Referee #1

In their manuscript "Global terrestrial water storage connectivity revealed using complex climate network analyses", the authors investigate correlation networks which reflect mutual dependencies between the temporal variation of terrestrial water storage at different locations. Estimates of terrestrial water storage (TWS) are based on two data products, the GRACE satellite mission and the NOAH model (GLDAS-NOAH). The network topology indicates different river basins to show different connectivity patterns, pointing to teleconnection structures which seem to be present for some basins (Ganges, Mississippi, Tigris) but not for others (Amazon, Congo, Yangtze). Maps of average node connection length (in Euclidean sense) reveal regions with long connections (such as the Pacific Northwest) and those with short connections (such as the Middle East). Nodes of large area-weighted connectivity (coined "supernode regions" by the authors) seem to reflect combined effects of climate variations and anthropogenic activities. The manuscript provides a network perspective on global terrestrial water storage which is a relevant and timely topic for readers of Nonlinear Processes in Geophysics. Combining insights gained from the network analysis might possibly help to identify TWS predictors and to improve land surface models. The employed methods - network measures as well as network inference by thresholding a matrix of Pearson correlation coefficients - are not new and have been frequently applied in many different scientific disciplines. Applying these methods to GRACE and GLDAS data products establishes the novelty of this work. I have only minor remarks and recommend publication after revision.

<u>Reply</u>: Thank you.

Minor remarks:
1) There is a certain ambiguity when defining network nodes which is natural for such datasets. Nevertheless, how does the spatial resolution of the used data products influence the results of the network analyses? Are the results reported in this study "stable" when network nodes are defined in a different way (e.g. after coarse graining)? How does network structures change when the number of nodes is increased or decreased? I expect the maximum number of reasonably defined nodes of a network to be constrained by the numbers of degrees of freedom captured in the data (which, in case of GRACE, seems to be related to the finite resolution of the measuring instruments).

<u>Reply</u>: These are excellent questions. Before addressing these questions, we first present a background on GRACE data processing <u>done by distribution centers</u> (this information is also available under Section 3 of the paper). The standard GRACE product are sets of spherical harmonic coefficients describing monthly variations of the earth's gravity field. Two filters are typically used to post-process the raw GRACE harmonic coefficients, a de-striping filter that removes correlations between certain spherical harmonic coefficients, and Gaussian averaging filter with a half-width of 300 km that reduces random errors in higher degree spherical harmonic coefficients not removed by ''de-striping'' process; finally, the gridded GRACE product used by this paper is obtained by converting spherical coordinates into geographic coordinates to create mass variations at 1-degree globally [Landerer and Swenson, 2012]. A direct consequence of such a smoothing process is that the pixel-wise correlation and the degree of centrality are increased. To ensure a fair comparison, a standard practice in the GRACE community is to apply the same filtering steps to the model data and, in our case, this is the GLDAS-NOAH data. Therefore, any discrepancy between the networks derived from the two products can be attributed

largely to model assumptions and deficiencies. Given this background, our answers to the particular questions are listed below.

How does the spatial resolution of the used data products influence the results of the network analyses?

In general, spatial resolution of a remotely sense product affect the level of details in the calculated measures. The spatial resolution of data products does not affect our comparison results because the two datasets have gone through the same spatial filtering process and considered to have the same "resolution". The spatial resolution is related to the inherent limitation of the GRACE instrumentation, not to any of the network techniques we adopted in this paper.

Are the results reported in this study "stable" when network nodes are defined in a different way (e.g. after coarse graining)?

Yes. The results are stable because the smoothing process essentially reduces the variations/noise and fixes the degree of freedom. We didn't do coarse graining or upscaling in this work; all the pre-processing has been done by NASA/JPL and other distribution centers. Each network node corresponds to a pixel in the dataset. The edge-density-based correlation cutoff threshold is the only parameter used for changing the complexity of networks. Figure 1 of the paper shows that the two products show very similar edge density functions. Because the edge density function is rather smooth, a slight change in cutoff threshold won't affect network. Also, the 300-km filtering radius only introduce local smoothing. The global hotspots are preserved.

How does network structures change when the number of nodes is increased or decreased?

The number of nodes is fixed in this study, which is equal to the number of non-Null, 1-degree pixels given in each dataset.

2) Figure 6 shows GRACE area-weighted connectivity derived from the maximum of the cross correlation functions. Does a corresponding map of the identified lags of the maximum cross correlation reveal any interesting information?

Reply: Accept. We will provide two additional plots in Appendix B of the paper. Figure B1 shows the maximum correlations for six basins chosen in Figure 2 and Figure B2 shows the corresponding phase lags. Recall these plots show the correlation between basin centroid and all other cells in the TWSA dataset. The phase lag plot (normalized by 18) shows that each river basin is in phase with itself and the immediate surrounding regions, but there are significant phase shifts between each river basin and other river basins.
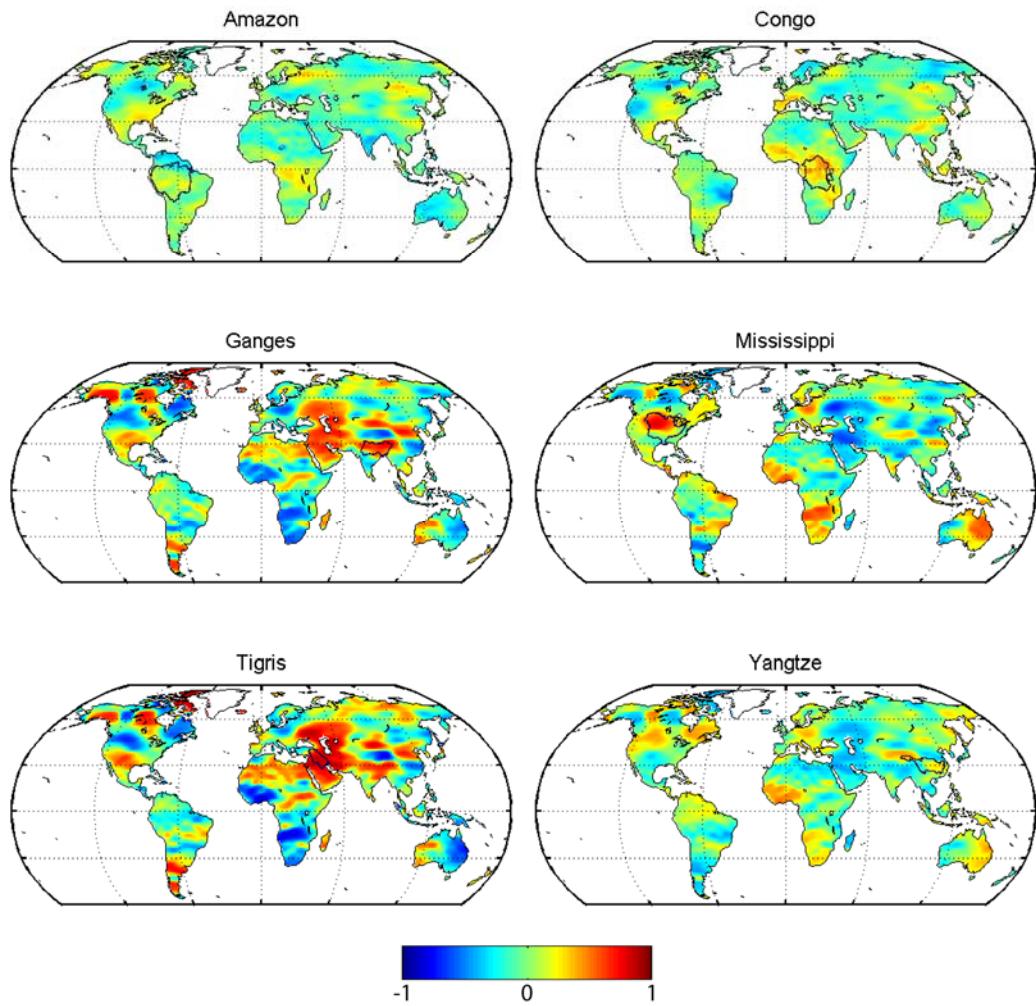
Figure B1. Degree centrality inferred from GRACE TWSA for six river basins, based on the maximum correlation between each basin centroid and all other cells in the grid, and within a window of [-18,18] months.
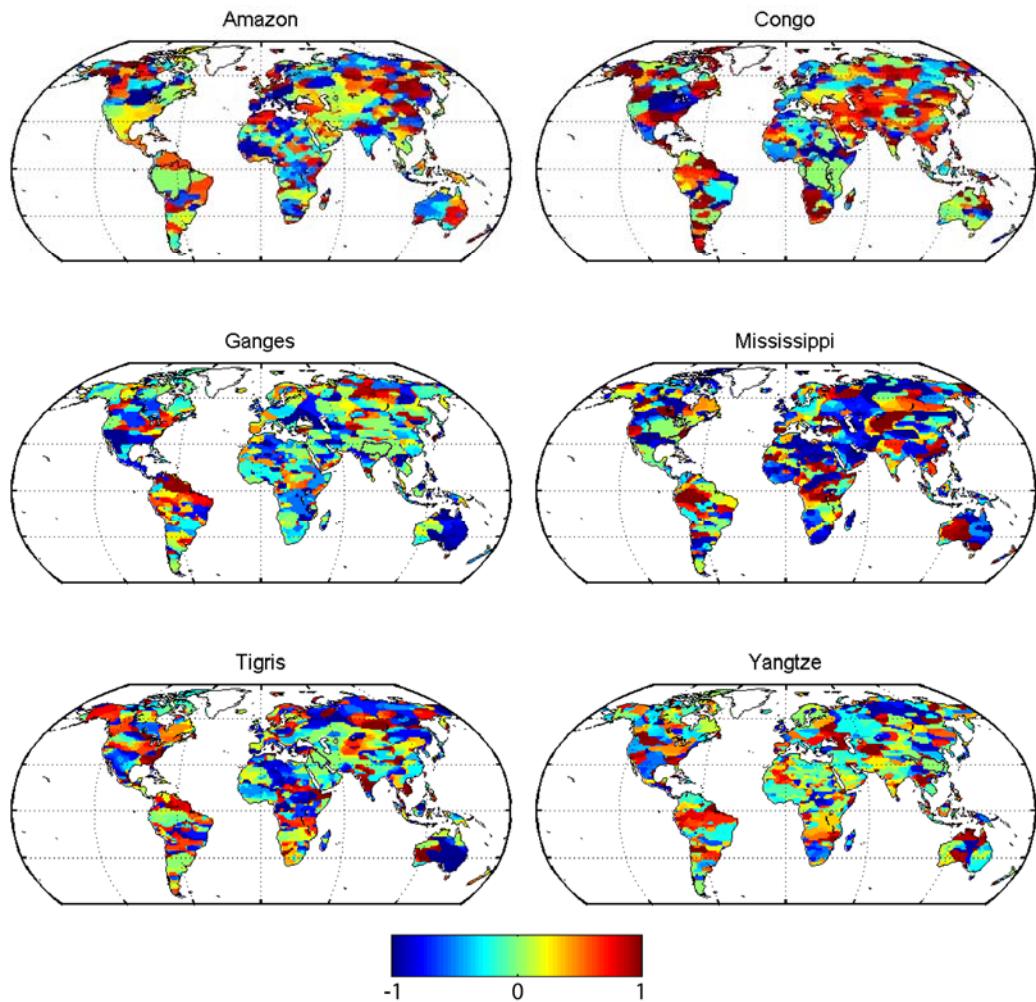
Figure B2. Phase lag of maximum correlations obtained for the six river basins shown in Figure B1 (normalized by 18).

