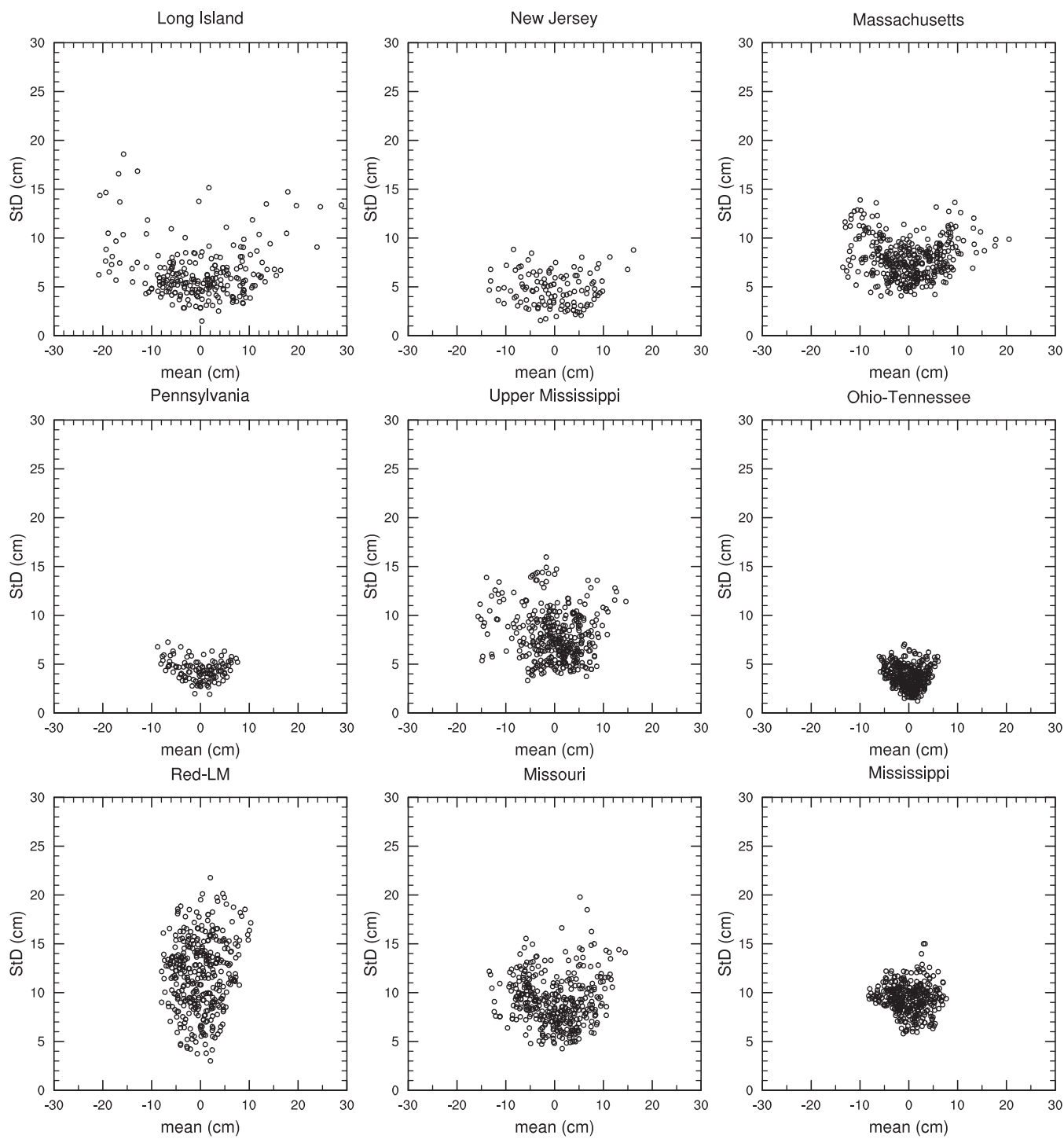


Figure 5. Spatial variability (standard deviation) of groundwater storage anomalies as a function of mean anomaly in each region.

