Interactive comment on "*Groundwater storage dynamics in the world's large aquifer systems from GRACE: uncertainty and role of extreme precipitation*" by Mohammad Shamsudduha and Richard G. Taylor

Marc Bierkens (Referee #1)

bierkens@geo.uu.nl
Received and published: 2 January 2020

Reviewer's comments are italicised, and our responses (R) and revisions (REVISION) are provided in normal fonts.

General comments

The authors use the results of three different GRACE-based TWS methods and 4 Land surface models to generate an ensemble of groundwater storage anomalies. These are subsequently analyzed by a non-parametric statistical method to separate seasonal signals from non-linear trends and residuals.

The main message of the paper is that trends in GWS anomalies (ΔGWS), if existing, are non-linear in the vast majority of main aquifer systems and that rainfall anomalies play an important role in explaining these non-linear trends.

I enjoyed reading the paper. I find that it is a well-written with an important message that deserves publication. However, I have a few comments.

Moderate comments:

1. *I find the lack of reference to estimates based on global hydrological models (GHMS) remarkable. The first spatially distributed global assessment of depletion rates where based on such models and, albeit indirect, should be used in the discussion. They are the basis for the "narratives on global groundwater depletion" that are mentioned in the discussion and the abstract (See https://iopscience.iop.org/article/10.1088/1748-9326/ab1a5f/meta for an overview of these studies). This is the more remarkable, given that the authors do use Land Surface Models (LSMs) to estimate ΔGWS from GRACE ΔTWS.*

Response to Reviewer 1 (R1) #1. Firstly, we thank the reviewer (Professor Bierkens) for his positive comments regarding the manuscript and for providing very constructive suggestions to improve the manuscript. Regarding the lack of reference to global hydrological models (GHMs) in "narratives on global groundwater depletion", we agree that this is a critical omission from the original manuscript, which focused on uncertainty in the estimation of GRACE-derived ΔGWS that is typically reliant on estimates of components of terrestrial storage from LSMs (e.g. Long et al., 2016) including commonly used models from NASA's Global Land Data Assimilation System (GLDAS). The revised manuscript will engage fully and directly with evidence from GHMs in describing narratives of "global groundwater depletion" including the recommended study by Bierkens and Wada (2019) and references therein (e.g. Wada et al., 2010; de Graaf et al., 2017).

REVISION R1 #1: see revised text on page 1 and lines 18-21, page 16 and lines 378-384, page 17 and lines 397-398, and page 19-20 and lines 458-466; also see the Conclusions on page 22 and lines 508-509.

2. *Regarding the estimation of ΔGWS from GRACE ΔTWS (Equation 1): I am quite doubtful that the surface water storage from integrating LSM runoff on a monthly basis is sufficiently accurate. Even a small basin as the Rhine has a discharge peak routing time of a week, while that of the Amazon amounts to 3 months. Apart from the lack of*

river routing, GLDAS LSMs do not include the storage and delayed discharge from reservoirs, lakes and inundated floodplains (GHMs do a better job in that respect; see https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL081836). This fact may lead to underestimation of ΔSWS and subsequently an overestimation of ΔGWS and its noisiness. Granted, comparison with piezometric data in the Limpopo and the Ganges-Brahmaputra is favourable, but this can be scaled easily by changing specific yield.

R1 #2. We agree with the reviewer's concerns regarding the use of GLDAS LSMs to account for surface water storage (ΔSWS) in the estimation of ΔGWS from GRACE, highlighted recently by Scanlon et al. (2019). The original manuscript first notes that most GRACE studies do not account for ΔSWS in the computation of ΔGWS with the assumption that its contribution to ΔTWS is limited. Consistent with previous studies (e.g. Bhanja et al., 2016; Thomas et al., 2017), this study applies time-series simulations of surface runoff from GLDAS LSMs as a proxy for ΔSWS in the absence of global-scale time-series monitoring of surface water storage changes in rivers, lakes, floodplains and reservoirs. Recognition of the problem of routing in the use of GLDAS LSM data for ΔSWS, identified by Reviewer 1, and its implications for the computation of ΔGWS will be made explicit in the revised manuscript. Together with this, we will expand related discussion of the performance of GHMs and LSMs related to GRACE ΔTWS as reviewed recently by Scanlon et al. (2018).

REVISION R1 #2: see revised text on page 16 and lines 378-384.

3. *The discussion related to the "narrative of global groundwater depletion" needs elaboration:*

   - *Not only piezometric studies show that groundwater depletion can be very local; this is also true for <u>model-based estimates of groundwater depletion</u>. See for instance results from Wada et al. 2012 (Figure S5) https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2012GL051230 and De Graaf et al 2017 (Figure 11): https://www.sciencedirect.com/science/article/pii/S030917081630656X#fig0011*

     This means that at the aquifer scale anomalous rainfall may cause an overall increase in groundwater storage, while groundwater depletion may locally still persist. Thus, the "<u>narrative of global groundwater depletion</u>" pertains to "<u>groundwater depletion as a global phenomenon</u>".

R1 #3a. We thank Reviewer 1 for their argument that it is not only piezometry but also global-scale modelling that shows that groundwater depletion can be localised so that depletion can occur alongside accumulation in the same (large) aquifer system. The revised manuscript will incorporate evidence from GHMs explicitly in its discussion of the nature of groundwater depletion assessed globally.

REVISION R1 #3a: see revised text on page 1 and lines 18-21, page 16 and lines 378-384, page 17 and lines 397-398, and page 19-20 and lines 458-466. also see the Conclusions on page 22 and lines 508-509.

   - *The current consensus seems to be that global ΔTWS has been increasing between 1950-1995 by dam building, decreasing from 1995-2005 by groundwater depletion and has been increasing again since then by increased land water storage due to a climate-change induced increase in precipitation: see the review by Wada et al: https://link.springer.com/chapter/10.1007%2F978-3-319-56490-6_7. Yet, at the same time groundwater depletion at the current hotspots has persisted. How do your findings relate to these insights?*

R1 #3b. We thank R1 for raising an interesting point on the dynamic nature of global ΔTWS due to spatiotemporal variability in anthropogenic influences including climate change on land-water budgets (e.g. irrigation abstraction, construction of dams and reservoirs, trends in precipitation, and land-use change). Our findings on ΔTWS and ΔGWS apply specifically to the period (2002-2016) observed by GRACE. We note from the recommended review by Wada et al. (2017) that recent groundwater-storage depletion has made a net positive contribution to global sea-level rise. Further, as highlighted by R1, Reager et al. (2016) apply GRACE data from 2002 to 2014 to show a trend towards enhanced precipitation on the land under climate change. Given this trend and the observed intensification of precipitation on land under climate change (Allan et al., 2010; Westra et al., 2013; Myhre et al., 2019; Zhang et al., 2013), we may expect that groundwater recharge to many large-scale aquifer systems will increase under climate change in light of the statistical relationships found in this study between ΔGWS and extreme precipitation. We propose to update the discussion on this specific point in the revised manuscript.

REVISION R1 #3b: see revised text on page 20 and lines 470-477.

4. *Line 146: I don't understand the 20 realisations. I would think: 3 GRACE products, 4 LSM estimates of ΔSMS and ΔSWS and one LSM with ΔSNS amount to 3x4x1 = 12 realisations? Or did you combine e.g. ΔSWS from one LSM with the ΔSMS from another? If you did this, this seems to be inconsistent as it would not preserve mass and overestimate the errors due to the LSM corrections.*

R1 #4. We thank Reviewer 1 for this query and will revise our explanation from where the 20 realisations derive. On lines 142-145 of the original manuscript, we write, "we apply 3 gridded GRACE products (CSR, JPL-Mascons, GRGS) and an ensemble mean of ΔTWS and individual storage component of ΔSMS and ΔSWS from 4 Land Surface Models (LSMs: CLM, Noah, VIC, Mosaic), and a single ΔSNS from Noah model (GLDAS version 2.1)." The breakdown of 20 realisations is given below with 12 realisations being the primary products, whereas the remaining 8 realisations derive from a combination of GRACE ΔTWS and different LSMs to demonstrate the range of uncertainty in the estimation of ΔGWS using GRACE-derived ΔTWS and GLDAS LSMs.

- 3 GRACE TWS x 4 LSMs (SWS, SMS) x 1 LSM (SNS) = 12 realisations
- CSR GRACE TWS x mean LSMs (SWS, SMS) x 1 LSM (SNS) = 1 realisation
- JPL GRACE TWS x mean LSMs (SWS, SMS) x 1 LSM (SNS) = 1 realisation
- GRGS GRACE TWS x mean LSMs (SWS, SMS) x 1 LSM (SNS) = 1 realisation
- Mean GRACE TWS x 4 LSMs (SWS, SMS) x 1 LSM (SNS) = 4 realisation
- Mean GRACE TWS x mean LSMs (SWS, SMS) x 1 LSM (SNS) = 1 realisation

REVISION R1 #4: see the revised text on page 7 and lines 147 and a new supplementary Table S5.

Small remarks:

- *The first sentence of the introduction: Doell et al (2012) is only one model-based study providing these numbers. I would advise using less significant numbers based on an overview of estimates by Hanasaki et al (2018): https://www.hydrol-earthsyst-sci.net/22/789/2018/*

R1 #5. We appreciate Reviewer 1's suggestion and will incorporate this recommended study by Hanasaki et al. (2018) in the revised manuscript.

REVISION R1 #5: see revised text on page 2 and lines 26-28.

- *Lines 212-215: Trying out different smoothing parameters. I feel that the results of this exercise should be shown, at least in the Supplementary Information (SI). Perhaps report the statistics of the residuals for a number of settings of the smoothing parameters to justify the values chosen.*

R1 #6. The original analysis evaluated the effect of seasonal and trend smoothing windows in applying the STL (Seasonal-Trend Decomposition using Loess) decomposition method. In the revised manuscript, we will expand discussion in the Methodology of the sensitivity analysis that applied smoothing windows at various lengths; we will also include a new figure to the supplementary information in addition to the current STL figure in Fig. S37.

REVISION R1 #6: see revised text on page 9 and lines 213-217.

- *On a related note: Looking at some of the plots in the SI I see that residuals are often far from white. In the time series literature this would be seen as a serious model insufficiency. Some discussion on how this would affect results is warranted as well.*

R1 #7. We thank Reviewer 1 for spotting the fact that some lines (e.g. uncertainty envelop around the mean of time-series records) are cut-off by figure margins. In the revised supplementary information, we will reproduce all the time-series plots (from Figs S1 to S36) with a full range of values.

REVISION R1 #7: see the reconstructed figures in supplementary Figs. S1-36.

Interactive comment on "*Groundwater storage dynamics in the world's large aquifer systems from GRACE: uncertainty and role of extreme precipitation*" by Mohammad Shamsudduha and Richard G. Taylor

Soumendra Bhanja (Referee #2)

soumendrabhanja@gmail.com
Received and published: 13 February 2020

Reviewer's comments are italicised, and our responses (R) and revisions (REVISION) are provided in normal fonts.

General comments:

The authors present a manuscript on GRACE-based terrestrial and groundwater storage changes in 37 major aquifer systems across the globe. I must say, the authors have done a commendable job to compare huge amount of data in all of those major global aquifers.

My major comments are provided below:

1. *The surface water storage was used from GLDAS estimates of surface runoff. How do the authors comment on surface water storage variations in natural structures such as, rivers, lakes and artificial structures like dams? I believe, the influence of surface water storage in natural and artificial structures can provide erroneous disaggregation of TWS. The smaller fraction of surface water storage is clearly visible in the figures on comparing the soil moisture and groundwater storages.*

R2 #1. We thank Reviewer 2 (R2, Dr. Bhanja) for his positive comments and providing valuable suggestions to improve the manuscript. As with R1, we agree with R2 that the representation of surface water storage (ΔSWS) by NASA's GLDAS (Global Land Data Assimilation System) models (i.e. CLM, Noah, VIC, Mosaic) is uncertain and limited. As with a recent study (Getirana et al., 2017), we note that the vast majority of studies estimating GRACE-derived ΔGWS disregard ΔSWS, assuming its contribution to ΔTWS variability is small (e.g. Long et al., 2016) in contrast to other studies (e.g. Shamsudduha et al., 2012) that show that contribution to ΔTWS from observed ΔSWS is substantial (22%). However, more precise estimates of the impacts of ΔSWS on ΔTWS globally and its spatial variability are unknown due to lack of global-scale observations of surface water storage including anthropogenic structures such as dams and reservoirs. Similar to the approaches taken in recent studies (e.g. Bhanja et al., 2016; Thomas et al., 2017), we apply surface runoff as a reasonable proxy for ΔSWS derived from GLDAS LSMs. We recognise differences exist in surface runoff estimates simulated by GLDAS LSMs as noted previously in inter-comparison studies (e.g. Scanlon et al., 2018; Scanlon et al., 2019; Zaitchik et al., 2010). We contend, however, that dismissing ΔSWS entirely in the estimation of ΔGWS from GRACE does not serve to reduce uncertainty.