3) p. 788, l. 16: "... \bar{L} provides a measure of network integration." I believe that readers not aware of the spatial aspect of this network measure will profit from a short interpretation of \bar{L}. For the grid-like arrangements of nodes considered in this work, low values of \bar{L} would indicate a grid of nodes which are only locally connected. In a topological sense, this would be a network which is not well integrated since it comes along with a large average shortest path length (which may be somewhat counterintuitive for some network scientists not investigating spatial networks).

Reply: In the context of our work, TWS is connected to the global precip patterns, although TWS is not a flux and does not move as atmospheric fluxes do. Therefore, it is not surprising that most connections are local. Nonetheless, as we mention in the conclusion of the paper, in some regions the connection length distribution is bimodal because of long-range teleconnections (in precipitation). We commented in the conclusion that "*…In terms of connection lengths, the Middle East region is dominated by local connections, whereas regions such as Pacific Northwest, North Central, Colorado River, and North East regions of the US, south Africa, and eastern Australia all have strong bimodal connections*"

4) p. 796, l. 5: "... which are extended for use with gridded datasets." The authors need to explain how exactly they extended the "classic degree of centrality and connection length measures" (l. 4). Is here anything novel that is not yet reported in studies on spatial networks?

Reply: Gridded datasets are special class of spatial networks. The word "classic" is used to distinguish networks that do not require area-weighting. To avoid future confusion, we will rephrase the sentence to "*The constructed networks are further analyzed using the degree of centrality and connection length measures.*"

5) p. 796, l. 19: "... and type of TWS connectivity ...". What do authors mean with "type" here?

Reply: The connection length distribution is bimodal. Loosely speaking, there are two types of connections, the local connection and long-range connection. We will add the following in parentheses after the word type, "*…(i.e., local connection vs. teleconnection)…*"

6) The authors use the notion "complex climate network theory" (CCN theory) at several places in their manuscript. Given that the employed network methods are pretty standard, it is unclear to me what exactly establishs a new theory. I recommend to refrain from using the notion "CCN theory". Instead, authors could use phrases like "applications of complex network theory to climate science" as they already did in their previous publications.

Reply: Accept. We will replace all appearances of "CCN theory" with either "CCN studies" or "CCN applications" in the text.

7) p. 787, l. 10: "... between edge (i,j) ...". Perhaps the authors wanted to write: "... between time series i and j ...".

Reply: Accept. We will replace "…*between edge (i,j)…*" to "…*between time series available at nodes i and j …*"

Referee #2

The paper applies complex network approach to study connectivity patterns in the global terrestrial water storage (TWS) data. The authors use two complementary TWS data bases – Gravity Recovery and Climate Experiment (GRACE) satellite mission data and the Global Land Data Assimilation (GLDAS) NOAH model – with the resolution of 1x1 degree. The network is constructed of the land nodes (cells of 1x1 degree) that might be connected pairwise depending on the correlation of the respective TWS time series. The authors analyze several network statistics – neighbor edge density, connectivity, and connection length – to characterize the principal river basins of the World and reveal some significant teleconnections in the global water dynamics. The findings of the study are consistent with the existing climate teleconnection literature, which supports the validity of the proposed approach. At the same time, the examined way of treating TWS time series seems to be novel for this type of data and may inform a range of studies focused on the global water cycle. In addition, the study quantifies the differences between the two examined global databases, which is an important independent contribution. The paper is clearly written and effectively organized. The main conclusions seem to be valid and robust with respect to the data noise and time series processing.

Reply: Thank you.

There are some minor remarks, which should be easily addressed by the authors:

1) p. 786, ll.1-2: The sentence is unclear. What is the definition of "relevant edges"? Probably the authors refer to the nodes (not edges) that are relevant to each other? This sentence should be revised.

Reply: Agree. The full sentence is rephrased to "*In the pruning step, an appropriate similarity threshold (\tau) is imposed to the edge set to retain only those connections that exceed the threshold value.*"

2) p. 787, ll.3-4: It is unclear what is meant by "all meaningful features". Please be more specific here.

Reply: Agree. The word "meaningful" is vague, we will change to significant. The full sentence is rephrased to "*Additional statistical analyses (see Section 4) are performed to ensure that all significant connections are retained in the constructed networks.*" The last paragraph of Section 4.1 also mentions "In this study, the threshold \tau is set to 0.57 because (a) the corresponding fraction of connected edges is relatively small (0.036), at which level more than 96% of edges is removed, (b) the edge densities of GRACE and GLDAS happen to be the same at that level; and importantly (c) the cutoff threshold is still below the maximum correlation exhibited at all separation distances, as suggested 10 by Fig. 1b. Thus, the selected \tau value ensures that all important network features are represented by the constructed networks."

3) p.787, l.10: "correlation between edge" probably means "correlation between nodes i and j"
Reply: Accept. Please see our reply to Comment #7 from Reviewer 1.

4) p. 787, l.10: Here and in other places: Please define what you mean by "correlation between time series". Is this Pearson cross-correlation at lag zero?

Reply: On p.787 L.10, the correlation coefficient doesn't have to be at lag zero, it can also mean maximum correlation. Therefore, we didn't specify the type of correlation there. Instead, on p.790, L18-19, we wrote "Note in the discussion below, R is calculated at zero lag unless otherwise specified."

5) p. 788, Eq. (5): Do we need this? There is a lot of network statistics that are not used in this study. Why does this particular one needs to be discussed with a dedicated equation?

Reply: We presented the classic average distance measure to give rationale for using the measure given in Eq 6. A direct application of Eq. 5 is computationally demanding when the number of nodes is large.

6) p. 788, Eq. (6): I'm not sure that this measure can quantify "average distance between node i and all other nodes". Do you have examples or theoretical argument in support of Eq. (6) being a proxy for Eq. (5)? Importantly, this statement is probably not necessary. Why not introducing Eq. (6) as a connectivity measure used in this study, without referring to the true average distance?

Reply: The description on p. 788, L. 10 is an oversight. We will rephrase it to "In this work, the average distance between node i and other nodes, L_i , is approximated according to…" because we only used the first neighbors of L_i to calculate the average distance. We feel that average (topological) distance is still an integration measure, instead of a connectivity measure as the reviewer suggested.

A reference for Eq. 6 can be found in Section 3.3.3 in Donges et al. 2009 [a],
Donges, J. F., Zou, Y., Marwan, N., and Kurths, J.: Complex networks in climate dynamics, Eur. Phys. J.-Spec. Top., 174, 157–179, 2009a.

7) p. 788, l. 12: "are included", not "is included"

Reply: Accept. Changed to "are included".

8) p. 789, l. 3: "linear interpolation". Do you refer to linear interpolation between the two neighboring values? Do you use deterministic or stochastic linear interpolation? What if more than one value in a row is missing? Please describe this process in more detail.

Reply: Linear interpolation is done between datasets of two months. It's temporal, not spatial. Such an approach has often been used in the GRACE community. Of course, one may seek to do cubic spline, but luckily as we mentioned on p789, L3, the missing months are not contiguous, so simple linear interpolation suffices for the mildly changing TWS time series.

9) p. 790, l. 24: Why using the maximum correlation coefficient is representative?