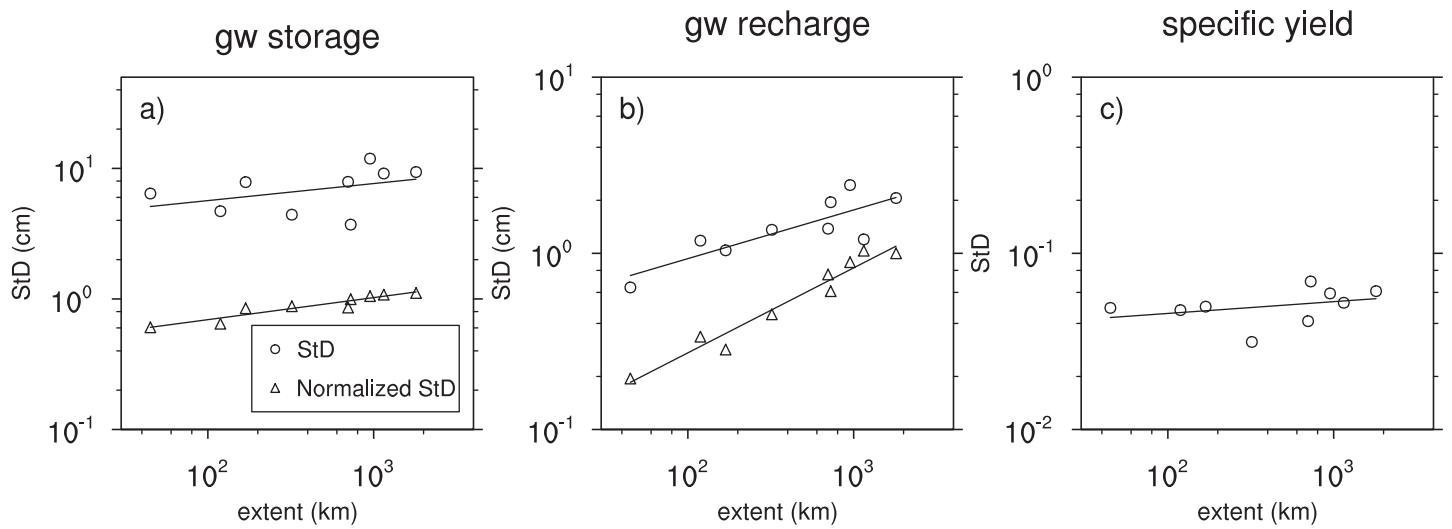


Figure 6. Spatial variability (standard deviation) of groundwater storage anomalies (a), groundwater recharge (b), and specific yield (c) as a function of spatial extent. Log scales were used in both x and y axes. Spatial standard deviations of groundwater storage and recharge were averaged over all months. Spatial extent was calculated as the square root of the area in each region. Normalized standard deviation in a region was computed by dividing the spatial standard deviation by the average (over all wells) of the long term temporal standard deviation.