REVISION R2 #1: see revised text on page 16 and lines 378-384.

2. *Lines 304-316: I am not totally agreeing with the arguments provided by the authors. They have not provided sufficient factual evidences in support of these arguments. There can be multiple reasons behind that. Surface water storage in the reservoirs can also play crucial role here, which is not considered in this study.*

R2 #2. On lines 304 to 316 in the original manuscript, we argue that abrupt rises or falls in calculated ΔGWS can result from simple arithmetic operations of numeric values, given the uncertainty that exists in estimates of water storage from GLDAS LSMs expressed as anomalies. We agree with R2 that uncertainty in the estimation of ΔSWS may be one of the causes of these abrupt rises or falls. Because ΔGWS is calculated from GRACE-derived ΔTWS by subtracting storage anomalies from other terrestrial water components such as soil moisture storage (ΔSMS), surface water storage (ΔSWS) and snow water storage (ΔSNS), 'mathematical artefacts' in the calculation of groundwater storage change (ΔGWS) from GRACE and GLDAS datasets can occur where over/underestimation in individual components can collectively lead to abrupt falls/rises relative to GRACE ΔTWS as per equation 1 in the original manuscript.

REVISION R2 #2: see revised text on page 16-17 and lines 386-390.

3. *Please include a limitation section mentioning all the limitations in this study. For soil moisture storage, one of the major limitations is that the simulated values are up to 3.4 m at max, soil moisture at deeper layers are not used. This is particularly important in arid, semi-arid regions where vadose zone thickness is much deeper. "Uncertainty is generally higher for aquifers systems located in arid to hyper-arid environments (Table 2, see supplementary Fig. S79)." This observation can be linked with the non-representation of deeper soil moisture.*

R2 #3. We agree with R2 that substantial variability and uncertainty exists in simulated soil moisture storage by GLDAS land surface models. R2 is correct in noting that the number of layers and depth of soil horizons in the four LSMs differ with a maximum depth of 3.5 m in Mosaic. We agree that the depth to the deepest part of the unsaturated zone and soils in semi-arid and arid environments could be well below the maximum depth of soil layers considered in these models. For example, the thickness of unsaturated zone in the Southern High Plains in the US can range from 0 to 134 m with a median thickness of 37 m (Scanlon et al., 2009). We will nevertheless expand discussion of uncertainty in the representation of ΔSMS by GLDAS LSMs making specific reference to their consequences for soils in semi-arid and arid environments.

REVISION R2 #3: see revised text on page 16 and lines 378-384.

4. *Sections 3.5, 4.2 and elsewhere: In general, while describing extreme precipitation, researchers normally use precipitation per day time-scale. The authors seem not to use the daily precipitation data. Please change the discussion topic to mention precipitation only.*

R2 #4. We agree with R2 that precipitation intensity is most commonly discussed in relation to daily precipitation but this is not exclusive. Statistically, extreme precipitation can also refer to annual, seasonal, and monthly precipitation. In the manuscript, we define statistically extreme precipitation (i.e. 90^{th} percentiles) on a monthly timestep over the period of 1901 to 2016 (116 years) consistent with the monthly timestep in the employed GRACE and GLDAS time-series datasets.

REVISION R2 #4: in the revised text, we added the word "monthly" on page 15 and line 341.

5. *Section 3.5: Observing non-significant, low correlation between precipitation and groundwater may indicate human interference. Central valley (16) is a clear exception here. This shows correlation analysis is not properly reflecting the observation.*

R2 #5. We agree with the R2 that the low correlation between precipitation and groundwater storage may indicate human interference (i.e. groundwater pumping) masking natural variability that may be caused by climate (i.e. precipitation variations). In the revised manuscript, we will expand discussion of potential factors overprinting natural variabilities in ΔGWS such as groundwater abstraction.

REVISION R2 #5: see revised text on page 19-20 and lines 458-466.

6. *Figure 3: Show the scale of variance.*

R2 #6. The variance in GRACE-derived TWS time-series records for all 37 large aquifer systems are presented in the Supplementary Table S1.

REVISION R2 #6: see Supplementary Table S1 for the variance in GRACE-TWS data. We also refer to page 10 and line 232.

7. *"For example, centennial-scale piezometry in the Ganges-Brahmaputra aquifer system (no. 24) reveals that recent groundwater depletion in NW India traced by GRACE (Fig. 5 and supplementary Fig. S23) follows more than a century of groundwater accumulation through leakage of surface water via a canal network constructed primarily during the 19th century (MacDonald et al., 2016)." Centennial-scale data are not present in the manuscript. Please include them in SI. This is not only from the recharge of canal irrigation, groundwater resources in this area got benefited also from a significant rate of annual rainfall. The present decline is clearly linked to irrigation-linked withdrawal. Please mention these.*

R2 #7. We thank R2 for their comments and suggestions regarding centennial-scale changes in groundwater storage in the Ganges-Brahmaputra Basin that contrast with short-term (2002-2016) declining trends in ΔGWS revealed by GRACE. We agree that recent declining trends result from intensive groundwater-fed irrigation in the region. The centennial-scale groundwater levels (1900 to 2010) are clearly shown in Figure 3b of the *Nature Geoscience* letter by MacDonald et al. (2016) and we will reproduce this figure, as proposed, in the revised Supplementary Information.

REVISION R2 #7: see a new figure in Supplementary Fig. S80. Reference to the figure is provided in the revised text on page 17 and line 400.

8. *Figure 8: Continuous rise in GWS observed in several basins including Amazon, where precipitation rates even show declining trends (Figure S18). Please discuss the probable reasons.*

R2 #8. We similarly note this interesting observation made by R2. Rising trends in GRACE-derived ΔGWS time series follows a similar trend in ΔTWS over the Amazon Basin that has been reported by Scanlon et al. (2018) at 41 to 44 km^3/year (period 2002 to 2014) and Rodell et al. (2018) at 51.9 ± 9.4 Gt/year (period 2002 to 2016). The magnitude of this rising trend in ΔTWS is explained by both the size of the region and the intensity of the Amazon water cycle (Chaudhari et al., 2019). Furthermore, Rodell et al. (2018) argue that large dam construction in southern Brazil and filling of reservoirs contributed to the rising trend in GRACE ΔTWS. We note that a slightly decreasing trend in soil moisture storage (ΔSMS) might be contributing to a positive change in ΔGWS over the Amazon Basin.

REVISION R2 #8: see revised text on page 19 and lines 444-449.

9. *Line 137: surface runoff or surface water storage (SNS).*

R2 #9. Thanks to R2 for pointing out this typo. We agree that it should read SWS, not SNS and will be corrected in the revised manuscript.

REVISION R2 #9: see revised text on page 6 and line 138.

References:

Allan, R. P., Soden, B. J., John, V. O., Ingram, W., and Good, P.: Current changes in tropical precipitation, Environmental Research Letters, 5, 025205, 10.1088/1748-9326/5/2/025205, 2010.

Bhanja, S. N., Mukherjee, A., Saha, D., Velicogna, I., and Famiglietti, J. S.: Validation of GRACE based groundwater storage anomaly using in-situ groundwater level measurements in India, Journal of Hydrology, 543, 729-738, 2016.

Bierkens, M. F. P., and Wada, Y.: Non-renewable groundwater use and groundwater depletion: a review, Environ. Res. Lett., 14, 063002, 10.1088/1748-9326/ab1a5f, 2019.

Chaudhari, S., Pokhrel, Y., Moran, E., and Miguez-Macho, G.: Multi-decadal hydrologic change and variability in the Amazon River basin: understanding terrestrial water storage variations and drought characteristics, Hydrol. Earth Syst. Sci., 23, 2841–2862, 10.5194/hess-23-2841-2019, 2019.

de Graaf, I. E. M., van Beek, L. P. H., Gleeson, T., Moosdorf, N., Schmitz, O., Sutanudjaja, E. H., and Bierkens, M. F. P.: A global-scale two layer transient groundwater model: development and application to groundwater depletion, Adv. Water Resour., 102, 53–67, 2017.

Getirana, A., Kumar, S., Girotto, M., and Rodell, M.: Rivers and Floodplains as Key Components of Global Terrestrial Water Storage Variability, Geophysical Research Letters, 44, 10359-10368, 2017.

Hanasaki, N., Yoshikawa, S., Pokhrel, Y., and Kanae, S.: A global hydrological simulation to specify the sources of water used by humans, Hydrol. Earth Syst. Sci., 22, 789–817, 2018.

Long, D., Chen, X., Scanlon, B. R., Wada, Y., Hong, Y., Singh, V. P., Chen, Y., Wang, C., Han, Z., and Yang, W.: Have GRACE satellites overestimated groundwater depletion in the Northwest India Aquifer?, Scientific Reports, 6, 24398, doi:10.1038/srep24398, 2016.

MacDonald, A. M., Bonsor, H. C., Ahmed, K. M., Burgess, W. G., Basharat, M., Calow, R. C., Dixit, A., Foster, S. S. D., Gopal, K., Lapworth, D. J., Lark, R. M., Moench, M., Mukherjee, A., Rao, M. S., Shamsudduha, M., Smith, L., Taylor, R. G., Tucker, J., van Steenbergen, F., and Yadav, S. K.: Groundwater quality and depletion in the Indo-Gangetic Basin mapped from in situ observations, Nature Geoscience, 9, 762-766, 2016.

Myhre, G., Alterskjær, K., Stjern, C. W., Hodnebrog, Ø., Marelle, L., Samset, B. H., Sillmann, J., Schaller, N., Fischer, E., Schulz, M., and Stohl, A.: Frequency of extreme precipitation increases extensively with event rareness under global warming, Scientific Reports, 9, 16063, 10.1038/s41598-019-52277-4, 2019.

Reager, J. T., Gardner, A. S., Famiglietti, J. S., Wiese, D. N., Eicker, A., and Lo, M.-H.: A decade of sea level rise slowed by climate-driven hydrology, Science, 351, 699-703, 2016.

Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., and Lo, M. H.: Emerging trends in global freshwater availability, Nature, 557, 651-659, 10.1038/s41586-018-0123-1, 2018.

Scanlon, B. R., Stonestrom, D. A., Reedy, R. C., Leaney, F. W., Gates, J., and Cresswell, R. G.: Inventories and mobilization of unsaturated zone sulfate, fluoride, and chloride related to land use change in semiarid regions, southwestern United States and Australia, Water Resour. Res., 45, W00A18, 10.1029/2008WR006963, 2009.

Scanlon, B. R., Zhang, Z., Save, H., Sun, A. Y., Müller Schmied, H., van Beek, L. P. H., Wiese, D. N., Wada, Y., Long, D., Reedy, R. C., Longuevergne, L., Döll, P., and Bierkens, M. F. P.: Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data, PNAS, 115 1080-1089, doi:10.1073/pnas.1704665115, 2018.

Scanlon, B. R., Zhang, Z., Rateb, A., Sun, A., Wiese, D., Save, H., Beaudoing, H., Lo, M. H., Müller-Schmied, H., Döll, P., van Beek, R., Swenson, S., Lawrence, D., Croteau, M., and Reedy, R. C.: Tracking seasonal fluctuations in land water storage using global models and GRACE satellites, Geophysical Research Letters, 46, 5254–5264, 10.1029/2018GL081836, 2019.

Shamsudduha, M., Taylor, R. G., and Longuevergne, L.: Monitoring groundwater storage changes in the highly seasonal humid tropics: validation of GRACE measurements in the Bengal Basin, Water Resour. Res., 48, W02508, doi:10.1029/2011WR010993, 2012.