Reply: This is because the network similarity is based on correlation and we want to make sure the maximum correlation at all distances are included. As we show on p790, L8-10, the main purpose of Figure 1b is to show that "the cutoff threshold \tau is still below the maximum correlation exhibited at all separation distances." So from Figure 1b we see that the cutoff value is 0.57, while the maximum correlations at all distances are greater than 0.57. Therefore, we can at least say that the most statistically significant node pairs are included.

10) p. 791, ll. 8-9, item (c): How the cutoff is related to the distribution of the correlations? (This is related to my comment immediately above)

<u>Reply:</u> In the edge density method (Donges et al., 2009a), the cutoff is a correlation value. This should be clear after reading Section 4.1

11) p. 791, ll. 10-11: Please explain what you mean by "all important network features"

<u>Reply:</u> Please see our reply to Comment # 9 in the above.

Global terrestrial water storage connectivity revealed using complex climate network analyses

A. Y. Sun[1]**, J. Chen**[2]**, and J. Donges**[3,4]

[1]Bureau of Economic Geology, the Jackson School of Geosciences, University of Texas at Austin, University Station, Box X, Austin, Texas, USA
[2]Center for Space Research, University of Texas at Austin, Austin, Texas, USA
[3]Potsdam Institute for Climate Impact Research, Potsdam, Germany
[4]Stockholm Resilience Center, Stockholm University, Sweden

Correspondence to: A. Y. Sun (alex.sun@beg.utexas.edu)

Abstract

Terrestrial water storage (TWS) exerts a key control in global water, energy, and biogeochemical cycles. Although certain causal ~~relationships exist~~ relationship exists between precipitation and TWS, the latter quantity also reflects impacts of anthropogenic activities.
Thus, quantification of the spatial patterns of TWS will not only help to understand feedbacks between climate dynamics and hydrologic cycle, but also provide new insights and model calibration constraints for improving the current land surface models. ~~In this work , the~~ This work is the first attempt to quantify the spatial connectivity of TWS ~~is quantified using the climate~~ using the complex network theory, which has received broad attention in the climate modeling community in recent years. Complex networks of TWS anomalies are built using two global TWS datasets, a remote-sensing product that is obtained from the Gravity Recovery and Climate Experiment (GRACE) satellite mission, and a model-generated dataset from the global land data assimilation system's NOAH model (GLDAS-NOAH). Both datasets have $1° \times 1°$ grid resolutions and cover most global land areas except for permafrost regions. TWS networks are built by first quantifying pairwise correlation among all valid TWS anomaly time series, and then applying a ~~statistical cutoff threshold~~ cutoff threshold derived from the edge density function to retain only the most important features in the network. Basinwise network connectivity maps are used to illuminate connectivity of individual river basins with other regions. The constructed network degree centrality maps show ~~TWS~~ the TWS anaomly hotspots around the globe and the patterns are consistent with recent GRACE studies. Parallel analyses of networks constructed using the two datasets ~~indicate~~ reveal that the GLDAS-NOAH model captures many of the spatial patterns shown by GRACE, although significant discrepancies exist in some regions. Thus, our results provide ~~important insights for constraining~~ further measures for constraining the current land surface models, especially in data sparse regions.

1 Introduction

Terrestrial water storage (TWS) is defined as vertically integrated water of all forms above and below the Earth's surface (e.g., surface water, soil moisture, groundwater, and snow and ice) (Famiglietti, 2004). It is not only a key control of global water, energy, and bio-
5 geochemical cycles, but also provides an integrated indicator of water availability and uses (Houborg et al., 2012; Lettenmaier and Famiglietti, 2006; Long et al., 2013; Voss et al., 2013; Guentner et al., 2007). Global TWS has been the subject of modeling studies for decades, however, validation of modeling results has been challenging historically because of limited availability of in situ data. Since its launch in 2002, the Gravity Recovery and
10 Climate Experiment (GRACE) satellite mission has provided an unprecedented opportunity to study TWS remotely. GRACE detects temporal variations of the Earth's gravity field which, over land, are mainly caused by short-term variations or TWS anomalies (TWSA). Numerous studies conducted in the past decade have confirmed the remarkable capability of GRACE in tracking continental- and regional-scale TWS changes (e.g., Famiglietti et al.,
15 2011; Sun et al., 2010; Yeh et al., 2006; Long et al., 2013; Rodell et al., 2009; Swenson and Wahr, 2003; Han et al., 2005; Long et al., 2014). So far, the monthly TWSA grids derived from GRACE have been used as an independent source of information for hydrologic model validation (Ramillien et al., 2008; Syed et al., 2008; Chen et al., 2005), calibration (Sun et al., 2012; Werth et al., 2009; Lo et al., 2010; Sun et al., 2010; Döll et al., 2014), and
20 data fusion (Zaitchik et al., 2010; Houborg et al., 2012; Sun, 2013; Forman et al., 2012; ; Li and Rodell, 2015).

The global GRACE dataset accumulated over the last decade ~~represents~~ is an important type of Big Data that can be mined for discovering information of global water/energy dynamics, and for helping to illuminate connections among major river basins and within
25 the river basins themselves. Such information will be complementary to existing physicically based TWS modeling efforts and will potentially provide calibration constraints (e.g., Guentner et al., 2007; Rodell et al., 2004)~~and will potentially serve as calibration constraints~~. In this study, the complex network theory is adopted to ~~represent GRACE TWSA as a~~ network

3

with a large set of interconnected nodes. Patterns of TWS construct a global TWSA network using GRACE data. The interannual spatial patterns of TWSA are then quantified through analyses of network topologies.

Complex network theory has long been used by scientists in various disciplines to study intricate connections in natural and social phenomena (Jackson, 2008; Newman and Girvan, 2004; Rubinov and Sporns, 2010). In recent years, the complex climate network field of complex climate networks (CCN) theory, which is an extension of the , which involves applications of traditional complex network analyses to climate systems (Tsonis and Roebber, 2004; Tsonis et al., 2006), has attracted significant attention. In CCN theorytypical CCN applications, cells of a gridded dataset are treated deemed as nodes of a complex network, and links (or edges) between nodes are established on the basis of statistical similarity of the time series associated with the cells. After a climate network is constructed, various descriptive measures derived from the classical complex network theory are then applied to quantify network topologies (Donges et al., 2009b; Tsonis et al., 2006; Steinhaeuser et al., 2011). One of the major main findings from the previous CCN studies is that climate networks manifest a "small-world" network property, akin to networks appear in many other fields (e.g., social networks). In CCN, this can be contributed to the existence of long-range connections that stabilize the climate system and enhance energy transfers within it (Donges et al., 2009a, b, 2011). TWS is closely intertwined with soil-vegetation-atmosphere interactions and is thus expected to show similar spatiotemporal patterns as observed from climate networks (e.g., precipitation network); however, it is well known that climate only plays a partial role in TWS changes. Land use changes and other anthropogenic activities (e.g., deforestation, aquifer mining, and water structures) increasingly stress water availability in many parts of the world and have been shown to produce global-scale impacts on the terrestrial water cycle (Vörösmarty and Sahagian, 2000). Such aspects are usually difficult to be fully captured and quantified without extensive in situ monitoring data. The global coverage of GRACE TWSA, thus, becomes especially important.

Different from the global circulation model outputs analyzed by many previous CCN studies, GRACE TWSA is a remote sensing productthat is subjected to errors , subjecting to

4

errors and uncertainties caused by instrumentation and data processing. As a result, the actual spatial resolution of GRACE TWSA is not $1° \times 1°$, but much coarser (Houborg et al., 2012). In other words, the intrinsic degrees of freedom of the GRACE TWS are ~~much~~ less than its grid dimension. An important question is then how well a complex network con-
5 structed using the GRACE TWSA can ~~reflect~~ represent the salient features of the global terrestrial water cycle. Importantly, how these patterns can be corroborated, at least partially, using other existing information. Toward this end, we use the TWS dataset ($1° \times 1°$) simulated by global land data assimilation system (GLDAS) ~~is used~~ for comparison. GLDAS is a global terrestrial modeling system jointly developed by US National Aeronautics and
10 Space Administration's (NASA) Goddard Space Flight Center and US National Oceanic and Atmospheric Administration's National Centers for Environmental Prediction. GLDAS incorporates satellite and in situ observations to produce optimal fields of land surface states and fluxes in near real time (Rodell et al., 2004). Although GLDAS is only a surrogate of in situ observations that are ultimately required to validate the GRACE results, previous studies
15 have shown that GLDAS represents the magnitudes and variability of TWS sufficiently well (Syed et al., 2008). Thus, GLDAS represents a valuable independent source of information for validating GRACE results and has been used by a number of global-scale GRACE studies (e.g., Syed et al., 2008; Landerer and Swenson, 2012; Chen et al., 2005). In this study, the network measures inferred from GRACE data are compared to those built from
20 the GLDAS outputs to cross-examine the two products. Note that GLDAS does not have an explicit representation of groundwater storage, an aspect that needs to be kept in mind when performing comparisions.

2 Methodology

2.1 Network construction

25 A network is commonly represented by a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, which is specified by its node set $\mathcal{V} = \{1, \ldots, N\}$ and edge set $\mathcal{E}$, with N the number of nodes. Thus, the number of possible

5

edges in an undirected graph (meaning the links are non-directional) is $N(N-1)/2$. In the current context, each node corresponds to a grid cell at which a valid monthly time series is available and N is the total number of such cells in a gridded dataset. Construction of a network generally proceeds in two steps, network growth and pruning. In the network growth

5 step, similarity between all potential node pairs (i.e., edges) in graph $\mathcal{G}$ is quantified. Common measures of similarity are statistical correlation (either Pearson or Spearman), mutual information, and synchronization (Boers et al., 2013; Donges et al., 2009a). In the pruning step, an appropriate similarity threshold (τ) is imposed to the edge set to retain only those ~~edges that are considered relevant to each other, out of all possible edges~~connections that

10 exceed the threshold. The main purpose of network pruning is to improve network analysis efficiency. If the correlation between two time series is used as a measure of statistical similarity, then τ represents the minimum correlation coefficient (R) above which a pair of nodes is considered connected. The absolute value of correlation is used such that both strongly positive and negative correlations are counted.