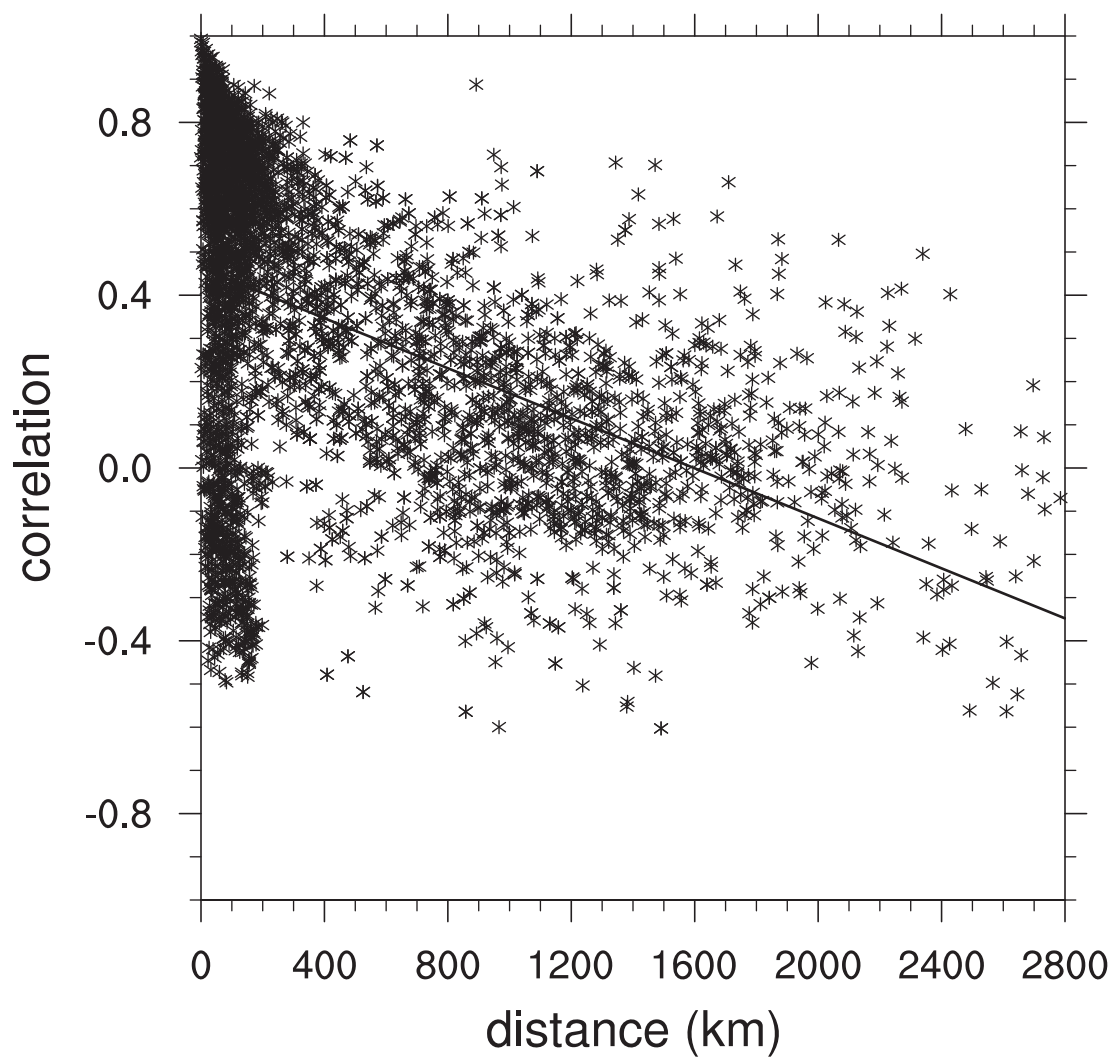


Figure 7. Coefficients of temporal correlation between groundwater storage anomalies of any two wells (located in the same region) as a function of their separation distance.

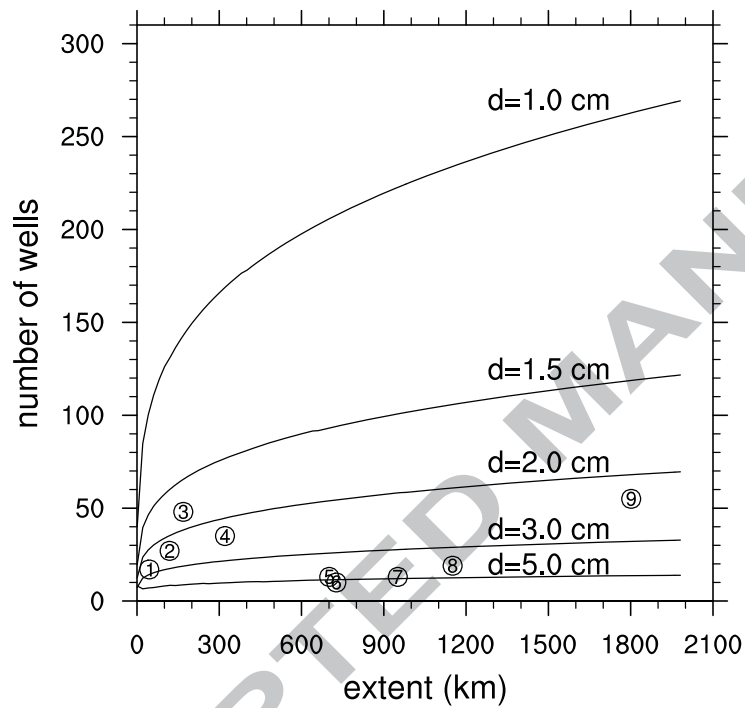


Figure 8. Number of wells required to represent the spatial mean at five different absolute error levels (" $d$ ") as a function of spatial extent. The x and y values of each circle (with region id enclosed) correspond to the extent and number of wells at each of the nine regions, respectively.