Thomas, B. F., Caineta, J., and Nanteza, J.: Global assessment of groundwater sustainability based on storage anomalies, Geophysical Research Letters, 44, 11445-11455, doi:10.1002/2017GL076005, 2017.

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

Wada, Y., Reager, J. T., Chao, B. F., Wang, J., Lo, M.-H., Song, C., Li, Y., and Gardner, A. S.: Recent Changes in Land Water Storage and its Contribution to Sea Level Variations, Surv Geophys, 38, 131–152, 10.1007/s10712-016-9399-6, 2017.

Westra, S., Alexander, L. V., and Zwiers, F. W.: Global Increasing Trends in Annual Maximum Daily Precipitation, J. Climate, 26, 3904–3918, 10.1175/JCLI-D-12-00502.1, 2013.

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, 10.1029/2009WR007811, 2010.

Zhang, X., Wan, H., Zwiers, F. W., Hegerl, G. C., and Min, S.-K.: Attributing intensification of precipitation extremes to human influence, Geophysical Research Letters, 40, 5252–5257, 10.1002/grl.51010, 2013.

Groundwater storage dynamics in the world's large aquifer systems from GRACE: uncertainty and role of extreme precipitation

Mohammad Shamsudduha[1,2,*] and Richard G. Taylor[1]

[1] Department of Geography, University College London, London, UK

[2] Department of Geography, University of Sussex, Falmer, Brighton, UK

[*] Corresponding author: M. Shamsudduha (M.Shamsudduha@sussex.ac.uk)

Abstract

Under variable and changing climates groundwater storage sustains vital ecosystems and enables freshwater withdrawals globally for agriculture, drinking-water, and industry. Here, we assess recent changes in groundwater storage (ΔGWS) from 2002 to 2016 in 37 of the world's large aquifer systems using an ensemble of datasets from the Gravity Recovery and Climate Experiment (GRACE) and Land Surface Models (LSMs). Ensemble GRACE-derived ΔGWS is well reconciled to in-situ observations (r = 0.62–0.86, p value <0.001) for two tropical basins with regional piezometric networks and contrasting climate regimes. Trends in GRACE-derived ΔGWS are overwhelmingly non-linear; indeed, linear declining trends adequately (R^2 >0.5, p value <0.001) explain variability in only two aquifer systems. Non-linearity in ΔGWS derives, in part, from the episodic nature of groundwater replenishment associated with extreme annual (>90th percentile, 1901–2016) precipitation and is inconsistent with prevailing narratives of global-scale groundwater depletion at the scale of GRACE footprint (~200,000 km^2). Substantial uncertainty remains in estimates of GRACE-derived ΔGWS, evident from 20 realisations presented here, but these data provide a regional context to changes in groundwater storage observed more locally through piezometry.

Commented [RT1]: Addresses R1#1 and R1#3a.

1 Introduction

Groundwater is estimated to account for between a quarter and a third of the world's annual freshwater withdrawals to meet agricultural, industrial and domestic demand (Döll et al., 2012; Wada et al., 2014; Hanasaki et al., 2018). As the world's largest distributed store of freshwater, groundwater plays a vital role in sustaining ecosystems and enabling adaptation to increased variability in rainfall and river discharge brought about by climate change (Taylor et al., 2013a). Sustained reductions in the volume of groundwater (i.e. groundwater depletion) resulting from human withdrawals or changes in climate have historically been observed as declining groundwater levels recorded in wells (Scanlon et al., 2012a; Castellazzi et al., 2016; MacDonald et al., 2016). The limited distribution and duration of piezometric records hinder, however, direct observation of changes in groundwater storage globally including many of the world's large aquifer systems (WHYMAP and Margat, 2008).

Commented [MS2]: The revised text addresses R1#5.

Since 2002 the Gravity Recovery and Climate Experiment (GRACE) has enabled large-scale ($\geq$ 200,000 km^2) satellite monitoring of changes in total terrestrial water storage (ΔTWS) globally (Tapley et al., 2004). As the twin GRACE satellites circle the globe ~15 times a day they measure the inter-satellite distance at a minute precision (within one micron) and provide ΔTWS for the entire earth approximately every 30 days. GRACE satellites sense movement of total terrestrial water mass derived from both natural (e.g. droughts) and anthropogenic (e.g. irrigation) influences globally (Rodell et al., 2018). Changes in groundwater storage (GRACE-derived ΔGWS) are computed from ΔTWS after deducting contributions (equation 1) that arise from other terrestrial water stores including soil moisture storage (ΔSMS), surface water storage (ΔSWS), and the snow water storage (ΔSNS) using data from Land Surface Models (LSMs) either exclusively (Rodell et al., 2009; Famiglietti et al., 2011; Scanlon et al., 2012a; Famiglietti and Rodell, 2013; Richey et al., 2015; Thomas et al., 2017) or in combination with in situ observations (Rodell et al., 2007; Swenson et al., 2008; Shamsudduha et al., 2012).

$$\Delta\text{GWS} = \Delta\text{TWS} - (\Delta\text{SMS} + \Delta\text{SWS} + \Delta\text{SNS}) \quad (1)$$

Substantial uncertainty persists in the quantification of changes in terrestrial water stores from GRACE measurements that are limited in duration (2002 to 2016), and the application of uncalibrated, global-scale LSMs (Shamsudduha et al., 2012; Döll et al., 2014; Scanlon et al., 2018). Computation of ΔGWS from GRACE ΔTWS is argued, nevertheless, to provide evaluations of large-scale changes in groundwater storage where regional-scale piezometric networks do not currently exist (Famiglietti, 2014).

Previous assessments of changes in groundwater storage using GRACE in the world's 37 large aquifer systems (Richey et al., 2015; Thomas et al., 2017) (Fig. 1, Table 1) have raised concerns about the sustainability of human use of groundwater resources. One analysis (Richey et al., 2015) employed a single GRACE ΔTWS product (CSR) in which changes in subsurface storage (ΔSMS + ΔGWS) were attributed to ΔGWS. This study applied linear trends without regard to their significance to compute values of GRACE-derived ΔGWS over 11 years from 2003 to 2013, and concluded that the majority of the world's aquifer systems (n = 21) are either "overstressed" or "variably stressed". A subsequent analysis (Thomas et al., 2017) employed a different GRACE ΔTWS product (Mascons) and estimated ΔSWS from LSM data for both surface and subsurface runoff, though the latter is normally considered to be groundwater recharge (Rodell et al., 2004). Using performance metrics normally applied to surface water systems including dams, this latter analysis classified nearly a third (n = 11) of the world's aquifer systems as having their lowest sustainability criterion.

Here, we update and extend the analysis of ΔGWS in the world's 37 large aquifer systems using an ensemble of three GRACE ΔTWS products (CSR, Mascons, GRGS) over a 14-year period from August 2002 to July 2016. To isolate GRACE-derived ΔGWS from GRACE ΔTWS, we employ estimates of ΔSMS, ΔSWS and ΔSNS from five LSMs (CLM, Noah, VIC, Mosaic, Noah v.2.1) run by NASA's Global Land Data Assimilation System (GLDAS). As such, we explicitly account for the contribution of ΔSWS to ΔTWS, which has been commonly overlooked (Rodell et al., 2009; Richey et al., 2015; Bhanja et al., 2016) despite evidence of its significant contribution to ΔTWS (Kim et al., 2009; Shamsudduha et al., 2012; Getirana et al., 2017). Further, we characterise trends in time-series records of GRACE-derived ΔGWS by employing a non-parametric, Seasonal-Trend decomposition procedure based on Loess (STL) (Cleveland et al., 1990) that allows for resolution of seasonal, trend and irregular components of GRACE-derived ΔGWS for each large aquifer system. In contrast to linear or multiple-linear regression-based techniques, STL assumes neither that data are normally distributed nor that the underlying trend is linear (Shamsudduha et al., 2009; Humphrey et al., 2016; Sun et al., 2017).

2 Data and Methods

2.1 Global large aquifer systems

We use the World-wide Hydrogeological Mapping and Assessment Programme (WHYMAP) Geographic Information System (GIS) dataset for the delineation of world's 37 Large Aquifer Systems (Fig. 1, Table1) (WHYMAP and Margat, 2008). The WHYMAP network, led by the German Federal Institute for Geosciences and Natural Resources (BGR), serves as a central repository and hub for global groundwater data, information, and mapping with a goal of assisting regional, national, and international efforts toward sustainable groundwater management (Richts et al., 2011). The largest aquifer system in this dataset (Supplementary Table S1) is the East European Aquifer System (WHYMAP no. 33; area: 2.9 million km^2) and the smallest one the California Central Valley Aquifer System (WHYMAP no. 16; area: 71,430 km^2), which is smaller than the typical sensing area of GRACE (~200,000 km^2). However, Longuevergne et al. (2013) argue that GRACE satellites are sensitive to total mass changes at a basin scale so ΔTWS measurements can be applied to smaller basins if the magnitude of temporal mass changes is substantial due to mass water withdrawals (e.g., intensive groundwater-fed irrigation). Mean and median sizes of these large aquifers are ~945,000 km^2 and ~600,000 km^2, respectively.

2.2 GRACE products

We use post-processed, gridded (1° × 1°) monthly GRACE TWS data from CSR land (Landerer and Swenson, 2012) and JPL Global Mascon (Watkins et al., 2015; Wiese et al., 2016) solutions from NASA's dissemination site (http://grace.jpl.nasa.gov/data), and a third GRGS GRACE solution (CNES/GRGS release RL03-v1) (Biancale et al., 2006) from the French Government space agency, Centre National D'études Spatiales (CNES). To address the uncertainty associated with different GRACE processing strategies (CSR, JPL-Mascons, GRGS), we apply an ensemble mean of the three GRACE solutions (Bonsor et al., 2018).

CSR land solution (version RL05.DSTvSCS1409) is post-processed from spherical harmonics released by the Centre for Space Research (CSR) at the University of Texas at Austin. CSR gridded datasets are available at a monthly timestep and a spatial resolution of 1° × 1° (~111 km at equator) though the actual spatial resolution of GRACE footprint (Scanlon et al., 2012a) is 450 km × 450 km or ~200,000 km^2. To amplify TWS signals we apply the dimensionless scaling factors provided as 1° × 1° bins that are derived from minimising differences between TWS estimated from GRACE and the hydrological fields from the Community Land Model (CLM4.0) (Landerer and Swenson, 2012). JPL-Mascons (version RL05M_1.MSCNv01) data processing involves the same glacial isostatic adjustment correction but applies no spatial filtering as JPL-RL05M directly relates inter-satellite range-rate data to mass concentration blocks (Mascons) to estimate monthly gravity fields in terms of equal area 3° × 3° mass concentration functions in order to minimise measurement errors. Gridded mascon fields are provided at a spatial sampling of 0.5° in both latitude and longitude (~56 km at the equator). Similar to CSR product, dimensionless scaling factors are provided as 0.5° × 0.5° bins (Shamsudduha et al., 2017) to apply to the JPL-Mascons product that also derive from the Community Land Model (CLM4.0) (Wiese et al., 2016). The scaling factors are multiplicative coefficients that minimize the difference between the smoothed and unfiltered monthly ΔTWS variations from the CLM4.0 hydrology model (Wiese et al., 2016). Finally, GRGS GRACE (version RL03-v1) monthly gridded solutions of a spatial resolution of 1° × 1° are extracted and aggregated time-series data are generated for each aquifer system. A description of the estimation method of ΔGWS from GRACE and in-situ observations is provided below.