15   Several methods have been used in the CCN literature to determine τ. In the significance testing method (Tsonis et al., 2006), τ is based on the two-sided Student's t test. The critical t value, t_c, for a given sample size n_s and user-defined significance level α are determined using the Student's t cumulative distribution function (CDF), from which the value of τ can be solved

20 $$t_c = \frac{\tau\sqrt{n_s-2}}{\sqrt{1-\tau^2}}. \tag{1}$$

A similar method uses the probability value (i.e., p value) of test statistics directly: a pair of nodes is considered connected if the p value is less than a critical value; for instance, Steinhaeuser et al. (2011) set the critical value to 10^{-10}. Yet another method defines τ from an edge density function $\rho(\tau)$ defined as

25 $$\rho(\tau) = \frac{n_c(\tau)}{N(N-1)/2}, \tag{2}$$

where n_c is the number of active edges retained in a network when the threshold is set to τ. Thus, edge density is closely related to the CDF of R.

Obviously, all methods involve certain degree of subjectivity. The selection of τ thus incurs a tradeoff between network maneuverability and preservation of network features: if too many edges are included, the main network features will be obscured, not to mention a significant increase in computational effort required to characterize a large network. In this work, the edge density method is used because it allows a direct comparison of network properties computed from different datasets (Donges et al., 2009a). Additional statistical analyses (see Sect. 4) are performed to ensure that all ~~meaningful~~ statistically significant features are retained in the constructed networks.

2.2 Network measures

The outcome of network construction process is a Boolean-valued, symmetric $N \times N$ matrix, referred to as the adjacency matrix and denoted by **A**. Elements of **A**, a_{ij}, are set according to the following rule

$$a_{ij} = \begin{cases} 1, & \text{if } |R_{ij}| > \tau \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

in which $|R_{ij}|$ is the absolute value of correlation between ~~edge (i,j)~~ time series at nodes i and j. A number of network measures can then be applied on **A** to quantify network topology. The main metrics adopted in this work include the degree of centrality and connection length.

The degree of centrality of a node, k_i, is defined as the number of first neighbors of node i and reflects the importance of node i in a network. Regions having high k_i values are referred to as "supernodes" in network theory because these nodes tend to have not only local connections, but also long-range connections or teleconnections. However, k_i itself does not reveal the actual type of connections. Because of nonuniformity of cell areas at different latitudes, the degree of centrality k_i is usually weighted by cell areas, leading to

7

the area-weighted connectivity, AC_i (Tsonis et al., 2006; Heitzig et al., 2012),

$$AC_i = \sum_{j \in n_i} \cos \lambda_j \Big/ \sum_{j=1}^{N} \cos \lambda_j, \quad i = 1, \ldots, N \tag{4}$$

where n_i is the set of all first neighbors of the node i, and λ_j is the latitude of its j-th first neighbor. Thus, AC_i is a normalized value representing the fraction of the Earth's surface area that a node is connected to.

A classic measure of network integration is the average distance between node i and all other nodes, D_i, and is defined as (Rubinov and Sporns, 2010)

$$D_i = \frac{1}{N-1} \sum_{j \in \mathcal{V}, j \neq i} d_{ij}, \quad i = 1, \ldots, N, \tag{5}$$

where d_{ij} is the number of edges traversed along the shortest path between node pair (i,j). If (i,j) is not connected, d_{ij} is defined as infinity. The characteristic path length of the network is obtained by taking average of all D_i and it represents the average number of edges to be traversed along the distance between two randomly selected nodes in a network. Calculation of pairwise shortest path lengths becomes computationally expensive when the number of node pairs is large. In this work, the average distance between node i and all other nodes, L_i, is ~~quantified~~ approximated according to

$$L_i = \frac{1}{k_i} \sum_{j \in n_i} l_{ij}, \tag{6}$$

where only the first neighbors of node i ~~is included~~ are included in the calculation, and l_{ij} is the physical distance between node pair (i,j) measured by using the respective cell-center latitudes and longitudes, (λ_i, ϕ_i) and (λ_j, ϕ_j). The physical-based characteristic path length of the network, $\overline{L}$, is simply the average of all L_i $(i = 1, \ldots, N)$. The probability distribution of L_i provides a sense of the average edge lengths in a network and $\overline{L}$ provides a measure of network integration.

3   Data and data processing

The GRACE TWSA dataset used in this study was downloaded from Jet Propulsion Laboratory (JPL)'s Tellus site, (http://grace.jpl.nasa.gov/index.cfm). The dataset is based on RL05 GRACE solutions (in the form of spherical harmonics) released by the Center for Space
5 Studies at the University of Texas Austin. It includes 121 epochs from January 2003 to July 2013 at approximately monthly intervals. The 6 missing months, which are not contiguous, were reconstructed using linear interpolation (temporal only). The grid dimensions are 360×180 and ocean area is masked out, resulting $\sim 25\,000$ cells in each TWSA grid. In generating the gridded TWSA product, a number of postprocessing algorithms have been
10 applied, as documented in details in Landerer and Swenson (2012). In particular, a destriping filter is applied to minimize the effect of north–south-oriented stripes in GRACE monthly solutions, and a 300 km Gaussian filter is then used to reduce random errors in high-degree spherical harmonic coefficients not removed by destriping. The GRACE gravity field solutions are typically truncated at a spectral degree less than 60. To restore signal attenuation
15 caused by truncation and filtering, the JPL dataset also includes a spatially distributed and temporally invariant scaling factor field. This scaling factor field is not used in this study because it does not affect pairwise correlations.

Outputs from GLDAS's NOAH model were obtained from NASA (http://disc.sci.gsfc.nasa.gov/services/grads-gds/gldas). GLDAS covers latitudes between -60° and 90°, and does
20 not model permafrost regions such as Greenland and Antarctica (Rodell et al., 2004). Its grid dimensions are 360×150 and the temporal span is from January 1979 to the present (GLDAS V1). The number of cells in each GLDAS monthly grid is $N = 14\,540$. The GLDAS TWS is defined as the sum of water mass from all four soil layers represented by NOAH (up to 2 m depth) and snow water equivalent. Thus, GLDAS TWS mainly includes surface
25 and root zone storages, but not the deeper groundwater stroage. The GRACE grids are masked using the smaller GLDAS coverage during network construction. To ensure a ~~fair~~ consistent comparison, the GLDAS dataset was processed using the same truncation and

filtering techniques applied to the GRACE data, which has been a standard practice in the literature (e.g., Chen et al., 2010; Rodell et al., 2009).

Monthly time series contains high-frequency noise. Because the main interest in this study is on interannual correlations of TWSA, the high frequency noise in each TWSA
5 time series are removed. Several methods have been used for such purpose. The z-score method has been employed in the CCN literature to remove seasonal variability (Donges et al., 2009b; Steinbach et al., 2003; Tsonis et al., 2006). It entails normalizing each monthly data point using the mean and standard deviation calculated for the corresponding month and over the entire record length. The least squares method, which is extensively used
10 in the GRACE literature (e.g., Yeh et al., 2006; Crowley et al., 2006), models the intraannual variability using Fourier series (two annual sine/cosine terms and two semi-annual sine/cosine terms) and then removes the variability, together with a slowly moving trend. Our numerical tests show the two methods give very similar results. Lags existing between time series may weaken linear correlation. Thus, to examine the effect of temporal lags, the
15 same interannual correlation analysis is repeated using a temporal window of 36 months (i.e., the maximum correlation observed within ± 1.5 years of the zero lag).

4 Results and discussion

4.1 Edge density

The number of possible edges represented by the TWS datasets is more than 100 million for
20 $N = 14\,540$. After removing seasonal trends from GRACE and GLDAS and calculating the correlation coefficient R for all node pairs, the edge density method is applied to determine a similarity threshold τ. Note in the discussion below, R is calculated at zero lag unless otherwise specified.

Figure 1a shows edge density functions constructed using GRACE and GLDAS TWS
25 data, respectively, both are monotonically decreasing (i.e., fewer connected edges at higher τ values) and are similar in shape. As mentioned in Sect. 2, edge density provides an in-

10

dicator of the fraction of connected edges at different threshold values. Figure 1b plots the maximum correlation coefficient as a function of edge length, which is defined as the shortest physical distance between a pair of nodes in this work. To arrive at Fig. 1b, all R values are first sorted according to nodal separation distances, a bin width of 250 km is applied to

5 the resulting distribution, and the maximum R value within each bin is recorded. Figure 1b suggests that the maximum correlation stays relatively high (> 0.7) for most distances. Recall that the main purpose of network pruning is to improve the computational efficiency of network characterization while preserving the most ~~important~~ significant network features. In this study, we set the threshold τ ~~is set~~ to 0.57 because (a) the corresponding fraction

10 of connected edges is relatively small (0.036), at which level more than 96 % of edges is removed, (b) the edge densities of GRACE and GLDAS happen to be the same at that level; and importantly (c) the cutoff τ threshold is still below the maximum correlation exhibited at all separation distances, as suggested by Fig. 1b. Thus, the selected τ value ensures that all ~~important~~ statistically signficant network features are ~~represented by~~ retained in the

15 constructed networks.