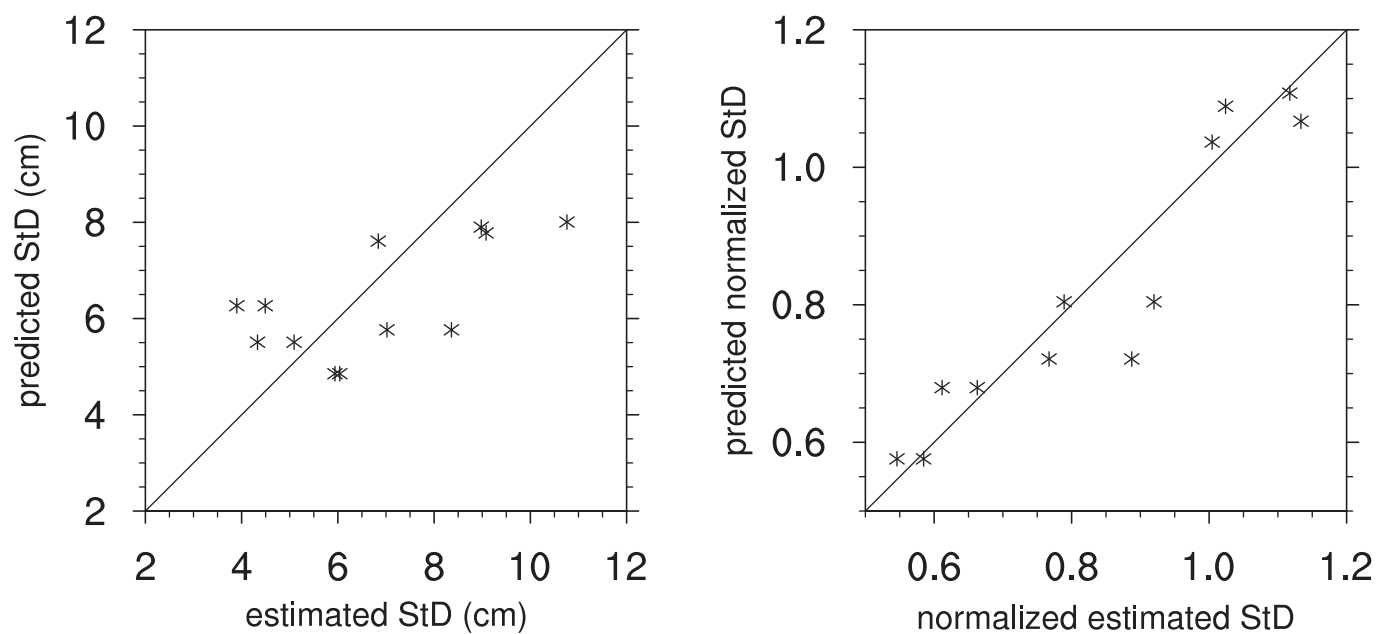


Fig.9. Predicted spatial variability (StD) of groundwater storage anomalies (left panel) based on equation (2) versus estimated spatial variability using in situ groundwater data in equally split regions of Long Island, New Jersey, Massachusetts, Pennsylvania and combined regions of Upper Mississippi and Ohio-Tennessee, Red-LM and Ohio-Tennessee, Upper Mississippi and Missouri, and Red-LM and Missouri. The right panel shows the similar results for normalized spatial variability.

Table 1. Region name, data period, number of wells, area, average specific yield ( $\bar{S}_y$ ), average well depth ( $\bar{d}_{well}$ ), average depth to water ( $\bar{d}_{gw}$ ), average annual NLDAS precipitation ( $\bar{P}$ ) and average annual groundwater recharge ( $\bar{R}$ ) in each region.

id	region	data period	# of wells	area (km <sup>2</sup> )	$\bar{S}_y$	$\bar{d}_{well}$ (m)	$\bar{d}_{gw}$ (m)	$\bar{P}$ (cm)	$\bar{R}$ (cm)
1	Long Island	1992-2011	16	2000	0.26	15	8	119	71
2	New Jersey	2002-2011	27	14200	0.17	27	6	124	62
3	Massachusetts	1980-2011	48	28400	0.20	9	4	122	68
4	Pennsylvania	2002-2011	35	102900	0.07	42	10	117	57
5	Upper Mississippi	1980-2010	13	491800	0.17	19	6	90	25
6	Ohio-Tennessee	1980-2010	10	528100	0.09	38	7	119	49
7	Red-LM	1980-2010	13	903900	0.16	86	17	97	30
8	Missouri	1980-2010	19	1324000	0.14	30	9	60	8
9	Mississippi	1980-2010	55	3247800	0.14	42	10	87	25

Table 2. Coefficients of correlation between regional mean groundwater storage anomalies and regional mean precipitation (2<sup>nd</sup> row) and regional mean recharge (3<sup>rd</sup> row), and correlation between regional mean changes in groundwater storage ( $\Delta GW$ ) and recharge (4<sup>th</sup> row).

region	Long Island	NJ	MA	PA	Upper Mississippi	Ohio-Tennessee	Red-LM	MO	MS
gw vs precip	0.06	0.05	0.31	0.17	0.24	0.28	0.07	0.37	0.25
gw vs recharge	0.22	0.53	0.73	0.80	0.55	0.75	0.38	0.43	0.43
$\Delta GW$ vs recharge	0.77	0.85	0.75	0.64	0.53	0.54	0.67	0.65	0.72

Table 3. Parameters obtained by fitting equation (2) to spatial variability (StD) of groundwater storage anomalies, recharge, and specific yields as a function of extents. Slopes in bold are significant at the 5% level.

	groundwater storage		recharge		specific yield
	StD	normalized StD	StD	normalized Std	StD
slope	0.13	<b>0.17</b>	<b>0.28</b>	<b>0.48</b>	0.067
y-intercept	1.13	-1.14	-1.34	-3.52	-3.39

- spatial variability of groundwater storage anomalies increases as a power function of extent
- spatial variability of groundwater storage anomalies depends on mean groundwater storage
- Seasonality of groundwater storage is similar to that of modeled recharge