2.3 Estimation of ΔGWS from GRACE

We apply monthly measurements of terrestrial water storage anomalies (ΔTWS) from Gravity Recovery and Climate Experiment (GRACE) satellites, and simulated records of soil moisture storage (ΔSMS), surface runoff or surface water storage (ΔSWS) and snow water equivalent (ΔSNS) from NASA's Global Land Data Assimilation System (GLDAS version 1.0) at 1° × 1° grids for the period of August 2002 to July 2016 to estimate (equation 1) groundwater storage changes (ΔGWS) in the 37 WHYMAP large aquifer systems. This approach is consistent with previous global (Thomas et al., 2017) and basin-scale (Rodell et al., 2009; Asoka et al., 2017; Feng et al., 2018) analyses of ΔGWS from GRACE. We apply 3 gridded GRACE products (CSR, JPL-Mascons, GRGS) and an ensemble mean of ΔTWS and

Commented [MS3]: This addresses R2#9

individual storage component of ΔSMS and ΔSWS from 4 Land Surface Models (LSMs: CLM, Noah, VIC, Mosaic), and a single ΔSNS from Noah model (GLDAS version 2.1) to derive a total of 20 realisations of ΔGWS (Table S5) for each of the 37 aquifer systems. We then averaged all the GRACE-derived ΔGWS estimates to generate an ensemble mean ΔGWS time-series record for each aquifer system. GRACE and GLDAS LSMs derived datasets are processed and analysed in R programming language (R Core Team, 2017).

Commented [RT4]: This new Table S5 addresses R1#4.

2.4 GLDAS Land Surface Models

To estimate GRACE-derived ΔGWS using equation (1), we use simulated soil moisture storage (ΔSMS), surface runoff, as a proxy for surface water storage ΔSWS (Getirana et al., 2017; Thomas et al., 2017), and snow water equivalent (ΔSNS) from NASA's Global Land Data Assimilation System (GLDAS). GLDAS system (https://ldas.gsfc.nasa.gov/gldas/) drives multiple, offline (not coupled to the atmosphere) Land Surface Models globally (Rodell et al., 2004), at variable grid resolutions (from 2.5° to 1 km), enabled by the Land Information System (LIS) (Kumar et al., 2006). Currently, GLDAS (version 1) drives four land surface models (LSMs): Mosaic, Noah, the Community Land Model (CLM), and the Variable Infiltration Capacity (VIC). We apply monthly ΔSMS (sum of all soil profiles) and ΔSWS data at a spatial resolution of 1° × 1° from 4 GLDAS LSMs: the Community Land Model (CLM, version 2.0) (Dai et al., 2003), Noah (version 2.7.1) (Ek et al., 2003), the Variable Infiltration Capacity (VIC) model (version 1.0) (Liang et al., 2003), and Mosaic (version 1.0) (Koster and Suarez, 1992). The respective total depths of modelled soil profiles are 3.4 m, 2.0 m, 1.9 m and 3.5 m in CLM (10 vertical layers), Noah (4 vertical layers), VIC (3 vertical layers), and Mosaic (3 vertical layers) (Rodell et al., 2004). For snow water equivalent (ΔSNS), we use simulated data from Noah (v.2.1) model (GLDAS version 2.1) that is forced by the global meteorological data set from Princeton University (Sheffield et al., 2006); LSMs under GLDAS (version 1) are forced by the CPC Merged Analysis of Precipitation (CMAP) data (Rodell et al., 2004).

2.5 Global precipitation datasets

To evaluate the relationships between precipitation and GRACE-derived ΔGWS, we use a high-resolution (0.5 degree) gridded, global precipitation dataset (version 4.01) (Harris et al., 2014) available from the Climatic Research Unit (CRU) at the University of East Anglia (https://crudata.uea.ac.uk/cru/data/hrg/). In light of uncertainty in observed precipitation datasets globally, we test the robustness of relationship between precipitation and groundwater storage using the GPCC (Global Precipitation Climatology Centre) precipitation dataset (Schneider et al., 2017) (https://www.esrl.noaa.gov/psd/data/gridded/data.gpcc.html) from 1901 to 2016. Time-series (January 1901 to July 2016) of monthly precipitation from CRU and GPCC datasets for the WHYMAP aquifer systems were analysed and processed in R programming language (R Core Team, 2017).

2.6 Seasonal-Trend Decomposition (STL) of GRACE ΔGWS

Monthly time-series records (Aug 2002 to Jul 2016; supplementary Figs. S1-S36) of the ensemble mean GRACE ΔTWS and GRACE-derived ΔGWS were decomposed to seasonal, trend and remainder or residual components using a non-parametric time series decomposition technique known as "Seasonal-Trend decomposition procedure based on a locally weighted regression method called LOESS (STL)" (Cleveland et al., 1990). Loess is a nonparametric method so that the fitted curve is obtained empirically without assuming the specific nature of any structure that may exist within the data (Jacoby, 2000). A key advantage of STL method is that it reveals relatively complex structures in time-series data that could easily be overlooked using traditional statistical methods such as linear regression.

STL decomposition technique has previously been used to analyse GRACE ΔTWS regionally (Hassan and Jin, 2014) and globally (Humphrey et al., 2016). GRACE-derived ΔGWS time-series records for each aquifer system were decomposed using the STL method (see equation 2) in the R programming language (R Core Team, 2017) as:

$$Y_t = T_t + S_t + R_t \tag{2}$$

where Y_t is the monthly ΔGWS at time t, T_t is the trend component; S_t is the seasonal component; and R_t is a remainder (residual or irregular) component.

The STL method consists of a series of smoothing operations with different moving window widths chosen to extract different frequencies within a time series, and can be regarded as an extension of classical methods for decomposing a series into its individual components (Chatfield, 2003). The nonparametric nature of the STL decomposition technique enables detection of nonlinear patterns in long-term trends that cannot be assessed through linear trend analyses (Shamsudduha et al., 2009). For STL decomposition, it is necessary to choose values of smoothing parameters to extract trend and seasonal components. Selection of parameters in STL decomposition is a subjective process. The choice of the seasonal smoothing parameter determines the extent to which the extracted seasonal component varies from year to year: a large value will lead to similar components in all years whereas a small value will allow the extracted component to track the observations more closely. Similar comments apply to the choice of smoothing parameter for the trend component. We experimented with several different choices of smoothing parameters (see supplementary Fig. S37) and checked the residuals (i.e. remainder component) for the overall performance of the STL decomposition model. We conducted the Shapiro-Wilk normality test on the residuals after fitting the STL smooth line with a range of trend-cycle (*t.window*) and seasonal (*s.window*) windows and compared the p values. Visualization of the results with several

Commented [MS5]: This addresses R1#6

smoothing parameters (supplementary Fig. S37) and the corresponding smaller *p* values (i.e., *p* value <0.01) of the normality test suggested that the overall structure of time series at all sites could be captured reasonably well using window widths of 13 for the seasonal component and 37 for the trend. We apply the STL decomposition with a robust fitting of the loess smoother (Cleveland et al., 1990) to ensure that the fitting of the curvilinear trend does not have an adverse effect due to extreme outliers in the time-series data (Jacoby, 2000). Finally, to make the interpretation and comparison of nonlinear trends across all 37 aquifer systems, smoothing parameters were then fixed for all subsequent STL analyses.

3 Results

3.1 Variability in ΔTWS of the large aquifer systems

Ensemble mean time series of GRACE ΔTWS for the world's 37 large aquifer systems are shown in Fig. 2 (High Plains Aquifer System, no. 17) and supplementary Figs. S1-S36 (remaining 36 aquifer systems). The STL decomposition of an ensemble GRACE ΔTWS in the High Plains Aquifer System (no. 17) decomposes the time series into seasonal, trend and residual components (see supplementary Fig. S37). Variance (square of the standard deviation) in monthly GRACE ΔTWS (Figs. 3a and 4, Supplementary Table S1) is highest (>100 cm^2) primarily under monsoonal precipitation regimes within the Inter-Tropical Convergence Zone (e.g. Upper Kalahari-Cuvelai-Zambezi-11, Amazon-19, Maranho-20, Ganges-Brahmaputra-24). The sum of individual components derived from the STL decomposition (i.e., seasonal, trend and irregular or residual) approximates the overall variance in time-series data. The majority of the variance (>50%) in ΔTWS is explained by seasonality (Fig. 3a); non-linear (curvilinear) trends represent <25% of the variance in ΔTWS with the exception of the Upper Kalahari-Cuvelai-Zambezi-11 (42%). In contrast, variance in

Commented [MS6]: This addresses R2#6

GRACE ΔTWS in most hyper-arid and arid basins is low (Fig. 3a), <10 cm^2 (e.g., Nubian-1, NW Sahara-2, Murzuk-Djado-3, Taodeni-Tanezrouft-4, Ogaden-Juba-9, Lower Kalahari-Stampriet-12, Karoo-13, Tarim-31) and largely (> 65%) attributed to ΔGWS (Supplementary Table S2). Overall, changes in ΔTWS (i.e., difference between two consecutive hydrological years) are correlated (Pearson correlation, r >0.5, p value <0.01) to annual precipitation for 25 of the 37 large aquifer systems (Table S1). GRACE ΔTWS in aquifer systems under monsoonal precipitation regimes is strongly correlated to rainfall with a lag of 2 months (r >0.65, p value <0.01).

3.2 GRACE-ΔGWS and evidence from in-situ piezometry

Evaluations of computed GRACE-derived ΔGWS using in situ observations are limited spatially and temporally by the availability of piezometric records (Swenson et al., 2006; Strassberg et al., 2009; Scanlon et al., 2012b; Shamsudduha et al., 2012; Panda and Wahr, 2015; Feng et al., 2018). Consequently, comparisons of GRACE and in situ ΔGWS remain opportunity-driven and, here, comprise the Limpopo Basin in South Africa and Bengal Basin in Bangladesh where we possess time series records of adequate duration and density. The Bengal Basin is a part of the Ganges-Brahmaputra aquifer system (aquifer no. 24) whereas the Limpopo Basin is located between the Lower Kalahari-Stampriet Basin (aquifer no. 12) and the Karoo Basin (aquifer no. 13). The two basins feature contrasting climates (i.e. tropical humid versus tropical semi-arid) and geologies (i.e. unconsolidated sands versus weathered crystalline rock) that represent key controls on the magnitude and variability expected in ΔGWS. Both basins are in the tropics and, as such, serve less well to test the computation of GRACE-derived ΔGWS at mid and high latitudes.

In the Bengal Basin, computed GRACE and in situ ΔGWS demonstrate an exceptionally strong seasonal signal associated with monsoonal recharge that is amplified by dry-season abstraction (Shamsudduha et al., 2009; Shamsudduha et al., 2012) and high storage of the regional unconsolidated sand aquifer, represented by a bulk specific yield (S_y) of 10% (Fig. S38a). Time-series of GRACE and LSMs are shown in Fig. S39. The ensemble mean time series of computed GRACE ΔGWS from three GRACE TWS solutions and five NASA GLDAS LSMs is strongly correlated (r = 0.86, p value <0.001) to in situ ΔGWS derived from a network of 236 piezometers (mean density of 1 piezometer per 610 km^2) for the period of 2003 to 2014. In the semi-arid Limpopo Basin where mean annual rainfall (469 mm for the period of 2003 to 2015) is one-fifth of that in the Bengal Basin (2,276 mm), the seasonal signal in ΔGWS, primarily in weathered crystalline rocks with a bulk S_y of 2.5%, is smaller (Fig. S38b). Time-series of GRACE and LSMs are shown in Fig. S40. Comparison of in situ ΔGWS, derived from a network of 40 piezometers (mean density of 1 piezometer per 1,175 km^2), and computed GRACE-derived ΔGWS shows broad correspondence (r = 0.62, p value <0.001) though GRACE-derived ΔGWS is 'noisier'; intra-annual variability may result from uncertainty in the representation of other terrestrial stores using LSMs that are used to compute GRACE-derived ΔGWS from GRACE ΔTWS. The magnitude of uncertainty in monthly ΔSWS, ΔSMS, and ΔSNS that are estimated by GLDAS LSMs to compute GRACE-derived ΔGWS in each large-scale aquifer system, is depicted in Fig. 2 and supplementary Figs. S1-S36. The favourable, statistically significant correlations between the computed ensemble mean GRACE-derived ΔGWS and in situ ΔGWS shown in these two contrasting basins indicate that, at large scales (~200,000 km^2), the methodology used to compute GRACE-derived ΔGWS has merit.