4.2 Basin analyses

A basin analysis is useful for helping visualize the TWSA connection patterns at the basin level. As some examples, Figure 2 shows the results for six river basins around the world. To generate a plot in Fig. 2, a cell is first fixed, and all its edges are colored according

20 to the actual R (not the absolute values). For our purpose, the centroid of each basin is used. While the basin centroid may not be representative of the connection patterns of an entire basin (especially when the basin spans several climatic regions), it serves as a basis for comparing multiple basins at a qualitative level. Figure 3 applies the cutoff threshold τ defined in Fig. 1 to all plots in Fig. 2. Results suggest that interannual TWSA connections

25 in Amazon and Congo Basins are dominated by local connections. The mid-latitude basins (Ganges, Mississippi, and Tigris) generally show more teleconnections, although Yangtze is an exception. In the case of Tigris basin, a large number of strongly positive and negative

correlations are observed and the local connections extend far beyond the basin boundary. A detailed interpretation of this observation will be given in the next section.

Extensive teleconnection is an advantage from forecasting perspective because climate indices, such as El Niño–Southern Oscillation (ENSO) and North Atlantic Oscillation (NAO),
can be used as possible indicators of future changes. For those basins without strong teleconnection, water resources planning must rely mainly on regional data. Such distinction sheds light on the significance of GRACE data to long-term basin planning and natural hazard mitigation strategies, as we will elaborate in the following sections.

As a sensitivity study, Fig. 4 (left column) shows the results of basin analysis for Mississippi basin, the largest basin in North America, using different thresholds corresponding to τ values of 0.41, 0.57 (the base case), and 0.76, respectively. The corresponding edge density is labeled in the figure. Because the cutoff threshold increases as ρ decreases, a significant reduction in number of edges can be observed. For comparison, the modeled TWS connections obtained from GLDAS are provided in the second column of Fig. 4. In general, the connections modeled by GLDAS are much weaker (i.e., smaller in spatial extent) than those obtained from GRACE. In some cases, the locations of connections are also different. For example, the negative correlation obtained by GLDAS in North Africa for $\rho = 0.1$ is not seen by GRACE. The complex networks thus provide a useful tool for examining the agreement, or the lack of it, between GLDAS and GRACE.

4.3 Connectivity

Using the selected cutoff τ, a network adjacency matrix **A** is formed and various network measures described under Sect. 2 are applied to quantify network topology. Figure 5a shows the area-weighted connectivity map constructed using GRACE data. On the map, red colors highlight regions of high connectivity. Recall that a high degree of connectivity indicates that a node interacts strongly with the rest of the nodes in a network (i.e., a supernode); however, the connectivity map itself does not tell the type of connections per se, and needs to be analyzed jointly with the connection length map to be shown in the next section. The largest cluster of supernodes appears in the Middle East region, where the

connected neighbors account for more than 0.16 of the global area. To a lesser extent, the Pacific Northwest and east coast of the US, southern Africa, southern South America, and eastern Australia show smaller supernode regions. In contrast, most of Asia, central US, and Europe exhibit little or no connectivity (blue color). These observations are consistent with patterns observed during basin analyses (see Figs. 3 and 4).

The supernode regions shown in Fig. 5a reflect the superposed effects of climate variations and anthropogenic activities. These can be explained in terms of global precipitation and atmospheric circulation patterns. In general, the poorly connected regions have stronger precipitation variations over shorter spatial scales, leading to the emergence of high precipitation gradients which, in turn, are responsible for regional extreme events that are more localized in time and space (Scarsoglio et al., 2013). Those with high connectivity tend to be directly influenced by ocean and climatic oscillations (e.g., ENSO and NAO). Kahya and Dracup (1993) studied streamflow variations in the contiguous US and identified Northeast, North Central, Pacific Northwest, and Gulf of Mexico states as regions with potentially significant streamflow responses to ENSO forcing. These four regions can be easily identified on Fig. 5a, among which the Gulf of Mexico region shows the weakest connection. Similarly, Chiew et al. (1998) reported that the ENSO can be used to help forecast spring runoff in south-east Australia and summer runoff in the north-east and east coasts of Australia. This teleconnection pattern is also indicated clearly by Fig. 5a.

At the global scale, Dai et al. (2009) studied the monthly streamflow records of the world's 925 largest ocean-reaching rivers from 1948 to 2004. They concluded that (a) the interannual variations of streamflows are correlated with the ENSO events for discharge into the Atlantic, Pacific, Indian, and the global ocean as a whole and (b) the effects of anthropogenic activities on annual streamflow are likely to be small compared to those of climate variations; however, anthropogenic activities can create more disturbances in arid and semi-arid regions, where the discharge magnitudes are low (e.g., Indus, Yellow, and Tigris–Euphrates River Basin). To elaborate the latter point further, Fig. A1 in Appendix A plots the proportion of total renewable water resources withdrawn by country for human uses in the agricultural, municipal, and industrial sectors, using long-term data compiled by the Food Agricultural

Organization of United Nations. Figure A1 indicates that the Middle East and North African countries show the highest withdraw proportions. In a recent GRACE study focusing on north-central Middle East, Voss et al. (2013) reported that GRACE data show an "alarming rate" of decrease in TWS of approximately 143.6 km³ during 2003–2009. Thus, the
5 resemblance between Fig. 5a and Fig. A1 in those regions is not coincidental and can be corroborated using multiple sources. Because interannual TWS anomalies are well connected in clustered supernode regions, these regions tend to exhibit more vulnerability to both climate and human-induced disturbances.

Having elaborated the close relationship between GRACE TWSA and climate patterns,
10 it is important to point out that the TWS also includes effects of soil moisture and groundwater storage (mostly unconfined aquifers) changes that may not synchronize with climate patterns.

Figure 5b shows the same area-weighted connectivity map, but constructed using the GLDAS-NOAH outputs. Although GLDAS-NOAH shows many of the similar patterns de-
15 tected by GRACE, it also indicates stronger connectivity in Arabian Peninsula, North Africa, and in middle South America, and much weaker connectivity in southern Africa. These discrepancies may be caused by GLDAS-NOAH's parameterization and other errors. The other main reason is ~~that GLDAS does not resolve groundwater storage well~~the lack of representation of the deeper groundwater storage in GLDAS. The discrepancies highlighted
20 here provide additional spatial calibration constraints for land surface models. In areas dominated by shallow TWS components, GLDAS needs to show similar patterns as those derived from GRACE, whereas discrepancies are only expected in areas dominated by deep TWS components and/or impacted by signficant anthropogenic activities. We emphasize here the connectivity maps shown in Fig. 5 are for TWSA. Thus, the high-precipitation
25 areas (e.g., Amazon basin) do not necessarily exhibit high anomaly connectivity after removing the intraannual variability.

So far, all results have been based on zero-lag correlations. The effect of temporal lag on connectivity is examined in Fig. 6, in which the connectivity map is built using the maximum (absolute) correlation found between −18 and +18 monthly lags of each node pair. The

14

figure suggests that incorporation of lagged correlation further strengthens connectivity. The supernode regions are more expanded in space, notably in eastern Australia and in the Colorado River Basin and Gulf Coast states in the US. Further, Appendix B shows the maximum correlation and phase lags for the six basins studied in Fig. 2, which suggest
5 that each river basin is in phase with most cells in itself and the immediate surroundings. However, significant phase lags exist between each river basin and other river basins.

4.4 Connection length

Figure 7a shows maps of the physical-based average nodal connection length L_i ($i = 1, \ldots, N$). Nodes that exhibit the longest connection lengths are mostly located in south-
10 ern part of South America ($\sim 12\,000$ km). Other regions with relatively long connections are found in Pacific Northwest, North Central, Colorado River, and North East regions of the US, south Africa, and eastern Australia. Interestingly, the Middle East region is mostly characterized by connection lengths less than 5000 km; thus, the supernodes in that region are dominated by local connections. The connection length patterns observed here support the
15 previous discussions in the context of teleconnection and forecasting potential. Importantly, the connection length map can help evaluate the influence of teleconnection on TWS for a particular region.

The average nodal connection length map constructed using GLDAS data suggests much wider connections, although most are local. Again this can be attributed to model param-
20 eterization schemes, forcing resolution, and spatial correlation constraints, as discussed before.

The probability distribution of the average connection length, L_i, is shown in Fig. 8. Most nodes in the GRACE network are dominated by short-range edges with lengths less than 2000 km, although several other smaller modes appear in the 4000–6000, 6000–8000, and
25 8000–10 000 km ranges. In contrast, the GLDAS network shows a weaker local connection mode in < 2000 km range, but a wider and more persistent second mode in 4000–6000 km. Interestingly, the two modes of GLDAS coincide with those of GRACE. The characteristic path length ($\overline{L}$) is 2300 km for GRACE and 4000 km for GLDAS, respectively.

15

5   Summary and conclusions

In this work, the ~~CCN~~ complex network theory is applied to analyzing spatial connection patterns in TWS. A comparative study is conducted using two global TWS datasets derived from GRACE and GLDAS, respectively, with an emphasis on interannual variability. Both
5 datasets are large and have more than 100 million potential connections. An edge-density method is adopted to define an appropriate network pruning threshold. The constructed networks are further analyzed using the ~~classic~~ degree of centrality and connection length measures~~, which are extended for use with gridded datasets~~.