3.3 Trends in GRACE-ΔGWS time series

Computation of GRACE-derived ΔGWS for the 37 large-scale aquifers globally is shown in Figs. 2 and 5. Figure 2 shows the ensemble GRACE ΔTWS and GLDAS LSM datasets used to compute GRACE-derived ΔGWS for the High Plains Aquifer System in the USA (aquifer no. 17 in Fig. 1); datasets used for all other large-scale aquifer systems are given in the Supplementary Material (Figs. S1–S36). In addition to the ensemble mean, we show uncertainty in GRACE-derived ΔGWS associated with 20 realisations from GRACE products and LSMs. Monthly time-series data of ensemble GRACE-derived ΔGWS for the other 36 large-scale aquifers are plotted (absolute scale) in Fig. 5 (in black) and fitted with a Loess-based trend (in blue). For all but five large aquifer systems (e.g., Lake Chad Basin-WHYMAP no. 7, Umm Ruwaba-8, Amazon-19, West Siberian Basin-25, and East European-33), the dominant time-series component explaining variance in GRACE-derived ΔGWS is trend (Fig. 3b, and supplementary Figs. S41-S77). Trends in GRACE-derived ΔGWS are, however, overwhelmingly non-linear (curvilinear); linear trends adequately (R^2 >0.5, p value <0.05) explain variability in GRACE-derived ΔGWS in just 5 of 37 large-scale aquifer systems and of these, only two (Arabian-22, Canning-37) are declining. GRACE-derived ΔGWS for three intensively developed, large-scale aquifer systems (Supplementary Table S1: California Central Valley-16, Ganges-Brahmaputra-24, North China Plains-29) show episodic declines (Fig. 5) though, in each case, their overall trend from 2002 to 2016 is declining but non-linear (Fig. 1).

3.4 Computational uncertainty in GRACE-ΔGWS

For several large aquifer systems primarily in arid and semi-arid environments, we identify anomalously negative or positive estimates of GRACE-derived ΔGWS that deviate substantially from underlying trends (Fig. 6 and supplementary Fig. S78). For example, the semi-arid Upper Kalahari-Cuvelai-Zambezi Basin (11) features an extreme, negative anomaly in GRACE-derived ΔGWS (Fig. 6a) in 2007-08 that is the consequence of simulated values of terrestrial stores (ΔSWS + ΔSMS) by GLDAS LSMs that exceed the ensemble GRACE ΔTWS signal. Inspection of individual time-series data for this basin (Fig. S11) reveals greater consistency in the three GRACE-ΔTWS time-series data (variance of CSR: 111 cm^2; Mascons: 164 cm^2; GRGS: 169 cm^2) compared to simulated ΔSMS among the 4 GLDAS LSMs (variance of CLM: 9 cm^2; Mosaic: 90 cm^2; Noah: 98 cm^2; VIC is 110 cm^2). In the humid Congo Basin (10), positive ΔTWS values in 2006-07 but negative ΔSMS values produce anomalously high values of GRACE-derived ΔGWS (Fig. 6b, Fig. S10). In the snow-dominated, humid Angara-Lena Basin (27), a strongly positive, combined signal of ΔSNS + ΔSWS exceeding ΔTWS leads to a very negative estimation of ΔGWS when groundwater is following a rising trend (Fig. 6c, Fig. S26).

3.5 GRACE ΔGWS and extreme precipitation

Non-linear trends in GRACE-derived ΔGWS (i.e., difference in STL trend component between two consecutive years) demonstrate a significant association with precipitation anomalies from CRU dataset for each hydrological year (i.e., percent deviations from mean annual precipitation between 2002 and 2016) in semi-arid environments (Fig. 7, Pearson correlation, $r = 0.62$, $p < 0.001$). These associations over extreme hydrological years are particularly strong in a number of individual aquifer systems (Fig. 5; Supplementary Tables S3 and S4) including the Great Artesian Basin (36) ($r = 0.93$), California Central Valley (16) ($r = 0.88$), North Caucasus Basin (34) ($r = 0.65$), Umm Ruwaba Basin (8) ($r = 0.64$), and Ogalalla (High Plains) Aquifer (17) ($r = 0.64$). In arid aquifer systems, overall associations between GRACE ΔGWS and precipitation anomalies are statistically significant but moderate ($r = 0.36$, $p < 0.001$); a strong association is found only for the Canning Basin (37) ($r = 0.52$). In humid (and sub-humid) aquifer systems, no overall statistically significant association is found yet strong correlations are noted for two temperate aquifer systems (Northern Great Plains Aquifer (14), $r = 0.51$; Angara−Lena Basin (27), $r = 0.54$); weak correlations are observed in the humid tropics for the Maranhao Basin (20, $r = 0.24$) and Ganges-Brahmaputra Basin (24, $r = 0.28$).

Distinct rises observed in GRACE-derived ΔGWS correspond with extreme seasonal (annual) precipitation (Fig. 5; Table S3 and Table S4). In the semi-arid Great Artesian Basin (aquifer no. 36) (Fig. 5 and supplementary Fig. S35), two consecutive years (2009–10 and 2010–11) of statistically extreme (i.e., $>90^{th}$ percentile, period: 1901 to 2016) monthly precipitation interrupt a multi-annual (2002 to 2009) declining trend. Pronounced rises in GRACE-derived ΔGWS in response to extreme annual rainfall are visible in other semi-arid, large aquifer systems including the Umm Ruwaba Basin (8) in 2007, Lower Kalahari-Stampriet Basin (12) in 2011, California Central Valley (16) in 2005, Ogalalla (High Plains) Aquifer (17) in 2015, and Indus Basin (23) in 2010 and 2015 (Tables S3 and S4 and Figs. S2, S8, S12, S16, S22). Similar rises in GRACE-derived ΔGWS in response to extreme annual rainfall in arid basins include the Lake Chad Basin (7) in 2012 and Ogaden-Juba Basin (9) in 2013 (Table S3 and Figs. S7, S9). In the Canning Basin, a substantial rise in GRACE-derived ΔGWS occurs in 2010–11 (Tables S3 and S4 and Fig. S36) in response to extreme annual rainfall though the overall trend is declining.

Commented [MS7]: This addresses R2#4

Non-linear trends that feature substantial rises in GRACE-derived ΔGWS in response to extreme annual precipitation under humid climates, are observed in the Maranhao Basin (20) in 2008-09, Guarani Aquifer System (21) in 2015-16, and North China Plains (29) in 2003. Consecutive years of extreme precipitation in 2012 and 2013 also generate a distinct rise in GRACE-derived ΔGWS in the Song-Liao Plain (30) (Tables S3 and S4 and Figs. S29). In the heavily developed (Table S2) Ganges-Brahmaputra Basin (24), a multi-annual (2002 to 2010) declining trend is halted by an extreme (i.e., highest over the GRACE period of 2002 to 2016 but 59^{th} percentile over the period of 1901 to 2016 using CRU dataset) annual precipitation in 2011 (Tables S3 and S4 and Figs. S23). Consecutive years from 2014 to 2015 of extreme annual precipitation increase GRACE-derived ΔGWS and disrupt a multi-annual declining trend in the West Siberian Artesian Basin (25) (Tables S3 and S4 and Figs. S24). In the sub-humid Northern Great Plains (14), distinct rises in GRACE-derived ΔGWS occur in 2010 (Tables S3 and S4 and Figs. S14) in response to extreme annual precipitation though the overall trend is linear and rising. The overall agreement in mean annual precipitation between the CRU and GPCC datasets for the period of 1901 to 2016 is strong (median correlation coefficient in 37 aquifer systems, $r = 0.92$).

4 Discussion

4.1 Uncertainty in GRACE-derived ΔGWS

We compute the range of uncertainty in GRACE-derived ΔGWS associated with 20 potential realisations from applied GRACE (CSR, JPL-Mascons, GRGS) products and LSMs (CLM, Noah, VIC, Mosaic). Uncertainty is generally higher for aquifers systems located in arid to hyper-arid environments (Table 2, see supplementary Fig. S79). Computation of GRACE-derived ΔGWS relies upon uncalibrated simulations of individual terrestrial water stores (i.e., ΔSWS, ΔSWS, ΔSNS) from LSMs to estimate ΔGWS from GRACE ΔTWS. A recent global-scale comparison of ΔTWS estimated by GLDAS LSMs and GRACE (Scanlon et al., 2018) indicates that LSMs systematically underestimate water storage changes. Further, the absence of river-routing and representation of lakes and reservoirs in the estimation of ΔSWS by LSMs constrains computation of GRACE ΔGWS as similarly recognised by Scanlon et al. (2019). Finally, substantial variability in ΔSMS among GLDAS models and the limited depth (<3.5 m below ground level) to the deepest soil layer over which these LSMs simulate ΔSMS also hamper estimation of GRACE ΔGWS, especially in drylands where the thickness of unsaturated zones may an order of magnitude greater (Scanlon et al., 2009).

Commented [RT8]: Addresses R1#2 and R2#1 and R2#3

We detect probable errors in GLDAS LSM data from events that produce large deviations in GWS (Fig. 5). These errors occur because GRACE-derived ΔGWS is computed as residual (equation 1); overestimation (or underestimation) of these combined stores produces negative (or positive) values of GRACE-derived ΔGWS when the aggregated value of other terrestrial water stores is strongly positive (or negative) and no lag is assumed (Shamsudduha et al., 2017). Evidence from limited piezometric data presented here and elsewhere (Panda and Wahr, 2015; Feng et al., 2018) suggests that the dynamics in computed GRACE-derived ΔGWS are nonetheless reasonable yet the amplitude in ΔGWS from piezometry is scalable due to uncertainty in the applied S_y (Shamsudduha et al., 2012).

Assessments of ΔGWS derived from GRACE are constrained by both their limited timespan (2002–2016) and course spatial resolution (>200,000 km^2). For example, centennial-scale piezometry in the Ganges-Brahmaputra aquifer system (no. 24) reveals that recent groundwater depletion, (i.e. groundwater withdrawals that are unlikely to be replenished within a century as per Bierkens and Wada (2019)), in NW India traced by GRACE (Fig. 5 and supplementary Fig. S23) (Rodell et al., 2009; Chen et al., 2014) follows more than a century of groundwater accumulation (see supplementary Fig. S80) through leakage of surface water via a canal network constructed primarily during the 19^{th} century (MacDonald et al., 2016). Long-term piezometric records from central Tanzania and the Limpopo Basin of South Africa (Supplementary Fig. S81) show dramatic increases in ΔGWS associated with extreme seasonal rainfall events that occurred prior to 2002 and thus provide a vital context to the more recent period of ΔGWS estimated by GRACE. At regional scales, GRACE-derived ΔGWS can differ substantially from more localised, in situ observations of ΔGWS from piezometry. In the Karoo Basin (aquifer no. 13), GRACE-derived ΔGWS is also rising (Fig. 5 and supplementary Fig. S13) over periods during which groundwater depletion has been reported in parts of the basin (Rosewarne et al., 2013). In the Guarani Aquifer System (21), groundwater depletion is reported from 2005 to 2009 in Ribeiro Preto near Sao Paulo as

Commented [MS9]: This addresses R2#2

Commented [MS10]: This addresses R1#1

Commented [MS11]: This new figure addresses R2#7 a result of intensive groundwater withdrawals for urban water supplies and irrigation of sugarcane (Foster et al., 2009) yet GRACE-derived ΔGWS over this same period is rising.

4.2 Variability in GRACE ΔGWS and role of extreme precipitation

Non-linear trends in GRACE-derived ΔGWS arise, in part, from inter-annual variability in precipitation which has similarly been observed in analyses of GRACE ΔTWS (Humphrey et al., 2016; Sun et al., 2017; Bonsor et al., 2018). Annual precipitation in the Great Artesian Basin (aquifer no. 36) provides a dramatic example of how years (2009–10, 2010–11 from both CRU and GPCC datasets) of extreme precipitation can generate anomalously high groundwater recharge that arrests a multi-annual declining trend (Fig. 5), increasing variability in GRACE-derived ΔGWS over the relatively short period (15 years) of GRACE data. The disproportionate contribution of episodic, extreme rainfall to groundwater recharge has previously been shown by (Taylor et al., 2013b) from long-term piezometry in semi-arid central Tanzania where nearly 20% of the recharge observed over a 55-year period resulted from a single season of extreme rainfall, associated with the strongest El Niño event (1997–1998) of the last century (Supplementary Fig. S81a). Further analysis from multi-decadal piezometric records in drylands across tropical Africa (Cuthbert et al., 2019) confirm this bias in response to intensive precipitation.

The dependence of groundwater replenishment on extreme annual precipitation indicated by GRACE-derived ΔGWS for many of the world's large aquifer systems is consistent with evidence from other sources. In a pan-tropical comparison of stable-isotope ratios of oxygen (^{18}O:^{16}O) and hydrogen (^{2}H:^{1}H) in rainfall and groundwater, Jasechko and Taylor (2015) show that recharge is biased to intensive monthly rainfall, commonly exceeding the 70th percentile. In humid Uganda, Owor et al. (2009) demonstrate that groundwater recharge observed from piezometry is more strongly correlated to daily rainfall exceeding a threshold

(10 mm) than all daily rainfalls. Periodicity in groundwater storage indicated by both GRACE and in situ data has been associated with large-scale synoptic controls on precipitation (e.g., El Niño Southern Oscillation, Pacific Decadal Oscillation,) in southern Africa (Kolusu et al., 2019), and have been shown to amplify recharge in major US aquifers (Kuss and Gurdak, 2014) and groundwater depletion in India (Mishra et al., 2016).

In some large-scale aquifer systems, GRACE-derived ΔGWS exhibits comparatively weak correlations to precipitation. In the semi-arid Iullemmeden-Irhazer Aquifer (6) variance in rainfall over the period of GRACE observation following the multi-decadal Sahelian drought is low (Table S1) and the net rise in GRACE-derived ΔGWS is associated with changes in the terrestrial water balance resulting from land-cover change (Ibrahim et al., 2014). In the Amazon (16), rising trends in GRACE-derived ΔGWS, which are aligned to ΔTWS reported previously by Scanlon et al. (2018) and Rodell et al. (2018), occur during a period (2010–2016; see supplementary Table S18) that is the driest since the 1980s (Chaudhari et al., 2019); analyses over the longer term (1980–2015) point nevertheless to an overall intensification of the Amazonian hydrological cycle.

Commented [RT12]: Addresses R2#8.

4.3 Trends in GRACE ΔGWS under global change

Our analysis identifies non-linear trends in GRACE-derived ΔGWS for the vast majority (32 of 37) of the world's large aquifer systems (Figs. 1, 5 and 8). Non-linearity reflects, in part, the variable nature of groundwater replenishment observed at the scale of the GRACE footprint that is consistent with more localised, emerging evidence from multi-decadal piezometric records (Taylor et al., 2013b) (Supplementary Fig. S81a). The variable and often episodic nature of groundwater replenishment complicates assessments of the sustainability of groundwater withdrawals and highlights the importance of long-term observations over decadal timescales in undertaking such evaluations. Dramatic rises in freshwater withdrawals, primarily associated with the expansion of irrigated agriculture in semi-arid environments, have nevertheless led to groundwater depletion, computed globally from hydrological models (e.g., Wada et al., 2010; de Graaf et al., 2017) and volumetric-based calculations (Konikow, 2011). Further, groundwater depletion globally has been shown to contribute to sea-level rise (e.g., Wada et al., 2016). However, as recognised in a comprehensive review by Bierkens and Wada (2019), groundwater depletion is often localised, occurring below the footprint (200,000 km^2) of GRACE as has been well demonstrated by detailed modelling studies in the California Central Valley (Scanlon et al., 2012a) and North China Plain (Cao et al., 2013).

Commented [RT13]: Addresses R1#1, R1#3a and R2#5. New references are cited here.

Projections of the sustainability of groundwater withdrawals under global change are complicated, in part, by uncertainty in how radiative forcing will affect large-scale, regional controls on extreme annual precipitation like El Niño Southern Oscillation (Latif and Keenlyside, 2009). Globally, Reager et al. (2016) show a trend towards enhanced precipitation on the land under climate change. Given this trend and the observed intensification of precipitation on land from global warming (Allan et al., 2010; Westra et al., 2013; Zhang et al., 2013; Myhre et al., 2019), groundwater recharge to many large-scale aquifer systems may increase under climate change as revealed by the statistical relationships found in this study between ΔGWS and extreme annual precipitation. The magnitude of this increase is, however, unlikely to offset the impact of human withdrawals in areas of intensive abstraction for irrigated agriculture as shown in NW India by Xie et al. (2020). The developed set of GRACE-derived ΔGWS time series data for the world's large aquifer systems provided here offers a consistent, additional benchmark alongside long-term piezometry to assess not only large-scale climate controls on groundwater replenishment but also opportunities to enhance groundwater storage through managed aquifer recharge.

Commented [RT14]: Addresses R1#3b. Additional references are cited here.

5 Conclusions

Changes in groundwater storage (ΔGWS) computed from GRACE satellite data continue to rely upon uncertain, uncalibrated estimates of changes in other terrestrial stores of water found in soil, surface water, and snow/ice from global-scale models. The application here of ensemble mean values of three GRACE ΔTWS processing strategies (CSR, JPL-Mascons, GRGS) and five land-surface models (GLDAS 1: CLM, Noah, VIC, Mosaic; GLDAS 2: Noah) is designed to reduce the impact of uncertainty in an individual model or GRACE product on the computation of GRACE-derived ΔGWS. We, nevertheless, identify a few instances where erroneously high or low values of GRACE-derived ΔGWS are computed; these occur primarily in arid and semi-arid environments where uncertainty in the simulation of terrestrial water balances is greatest. Over the period of GRACE observation (2002 to 2016), we show favourable comparisons between GRACE-derived ΔGWS and piezometric observations ($r = 0.62$ to 0.86) in two contrasting basins (i.e., semi-arid Limpopo Basin, tropical humid Bengal Basin) for which in situ data are available. This study thus contributes to a growing body of research and observations reconciling computed GRACE-derived ΔGWS to ground-based data.

GRACE-derived ΔGWS from 2002 to 2016 for the world's 37 large-scale aquifer systems shows substantial variability as revealed explicitly by 20 potential realisations from GRACE products and LSMs computed here; trends in ensemble mean GRACE-derived ΔGWS are overwhelmingly (87%) non-linear. Linear trends adequately explain variability in GRACE-derived ΔGWS in just 5 aquifer systems for which linear declining trends, indicative of groundwater depletion, are observed in 2 aquifer systems (Arabian, Canning); overall trends for three intensively developed, large-scale aquifer systems (California Central Valley, Ganges-Brahmaputra, North China Plains) are declining but non-linear. This non-linearity in GRACE-derived ΔGWS for the vast majority of the world's large aquifer systems is inconsistent with previous analyses at the scale of GRACE (~200,000 km^2) asserting global-scale groundwater depletion. Groundwater depletion, more commonly observed by piezometry, is experienced at scales well below the GRACE footprint and is likely to be more pervasive than suggested by the presented analysis of large-scale aquifers. Non-linearity in GRACE-derived ΔGWS arises, in part, from episodic recharge associated with extreme (>90th percentile) annual precipitation. This episodic replenishment of groundwater, combined with natural discharges that sustain ecosystem functions and human withdrawals, produces highly dynamic aquifer systems that complicate assessments of the sustainability of groundwater withdrawals from large aquifer systems. These findings highlight, however, potential opportunities for sustaining groundwater withdrawals through induced recharge from extreme precipitation and managed aquifer recharge.

Commented [RT15]: Addresses R1#1 and R1#3a.

References

Allan, R. P., Soden, B. J., John, V. O., Ingram, W., and Good, P.: Current changes in tropical precipitation, Environmental Research Letters, 5, 025205, 10.1088/1748-9326/5/2/025205, 2010.

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

Bhanja, S. N., Mukherjee, A., Saha, D., Velicogna, I., and Famiglietti, J. S.: Validation of GRACE based groundwater storage anomaly using in-situ groundwater level measurements in India, Journal of Hydrology, 543, 729-738, 2016.

Biancale, R., Lemoine, J.-M., Balmino, G., Loyer, S., Bruisma, S., Perosanz, F., Marty, J.-C., and Gégout, P.: 3 Years of Geoid Variations from GRACE and LAGEOS Data at 10-day Intervals from July 2002 to March 2005, CNES/GRGS, 2006.

Bierkens, M. F. P., and Wada, Y.: Non-renewable groundwater use and groundwater depletion: a review, Environ. Res. Lett., 14, 063002, 10.1088/1748-9326/ab1a5f, 2019.

Bonsor, H. C., Shamsudduha, M., Marchant, B. P., MacDonald, A. M., and Taylor, R. G.: Seasonal and Decadal Groundwater Changes in African Sedimentary Aquifers Estimated Using GRACE Products and LSMs, Remote Sensing, 10, 2018.

Cao, G., Zheng, C., Scanlon, B. R., Liu, J., and Li, W.: Use of flow modeling to assess sustainability of groundwaterresources in the North China Plain, Water Resour. Res., 49, 10.1029/2012WR011899, 2013.

Castellazzi, P., Martel, R., Rivera, A., Huang, J., Pavlic, G., Calderhead, A. I., Chaussard, E., Garfias, J., and Salas, J.: Groundwater depletion in Central Mexico: Use of GRACE and InSAR to support water resources management, Water Resour. Res., 52, 5985-6003, doi:10.1002/2015WR018211, 2016.

Chatfield, C.: The analysis of time series - an introduction, 6th ed., Chapman and Hall, CRC Press, Boca Raton, 2003.

Chaudhari, S., Pokhrel, Y., Moran, E., and Miguez-Macho, G.: Multi-decadal hydrologic change and variability in the Amazon River basin: understanding terrestrial water storage variations and drought characteristics, Hydrol. Earth Syst. Sci., 23, 2841–2862, 10.5194/hess-23-2841-2019, 2019.

Chen, J., Li, J., Zhang, Z., and Ni, S.: Long-term groundwater variations in Northwest India from satellite gravity measurements, Global and Planetary Change, 116, 130-138, 2014.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I.: STL: A Seasonal Trend Decomposition Procedure Based on LOESS, J. Official Statistics, 6, 3-33, 1990.

Cuthbert, M. O., Taylor, R. G., Favreau, G., Todd, M. C., Shamsudduha, M., Villholth, K. G., MacDonald, A. M., Scanlon, B. R., Kotchoni, D. O. V., Vouillamoz, J.-M., Lawson, F. M. A., Adjomayi, P. A., Kashaigili, J., Seddon, D., Sorensen, J. P. R., Ebrahim, G. Y., Owor, M., Nyenje, P. M., Nazoumou, Y., Goni, I., Ousmane, B. I., Sibanda, T., Ascott, M. J., Macdonald, D. M. J., Agyekum, W., Koussoubé, Y., Wanke, H., Kim, H., Wada, Y., Lo, M.-H., Oki, T., and Kukuric, N.: Observed controls on resilience of groundwater to climate variability in sub-Saharan Africa, Nature, 572, 230–234, https://doi.org/10.1038/s41586-019-1441-7, 2019.

Dai, Y., Zeng, X., Dickinson, R. E., Baker, I., Bonan, G. B., Bosilovich, M. G., Denning, A. S., Dirmeyer, P. A., Houser, P. R., Niu, G., Oleson, K. W., Schlosser, C. A., and Yang, Z.-L.: The common land model (CLM), Bull. Am. Meteorol. Soc., 84, 1013-1023, 2003.

de Graaf, I. E. M., van Beek, L. P. H., Gleeson, T., Moosdorf, N., Schmitz, O., Sutanudjaja, E. H., and Bierkens, M. F. P.: A global-scale two layer transient groundwater model: development and application to groundwater depletion, Adv. Water Resour., 102, 53–67, 2017.

Döll, P., Hoffmann-Dobrev, H., Portmann, F. T., Siebert, S., Eicker, A., Rodell, M., Strassberg, G., and Scanlon, B. R.: Impact of water withdrawals from groundwater and surface water on continental water storage variations, Journal of Geodynamics, 59-60, 143-156, 2012.

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

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., and Tarpley, J. D.: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, J. Geophys. Res., 108(D22), 8851, 10.1029/2002JD003296, 2003.

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

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

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

Feng, W., Shum, C. K., Zhong, M., and Pan, Y.: Groundwater Storage Changes in China from Satellite Gravity: An Overview, Remote Sensing, 10, 674, 2018.