Our results show that ~~CCN theory provides a powerful tool for characterizing~~ complex
10 networks and GRACE TWSA can be used to identify global TWSA hotspots or supernode regions. The area-weighted connectivity is a local measure that reveals nodes with a large number of connections (edges), whereas the connection length helps identify the dominating type of connections ~~.~~ (i.e., local connections vs. teleconnections). In terms of connectivity, the largest cluster of supernodes appears in the Middle East region, while other promi-
15 nent ones are found in Pacific Northwest and eastern US, southern Africa, southern South America, and eastern Australia. In terms of connection lengths, the Middle East region is dominated by local connections, whereas regions such as Pacific Northwest, North Central, Colorado River, and North East regions of the US, south Africa, and eastern Australia all have strong bimodal connections.

20 While many of the TWSA network features found here can be explained by established climate teleconnection theories, the TWS, as an integrated indicator of global water storage, is unique in its own way. It shows the impact of both climate and anthropogenic activities. Knowledge of both the strength and type of TWS connectivity can help identify useful TWS predictors and provide insight to further improve current land surface models.

25 GLDAS outputs have been used extensively in validating GRACE results at various scales. Less focused is the consistency of spatial correlation ~~between~~ represented by GLDAS and GRACE data. Results from this study statistically quantify the similarity and discrepancies between the two datasets. In this case, the observed discrepancies may

be attributed to missing surface and groundwater components in the GLDAS model, or to GRACE uncertainties (Syed et al., 2008; Li and Rodell, 2015). Although data assimilation has been used to reduce discrepancies in land surface models, the geometrical, spatial connection patterns have not been used before. A main conclusion from this work is that

5 network connectivity measures should be incorporated as an additional model calibration and validation criterion when developing the future-generation of GLDAS models.

Appendix A

According to FAO, the proportion of total renewable water resources withdrawn is defined as the total volume of fresh groundwater and surface water withdrawn from their sources for

10 human use (in the agricultural, municipal and industrial sectors), expressed as a percentage of the total actual renewable water resources. The data used in Fig. A1 are compiled from 2005 data published by FAO http://www.fao.org/nr/aquastat. In several cases where 2005 data are not available, 2000 data are used as best estimates.

Appendix B

15 Fig. B1 shows the maximum correlations for the six basins chosen in Fig. 2, and Fig. B2 shows the corresponding phase lags. Recall these plots show the correlation between each basin centroid and all other cells in the TWSA dataset. The phase lag plot (normalized by 18 months) shows that each river basin is in phase with most cells in itself and the immediate surrounding regions, but there can be significant phase shifts between each river basin and

20 other river basins.

17

References

Boers, N., Bookhagen, B., Marwan, N., Kurths, J., and Marengo, J.: Complex networks identify spatial patterns of extreme rainfall events of the South American Monsoon System, Geophys. Res. Lett., 40, 4386–4392, 2013.

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

Chen, J. L., Wilson, C. R., Tapley, B. D., Longuevergne, L., Yang, Z. L., and Scanlon, B. R.: Recent La Plata basin drought conditions observed by satellite gravimetry, J. Geophys. Res., 115, D22108,
10 doi:10.1029/2010JD014689, 2010.

Chiew, F. H. S., Piechota, T. C., Dracup, J. A., and McMahon, T. A.: El Nino/Southern Oscillation and Australian rainfall, streamflow and drought: links and potential for forecasting, J. Hydrol., 204, 138–149, doi:10.1016/S0022-1694(97)00121-2, 1998.

Crowley, J. W., Mitrovica, J. X., Bailey, R. C., Tamisiea, M. E., and Davis, J. L.: Land water stor-
15 age within the Congo Basin inferred from GRACE satellite gravity data, Geophys. Res. Lett., 33, L19402, doi:10.1029/2006gl027070, 2006.

Dai, A., Qian, T., Trenberth, K. E., and Milliman, J. D.: Changes in continental freshwater discharge from 1948 to 2004, J. Climate, 22, 2773–2792, doi:10.1175/2008JCLI2592.1, 2009.

Döll, P., Fritsche, M., Eicker, A., and Schmied, H. M.: Seasonal water storage variations as impacted
20 by water abstractions: comparing the output of a global hydrological model with GRACE and GPS observations, Surv. Geophys., 35, 1311–1331, 2014.

Donges, J. F., Zou, Y., Marwan, N., and Kurths, J.: Complex networks in climate dynamics, Eur. Phys. J.-Spec. Top., 174, 157–179, 2009a.

Donges, J. F., Zou, Y., Marwan, N., and Kurths, J.: The backbone of the climate network, Europhys.
25 Lett., 87, 48007, doi:10.1209/0295-5075/87/48007, 2009b.

Donges, J. F., Schultz, H. C., Marwan, N., Zou, Y., and Kurths, J.: Investigating the topology of interacting networks, Eur. Phys. J. B, 84, 635–651, 2011.

Famiglietti, J. S.: Remote sensing of terrestrial water storage, soil moisture and surface waters, Geoph. Monog. Series, 197–207, doi:10.1029/150GM16, 2004.

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

18

Forman, B. A., Reichle, R., and Rodell, M.: Assimilation of terrestrial water storage from GRACE in a snow-dominated basin, Water Resour. Res., 48, W01507, doi:10.1029/2011WR011239, 2012.

Guentner, A., Stuck, J., Werth, S., Doell, P., Verzano, K., and Merz, B.: A global analysis of temporal and spatial variations in continental water storage, Water Resour. Res., 43, W05416, doi:10.1029/2006WR005247, 2007.

Han, S. C., Shum, C., Jekeli, C., and Alsdorf, D.: Improved estimation of terrestrial water storage changes from GRACE, Geophys. Res. Lett., 32, L07302, doi:10.1029/2005GL022382, 2005.

Heitzig, J., Donges, J. F., Zou, Y., Marwan, N., and Kurths, J.: Node-weighted measures for complex networks with spatially embedded, sampled, or differently sized nodes, Eur. Phys. J. B, 85, 1–22, 2012.

Houborg, R., Rodell, M., Li, B., Reichle, R., and Zaitchik, B. F.: Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations, Water Resour. Res., 48, W07525, doi:10.1029/2011wr011291, 2012.

Jackson, M. O.: Social and Economic Networks, Princeton University Press, Princeton, NJ, XIII, 504 pp., 2008.

Kahya, E. and Dracup, J. A.: US streamflow patterns in relation to the El Niño/Southern Oscillation, Water Resour. Res., 29, 2491–2503, 1993.

Landerer, F. W. and Swenson, S. C.: Accuracy of scaled GRACE terrestrial water storage estimates, Water Resour. Res., 48, W04531, doi:10.1029/2011wr011453, 2012.

Lettenmaier, D. P. and Famiglietti, J. S.: Hydrology: water from on high, Nature, 444, 562–563, 2006.

Li, B. and Rodell, M.: Evaluation of a model-based groundwater drought indicator in the conterminous US, J. Hydrol., 526, 78-88, 2015.

Lo, M. H., Famiglietti, J. S., Yeh, P. J. F., and Syed, T. H.: Improving parameter estimation and water table depth simulation in a land surface model using GRACE water storage and estimated base flow data, Water Resour. Res., 46, W05517, doi:10.1029/2009wr007855, 2010.

Long, D., Scanlon, B. R., Longuevergne, L., Sun, A. Y., Fernando, D. N., and Himanshu, S.: GRACE satellites monitor large depletion in water storage in response to the 2011 drought in Texas, Geophys. Res. Lett., 40, 3395–3401, doi:10.1002/grl.50655, 2013.

Long, D., Shen, Y., Sun, A., Hong, Y., Longuevergne, L., Yang, Y., Li, B., and Chen, L.: Drought and flood monitoring for a large karst plateau in Southwest China using extended GRACE data, Remote Sens. Environ., 155, 145–160, doi:10.1016/j.rse.2014.08.006, 2014.

Newman, M. E. and Girvan, M.: Finding and evaluating community structure in networks, Phys. Rev. E, 69, 026113, doi:10.1103/PhysRevE.69.026113, 2004.

Ramillien, G., Famiglietti, J., and Wahr, J.: Detection of continental hydrology and glaciology signals from GRACE: a review, Surv. Geophys., 29, 361–374, 2008.

Rodell, M., Houser, P., and Jambor, U.: The global land data assimilation system, B. Am. Meteorol. Soc., 85, 381–394, 2004.

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

Rubinov, M. and Sporns, O.: Complex network measures of brain connectivity: uses and interpretations, Neuroimage, 52, 1059–1069, 2010.

Scarsoglio, S., Laio, F., and Ridolfi, L.: Climate dynamics: a network-based approach for the analysis of global precipitation, PloS one, 8, e71129, doi:10.1371/journal.pone.0071129, 2013.

Steinbach, M., Tan, P.-N., Kumar, V., Klooster, S., and Potter, C.: Discovery of climate indices using clustering, in: Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 24–27 August, 446–455, 2003.

Steinhaeuser, K., Chawla, N. V., and Ganguly, A. R.: Complex networks as a unified framework for descriptive analysis and predictive modeling in climate science, Statistical Analysis and Data Mining, 4, 497–511, 2011.

Sun, A. Y.: Predicting groundwater level changes using GRACE data, Water Resour. Res., 49, 5900–5912, 2013.

Sun, A. Y., Green, R., Rodell, M., and Swenson, S.: Inferring aquifer storage parameters using satellite and in situ measurements: estimation under uncertainty, Geophys. Res. Lett., 37, L10401, doi:10.1029/2010gl043231, 2010.