Foster, S., Hirata, R., Vidal, A., Schmidt, G., and Garduño, H.: The Guarani Aquifer Initiative - Towards Realistic Groundwater Management in a Transboundary Context, The World Bank, Washington D.C., 28, 2009.

Getirana, A., Kumar, S., Girotto, M., and Rodell, M.: Rivers and Floodplains as Key Components of Global Terrestrial Water Storage Variability, Geophysical Research Letters, 44, 10359-10368, 2017.

Hanasaki, N., Yoshikawa, S., Pokhrel, Y., and Kanae, S.: A global hydrological simulation to specify the sources of water used by humans, Hydrol. Earth Syst. Sci., 22, 789-817, 10.5194/hess-22-789-2018, 2018.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset, International Journal of Climatology, 34, 623-642, 2014.

Hassan, A. A., and Jin, S.: Lake level change and total water discharge in East Africa Rift Valley from satellite-based observations, Global Planet Change, 117, 79-90, 10.1016/j.gloplacha.2014.03.005, 2014.

Humphrey, V., Gudmundsson, L., and Seneviratne, S. I.: Assessing Global Water Storage Variability from GRACE: Trends, Seasonal Cycle, Subseasonal Anomalies and Extremes, Surveys in Geophysics, 37, 357-395, doi:10.1007/s10712-016-9367-1, 2016.

Ibrahim, M., Favreau, G., Scanlon, B. R., Seidel, J. L., Coz, M. L., Demarty, J., and Cappelaere, B.: Long-term increase in diffuse groundwater recharge following expansion of rainfed cultivation in the Sahel, West Africa, Hydrogeology Journal, 22, 1293-1305, 2014.

Jacoby, W. G.: Loess::a nonparametric, graphical tool for depicting relationships between variables, Electoral Studies, 19, 577-613, 2000.

Jasechko, S., and Taylor, R. G.: Intensive rainfall recharges tropical groundwaters, Environmental Research Letters, 10, 124015, 2015.

Kim, H., Yeh, P. J.-F., Oki, T., and Kanae, S.: Role of rivers in the seasonal variations of terrestrial water storage over global basins, Geophys. Res. Lett., 36, L17402, doi:10.1029/2009GL039006, 2009.

Kolusu, S. R., Shamsudduha, M., Todd, M. C., Taylor, R. G., Seddon, D., Kashaigili, J. J., Girma, E., Cuthbert, M., Sorensen, J. P. R., Villholth, K. G., MacDonald, A. M., and MacLeod, D. A.: The El Niño event of 2015-16: Climate anomalies and their impact on groundwater resources in East and Southern Africa, Hydrol. Earth Syst. Sci., 23, 1751-1762, 2019.

Konikow, L. F.: Contribution of global groundwater depletion since 1900 to sea-level rise, Geophys. Res. Lett. , 38, L17401, 10.1029/2011GL048604, 2011.

Koster, R. D., and Suarez, M. J.: Modeling the land surface boundary in climate models as a composite of independent vegetation stands, J. Geophys. Res., 97, 2697-2715, 1992.

Kumar, S. V., Peters-Lidard, C. D., Tian, Y., Houser, P. R., Geiger, J., Olden, S., Lighty, L., Eastman, J. L., Doty, B., Dirmeyer, P., Adams, J., Mitchell, K., Wood, E. F., and Sheffield, J.: Land information system: An interoperable framework for high resolution land surface modeling, Environmental Modelling & Software, 21, 1402-1415, 2006.

Kuss, A. J. M., and Gurdak, J. J.: Groundwater level response in U.S. principal aquifers to ENSO, NAO, PDO, and AMO, Journal of Hydrology, 519 (Part B), 1939-1952, 2014.

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

Latif, M., and Keenlyside, N. S.: El Niño/Southern Oscillation response to global warming, Proceedings of the National Academy of Sciences, 106, 20578-20583, doi:10.1073/pnas.0710860105, 2009.

Liang, X., Xie, Z., and Huang, M.: A new parameterization for surface and groundwater interactions and its impact on water budgets with the variable infiltration capacity (VIC) land surface model, J. Geophys. Res., 108(D16), 8613, 10.1029/2002JD003090, 2003.

Longuevergne, L., Wilson, C. R., Scanlon, B. R., and Crétaux, J. F.: GRACE water storage estimates for the Middle East and other regions with significant reservoir and lake storage, Hydrol. Earth Syst. Sci., 17, 4817-4830, doi:10.5194/hess-17-4817-2013, 2013.

MacDonald, A. M., Bonsor, H. C., Ahmed, K. M., Burgess, W. G., Basharat, M., Calow, R. C., Dixit, A., Foster, S. S. D., Gopal, K., Lapworth, D. J., Lark, R. M., Moench, M., Mukherjee, A., Rao, M. S., Shamsudduha, M., Smith, L., Taylor, R. G., Tucker, J., van

Steenbergen, F., and Yadav, S. K.: Groundwater quality and depletion in the Indo-Gangetic Basin mapped from in situ observations, Nature Geoscience, 9, 762-766, 2016.

Mishra, V., Aadhar, S., Asoka, A., S. Pai, and Kumar, R.: On the frequency of the 2015 monsoon season drought in the Indo-Gangetic Plain, Geophys. Res. Lett., 43, 12102–12112, 10.1002/GL071407, 2016.

Myhre, G., Alterskjær, K., Stjern, C. W., Hodnebrog, Ø., Marelle, L., Samset, B. H., Sillmann, J., Schaller, N., Fischer, E., Schulz, M., and Stohl, A.: Frequency of extreme precipitation increases extensively with event rareness under global warming, Scientific Reports, 9, 16063, 10.1038/s41598-019-52277-4, 2019.

Owor, M., Taylor, R. G., Tindimugaya, C., and Mwesigwa, D.: Rainfall intensity and groundwater recharge: empirical evidence from the Upper Nile Basin, Environmental Research Letters, 1-6, 2009.

Panda, D. K., and Wahr, J.: Spatiotemporal evolution of water storage changes in India from the updated GRACE-derived gravity records, Water Resour. Res., 51, 135–149, doi:10.1002/2015WR017797, 2015.

Reager, J. T., Gardner, A. S., Famiglietti, J. S., Wiese, D. N., Eicker, A., and Lo, M.-H.: A decade of sea level rise slowed by climate-driven hydrology, Science, 351, 699-703, 2016.

Richey, A. S., Thomas, B. F., Lo, M.-H., Reager, J. T., Famiglietti, J. S., Voss, K., Swenson, S., and Rodell, M.: Quantifying renewable groundwater stress with GRACE, Water Resour. Res., 51, 5217-5238, doi:10.1002/2015WR017349, 2015.

Richts, A., Struckmeier, W. F., and Zaepke, M.: WHYMAP and the Groundwater Resources of the World 1:25,000,000, in: Sustaining Groundwater Resources. International Year of Planet Earth, edited by: J., J., Springer, Dordrecht, 159-173, 2011.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, Bull. Am. Meteorol. Soc., 85, 381-394, 2004.

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

Rodell, M., Velicogna, I., and Famiglietti, J. S.: Satellite-based estimates of groundwater depletion in India, Nature, 460, 999-1003, doi:10.1038/nature08238, 2009.

Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., and Lo, M. H.: Emerging trends in global freshwater availability, Nature, 557, 651-659, 10.1038/s41586-018-0123-1, 2018.

Rosewarne, P. N., Woodford, A. C., O'Brien, R., Tonder, G. V., Esterhuyse, C., Goes, M., Talma, A. S., Tredoux, G., and Visser, D.: Karoo Groundwater Atlas, Volume 2. Karoo Groundwater Expert Group (KGEG), Ground Water Division, The Geological Society of South Africa, Cape Town, South Africa, 1-35, 2013.

Scanlon, B. R., Stonestrom, D. A., Reedy, R. C., Leaney, F. W., Gates, J., and Cresswell, R. G.: Inventories and mobilization of unsaturated zone sulfate, fluoride, and chloride related to land use change in semiarid regions, southwestern United States and Australia, Water Resour. Res., 45, W00A18, 10.1029/2008WR006963, 2009.

Scanlon, B. R., Faunt, C. C., Longuevergne, L., Reedy, R. C., Alley, W. M., McGuire, V. L., and McMahon, P. B.: Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley, Proc. Natl. Acad. Sci. USA, 109, 9320-9325, 2012a.

Scanlon, B. R., Longuevergne, L., and Long, D.: Ground referencing GRACE satellite estimates of groundwater storage changes in the California Central Valley, USA, Water Resour. Res., 48, W04520, 2012b.

Scanlon, B. R., Zhang, Z., Save, H., Sun, A. Y., Müller Schmied, H., van Beek, L. P. H., Wiese, D. N., Wada, Y., Long, D., Reedy, R. C., Longuevergne, L., Döll, P., and Bierkens, M. F. P.: Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data, PNAS, 115 1080-1089, doi:10.1073/pnas.1704665115, 2018.

Scanlon, B. R., Zhang, Z., Rateb, A., Sun, A., Wiese, D., Save, H., Beaudoing, H., Lo, M. H., Müller-Schmied, H., Döll, P., van Beek, R., Swenson, S., Lawrence, D., Croteau, M., and Reedy, R. C.: Tracking seasonal fluctuations in land water storage using global models and GRACE satellites, Geophysical Research Letters, 46, 5254–5264, 10.1029/2018GL081836, 2019.

Schneider, U., Finger, P., Meyer-Christoffer, A., Rustemeier, E., Ziese, M., and Becker, A.: Evaluating the Hydrological Cycle over Land Using the Newly-Corrected Precipitation Climatology from the Global Precipitation Climatology Centre (GPCC), Atmosphere, 8, 52, 10.3390/atmos8030052, 2017.

Shamsudduha, M., Chandler, R. E., Taylor, R. G., and Ahmed, K. M.: Recent trends in groundwater levels in a highly seasonal hydrological system: the Ganges-Brahmaputra-Meghna Delta, Hydrol. Earth Syst. Sci., 13, 2373-2385, doi:10.5194/hess-13-2373-2009, 2009.

Shamsudduha, M., Taylor, R. G., and Longuevergne, L.: Monitoring groundwater storage changes in the highly seasonal humid tropics: validation of GRACE measurements in the Bengal Basin, Water Resour. Res., 48, W02508, doi:10.1029/2011WR010993, 2012.

Shamsudduha, M., Taylor, R. G., Jones, D., Longuevergne, L., Owor, M., and Tindimugaya, C.: Recent changes in terrestrial water storage in the Upper Nile Basin: an evaluation of commonly used gridded GRACE products, Hydrol. Earth Syst. Sci., 21, 4533-4549, 2017.

Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling, Journal of Climate, 19, 3088-3111, 2006.

Strassberg, G., Scanlon, B. R., and Chambers, D.: Evaluation of groundwater storage monitoring with the GRACE satellite: Case study of the High Plains aquifer, central United States, Water Resour. Res., 45, W05410, 2009.

Sun, A. Y., Scanlon, B. R., AghaKouchak, A., and Zhang, Z.: Using GRACE Satellite Gravimetry for Assessing Large-Scale Hydrologic Extremes, Remote Sensing, 9, 1287, doi:10.3390/rs9121287, 2017.

Swenson, S., Yeh, P. J.-F., Wahr, J., and Famiglietti, J. S.: A comparison of terrestrial water storage variations from GRACE with in situ measurements from Illinois, Geophys. Res. Lett., 33, L16401, doi:10.1029/2006GL026962, 2006.

Swenson, S., Famiglietti, J., Basara, J., and Wahr, J.: Estimating profile soil moisture and groundwater variations using GRACE and Oklahoma Mesonet soil moisture data, Water Resour. Res., 44, W01413, doi:10.1029/2007WR006057, 2008.

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

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

Taylor, R. G., Todd, M. C., Kongola, L., Maurice, L., Nahozya, E., Sanga, H., and MacDonald, A.: Evidence of the dependence of groundwater resources on extreme rainfall in East Africa, Nature Climate Change, 3, 374-378, doi:10.1038/nclimate1731, 2013b.

Thomas, B. F., Caineta, J., and Nanteza, J.: Global assessment of groundwater sustainability based on storage anomalies, Geophysical Research Letters, 44, 11445-11455, doi:10.1002/2017GL076005, 2017.