Sun, A. Y., Green, R., Swenson, S., and Rodell, M.: Toward calibration of regional groundwater models using GRACE data, J. Hydrol., 422–423, 1–9, doi:10.1016/j.jhydrol.2011.10.025, 2012.

Swenson, S. and Wahr, J.: Monitoring changes in continental water storage with GRACE, Space Sci. Rev., 108, 345–354, 2003.

Syed, T. H., Famiglietti, J. S., Rodell, M., Chen, J., and Wilson, C. R.: Analysis of terrestrial water storage changes from GRACE and GLDAS, Water Resour. Res., 44, W02433, doi:10.1029/2006WR005779, 2008.

Tsonis, A. A. and Roebber, P. J.: The architecture of the climate network, Physica A, 333, 497–504, 2004.

Tsonis, A. A., Swanson, K. L., and Roebber, P. J.: What do networks have to do with climate?, B. Am. Meteorol. Soc., 87, 585–595, 2006.

20

Vörösmarty, C. J. and Sahagian, D.: Anthropogenic disturbance of the terrestrial water cycle, Bioscience, 50, 753–765, 2000.

Voss, K. A., Famiglietti, J. S., Lo, M., de Linage, C., Rodell, M., and Swenson, S. C.: 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, 904–914, doi:10.1002/wrcr.20078, 2013.

Werth, S., Güntner, A., Petrovic, S., and Schmidt, R.: Integration of GRACE mass variations into a global hydrological model, Earth Planet. Sc. Lett., 277, 166–173, 2009.

Yeh, P. J. F., Swenson, S. C., Famiglietti, J. S., and Rodell, M.: 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, 2006.

Zaitchik, B. F., Rodell, M., and Olivera, F.: Evaluation of the Global Land Data Assimilation System using global river discharge data and a source-to-sink routing scheme, Water Resour. Res., 46, W06507, doi:10.1029/2009WR007811, 2010.
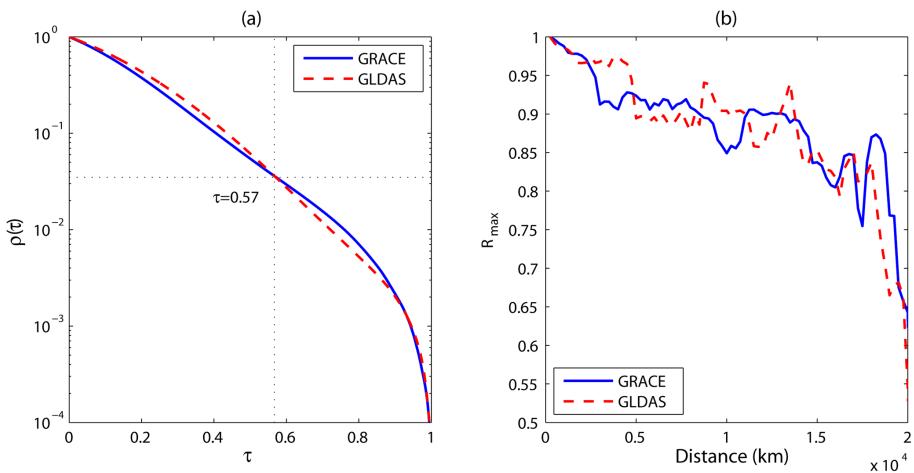
Figure 1. (a) Edge density function $\rho(\tau)$ of GRACE and GLDAS (the value of τ selected for network pruning is 0.57, corresponding to an edge density 0.036); **(b)** maximum correlation as a function of edge lengths.
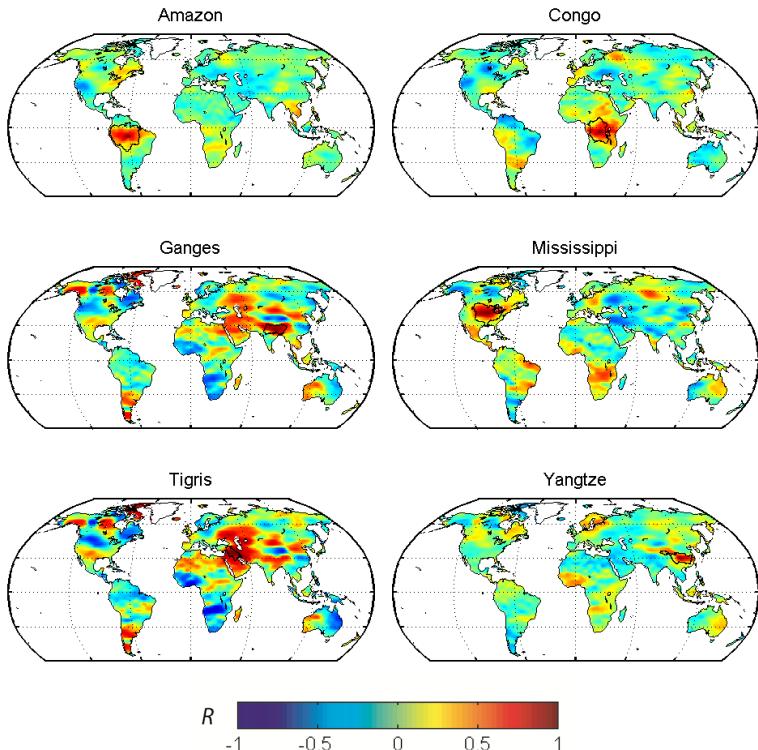
Figure 2. Patterns of connection inferred from GRACE TWSA for six river basins, in which connection pattern is based on correlation between the basin centroid and all other cells in the grid.
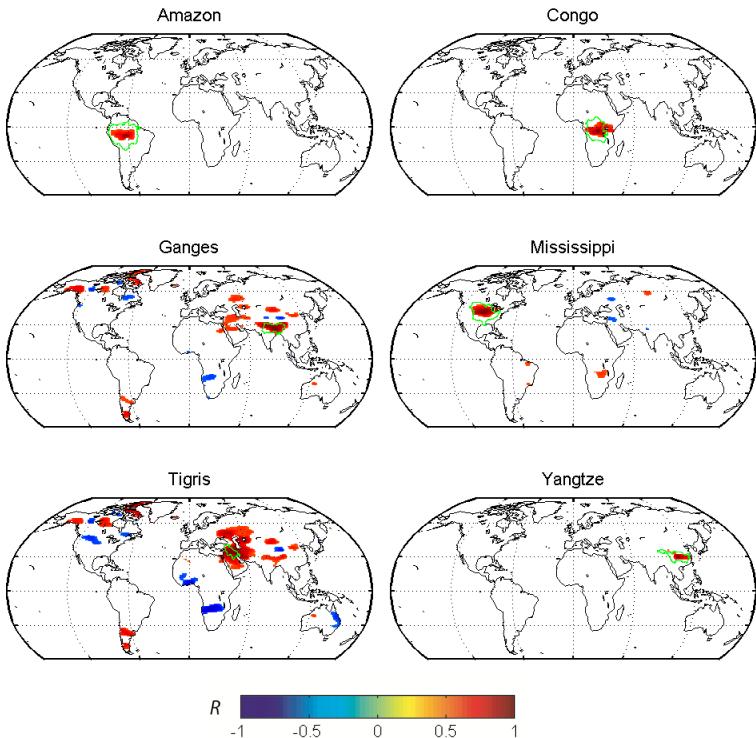
Figure 3. GRACE connection patterns after cutoff threshold $\tau = 0.57$ is applied (the green solid line delineates basin boundaries).
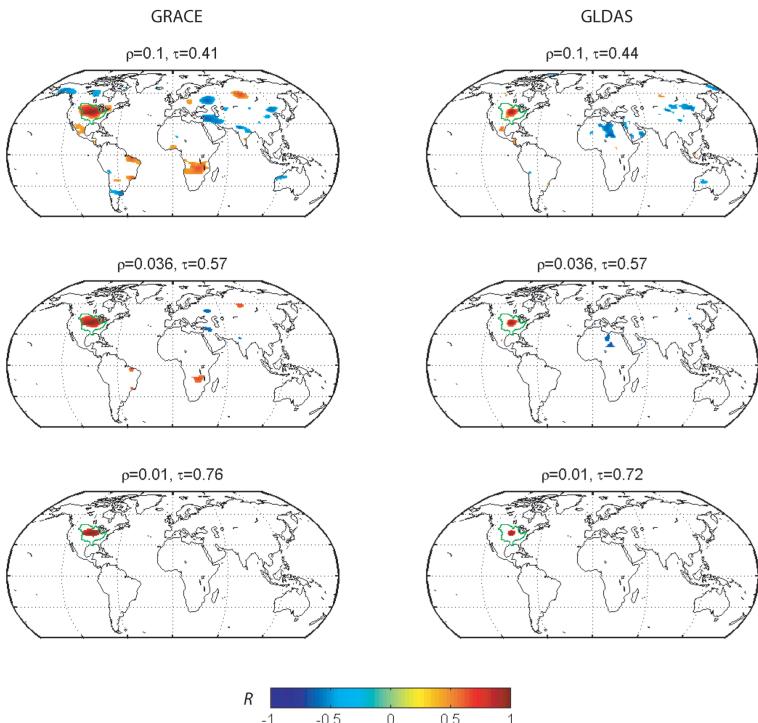
Figure 4. Sensitivity of connection patterns to cutoff threshold, demonstrated using Mississippi River Basin's centroid. Left column, GRACE results; right column, GLDAS results.
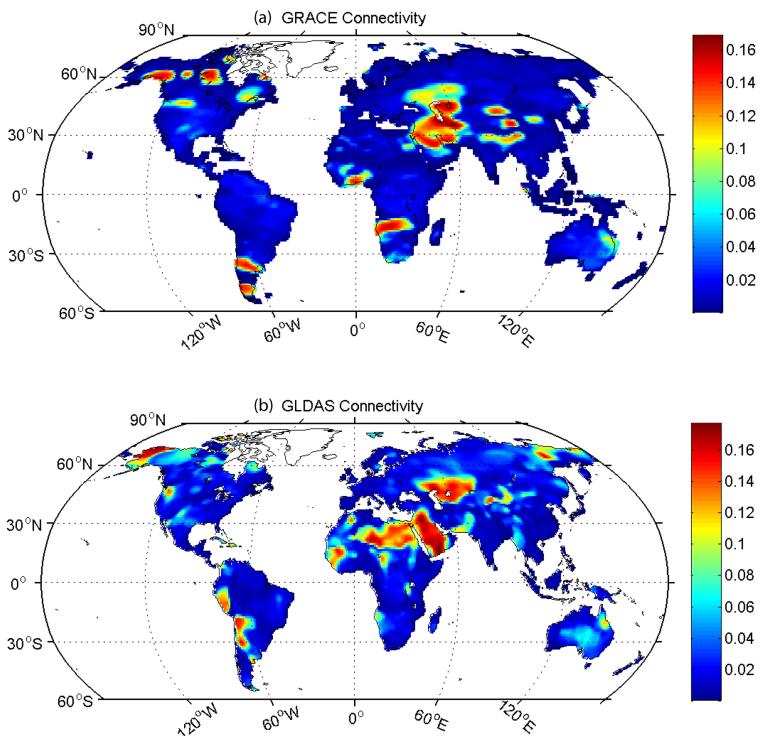
Figure 5. Area-weighted connectivity map obtained using **(a)** GRACE and **(b)** GLDAS data (zero-lag correlation).
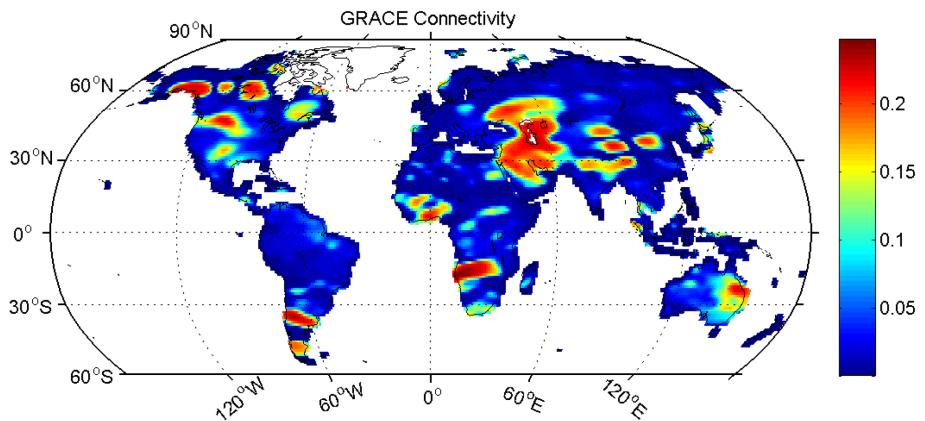
Figure 6. Effect of lagged-correlation on GRACE area-weighted connectivity, where the window of lagged correlation is [−18, 18] months.
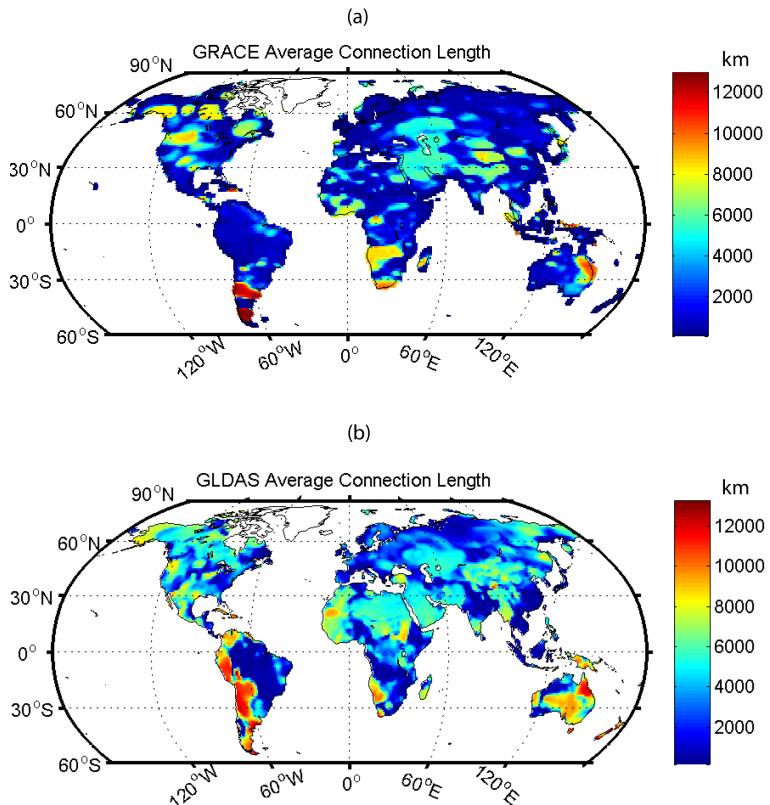
Figure 7. Map of average node connection lengths derived based on **(a)** GRACE and **(b)** GLDAS.

28
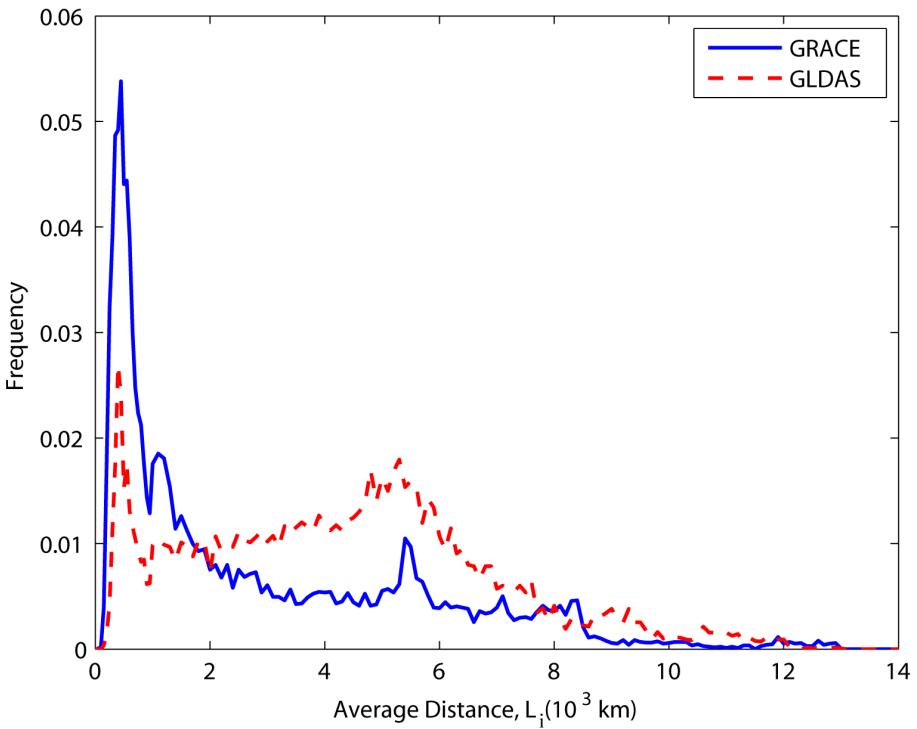
Figure 8. Distribution of average edge lengths in GRACE and GLDAS networks, where L_i denotes the average distance between node i and its neighbors.
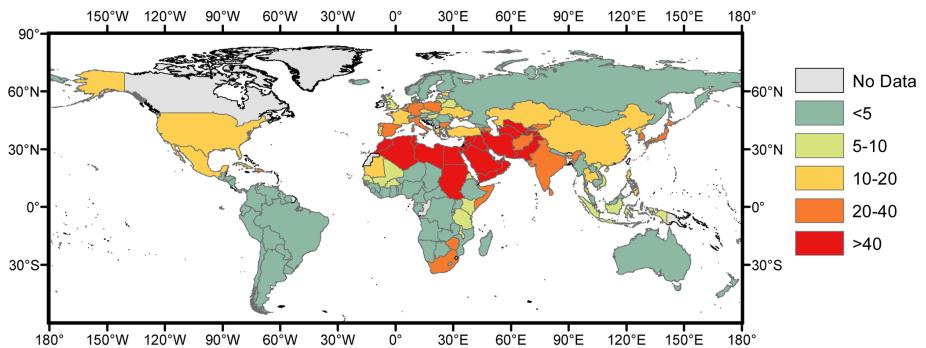
Figure A1. Proportion of total renewable water resources used by country (Data source: Food Agricultural Organization (FAO) of the United Nations, http://www.fao.org/nr/aquastat)

Figure B1. Degree centrality inferred from GRACE TWSA for six river basins, based on the maximum correlation between each basin centroid and all other cells in the grid, and within a window of [−18,18] months.

Figure B2. Phase lag of maximum correlations obtained for the six river basins shown in Figure B1 (normalized by 18).