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

Wada, Y., Wisser, D., and Bierkens, M. F. P.: Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources, Earth Syst. Dynam., 5, 15-40, 10.5194/esd-5-15-2014, 2014.

Wada, Y., Lo, M.-H., Yeh, P. J. F., Reager, J. T., Famiglietti, J. S., Wu, R.-J., and Tseng, Y.-H.: Fate of water pumped from underground and contributions to sea-level rise, Nature Climate Change, 6, 777-780, 10.1038/nclimate3001, 2016.

Watkins, M. M., Wiese, D. N., Yuan, D.-N., Boening, C., and Landerer, F. W.: Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons, J. Geophys. Res. Solid Earth, 120, 2648–2671, doi:10.1002/2014JB011547, 2015.

Westra, S., Alexander, L. V., and Zwiers, F. W.: Global Increasing Trends in Annual Maximum Daily Precipitation, J. Climate, 26, 3904–3918, 10.1175/JCLI-D-12-00502.1, 2013.

Wiese, D. N., Landerer, F. W., and Watkins, M. M.: Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution, Water Resour. Res., 52, 7490-7502, 10.1002/2016WR019344, 2016.

Xie, H., Longuevergne, L., Ringler, C., and Scanlon, B. R.: Integrating groundwater irrigation into hydrological simulation of India: Case of improving model representation of anthropogenic water use impact using GRACE, Journal of Hydrology: Regional Studies, 29, 100681, 10.1016/j.ejrh.2020.100681, 2020.

Zhang, X., Wan, H., Zwiers, F. W., Hegerl, G. C., and Min, S.-K.: Attributing intensification of precipitation extremes to human influence, Geophysical Research Letters, 40, 5252–5257, 10.1002/grl.51010, 2013.

Acknowledgements

M.S. and R.T. acknowledge support from NERC-ESRC-DFID UPGro '*GroFutures*' (Ref. NE/M008932/1; www.grofutures.org); R.T. also acknowledges the support of a Royal Society – Leverhulme Trust Senior Fellowship (Ref. LT170004).

Data Availability

Supplementary information is available for this paper as a single PDF file. Data generated and used in this study can be made available upon request to the corresponding author.

Tables and Figures

Table 1. Identification number, name and general location of the world's 37 large aquifer systems as provided in the WHYMAP database (https://www.whymap.org/). Mean climatic condition of each of the 37 aquifer systems based on the aridity index is tabulated.

WHYMAP aquifer no.	**WHYMAP Aquifer name**	**Continent**	**Climate zones based on Aridity index**	**WHYMAP aquifer no.**	**WHYMAP Aquifer name**	**Continent**	**Climate zones based on Aridity index**
1	Nubian Sandstone Aquifer System	Africa	Hyper-arid	20	Maranhao Basin	South America	Humid
2	Northwestern Sahara Aquifer System	Africa	Arid	21	Guarani Aquifer System (Parana Basin)	South America	Humid
3	Murzuk-Djado Basin	Africa	Hyper-arid	22	Arabian Aquifer System	Asia	Arid
4	Taoudeni-Tanezrouft Basin	Africa	Hyper-arid	23	Indus River Basin	Asia	Semi-arid
5	Senegal-Mauritanian Basin	Africa	Semi-arid	24	Ganges-Brahmaputra Basin	Asia	Humid
6	Iullemmeden-Irhazer Aquifer System	Africa	Arid	25	West Siberian Artesian Basin	Asia	Humid
7	Lake Chad Basin	Africa	Arid	26	Tunguss Basin	Asia	Humid
8	Umm Ruwaba Aquifer (Sudd Basin)	Africa	Semi-arid	27	Angara-Lena Basin	Asia	Humid
9	Ogaden-Juba Basin	Africa	Arid	28	Yakut Basin	Asia	Humid
10	Congo Basin	Africa	Humid	29	North China Plains Aquifer System	Asia	Humid
11	Upper Kalahari-Cuvelai-Zambezi Basin	Africa	Semi-arid	30	Song-Liao Plain	Asia	Humid
12	Lower Kalahari-Stampriet Basin	Africa	Arid	31	Tarim Basin	Asia	Arid
13	Karoo Basin	Africa	Semi-arid	32	Paris Basin	Europe	Humid
14	Northern Great Plains Aquifer	North America	Sub-humid	33	East European Aquifer System	Europe	Humid
15	Cambro-Ordovician Aquifer System	North America	Humid	34	North Caucasus Basin	Europe	Semi-arid
16	California Central Valley Aquifer System	North America	Semi-arid	35	Pechora Basin	Europe	Humid
17	Ogallala Aquifer (High Plains)	North America	Semi-arid	36	Great Artesian Basin	Australia	Semi-arid
18	Atlantic and Gulf Coastal Plains Aquifer	North America	Humid	37	Canning Basin	Australia	Arid
19	Amazon Basin	South America	Humid				

Table 2. Variability (expressed as standard deviation) in GRACE-derived estimates of GWS from 20 realisations (3 GRACE-TWS and an ensemble mean of TWS, and 4 LSMs and an ensemble mean of surface water and soil moisture storage, and a snow water storage) and their reported range of uncertainty (% deviation from the ensemble mean) in world's 37 large aquifer systems.

WHYMAP aquifer no.	WHYMAP Aquifer name	Std. deviation in GRACE-GWS (cm)	Range of uncertainty (%)	WHYMAP aquifer no.	WHYMAP Aquifer name	Std. deviation in GRACE-GWS (cm)	Range of uncertainty (%)
1	Nubian Sandstone Aquifer System	1.05	83	20	Maranhao Basin	5.68	136
2	Northwestern Sahara Aquifer System	1.29	121	21	Guarani Aquifer System (Parana Basin)	3.37	77
3	Murzuk-Djado Basin	1.17	189	22	Arabian Aquifer System	2.01	163
4	Taoudeni-Tanezrouft Basin	0.99	193	23	Indus River Basin	3	78
5	Senegal-Mauritanian Basin	3.23	96	24	Ganges-Brahmaputra Basin	9.84	58
6	Iullemmeden-Irhazer Aquifer System	1.52	116	25	West Siberian Artesian Basin	7.53	79
7	Lake Chad Basin	2.23	91	26	Tunguss Basin	7.4	103
8	Umm Ruwaba Aquifer (Sudd Basin)	4.95	113	27	Angara-Lena Basin	3.73	48
9	Ogaden-Juba Basin	1.52	57	28	Yakut Basin	4.15	83
10	Congo Basin	5.09	98	29	North China Plains Aquifer System	3.93	77
11	Upper Kalahari-Cuvelai-Zambezi Basin	10.03	36	30	Song-Liao Plain	2.63	62
12	Lower Kalahari-Stampriet Basin	1.76	106	31	Tarim Basin	1.37	219
13	Karoo Basin	3.06	74	32	Paris Basin	4.06	84
14	Northern Great Plains Aquifer	4.18	111	33	East European Aquifer System	5.91	75
15	Cambro-Ordovician Aquifer System	4.56	44	34	North Caucasus Basin	4.67	66
16	California Central Valley Aquifer System	9.73	55	35	Pechora Basin	8.55	94
17	Ogallala Aquifer (High Plains)	4.05	104	36	Great Artesian Basin	2.77	69
18	Atlantic and Gulf Coastal Plains Aquifer	2.56	193	37	Canning Basin	5.34	57
19	Amazon Basin	10.93	58				

Main Figures:

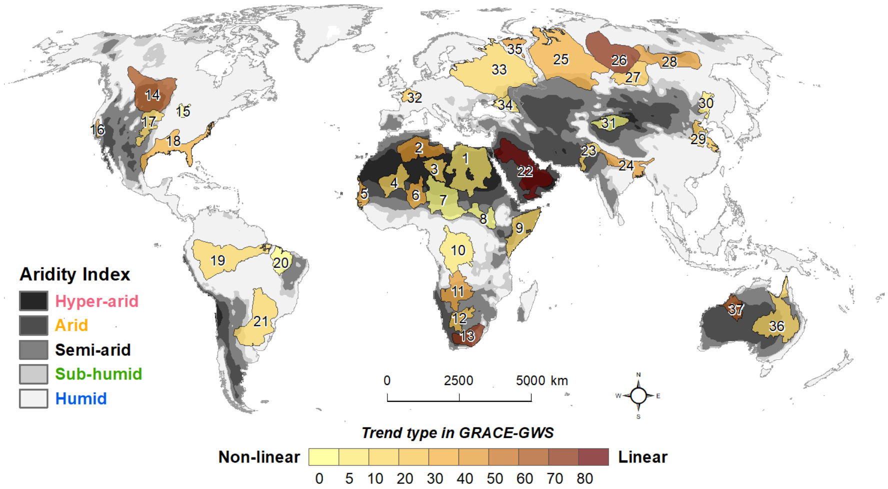

Fig. 1. Global map of 37 large aquifer systems from the GIS database of the World-wide Hydrogeological Mapping and Assessment Programme (WHYMAP); names of these aquifer systems are listed in Table 1 and correspond to numbers shown on this map for reference. Grey shading shows the aridity index based on CGIAR's database of the Global Potential Evapo-Transpiration (Global-PET) and Global Aridity Index (https://cgiarcsi.community/); the proportion (as a percentage) of long-term trends in GRACE-derived ΔGWS of these large aquifer systems that is explained by linear trend fitting is shown in colour (i.e. linear trends toward red and non-linear trends toward blue).

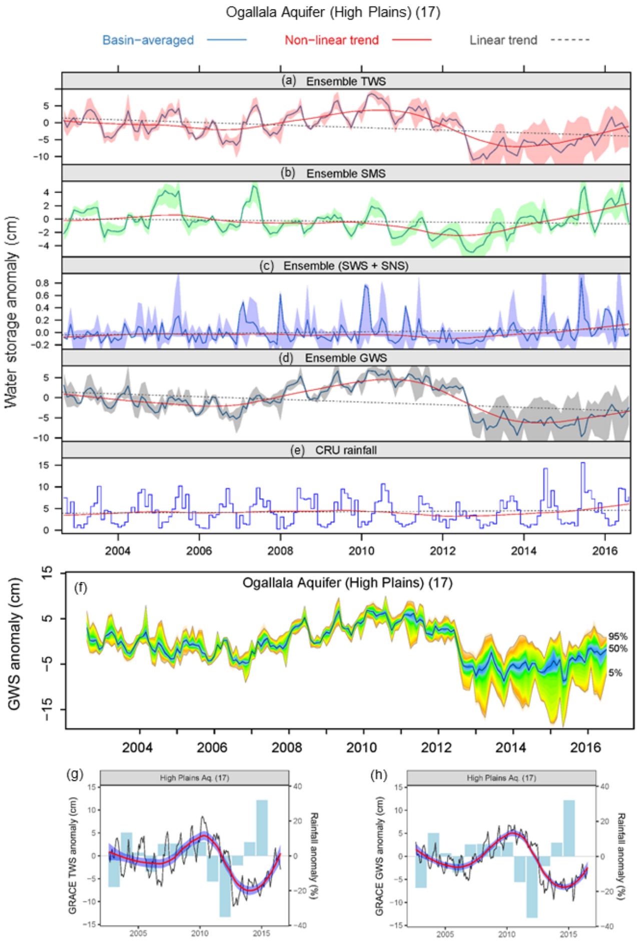

Fig. 2. Time-series data of terrestrial water storage anomaly (ΔTWS) from GRACE and individual water stores from GLDAS Land Surface Models (LSMs): (a) Ensemble monthly GRACE ΔTWS from three solutions (CSR, Mascons, GRGS), (b-c) ensemble monthly ΔSMS and ΔSWS + ΔSNS from four GLDAS LSMs (CLM, Noah, VIC, Mosaic), (d) computed monthly ΔGWS and (e) monthly precipitation from August 2002 to July 2016, (f) range of uncertainty in GRACE-derived GWS from 20 realisations, (g) ensemble TWS and annual precipitation, and (h) ensemble GRACE-derived GWS and annual precipitation for the High Plains Aquifer System in the USA (WHYMAP aquifer no. 17). Values in the Y-axis of the top four panels show monthly water-storage anomalies (cm) and the bottom panel shows monthly precipitation (cm). Time-series data (a-e) for the 36 large aquifer systems can be found in supplementary Figs. S1-S36.

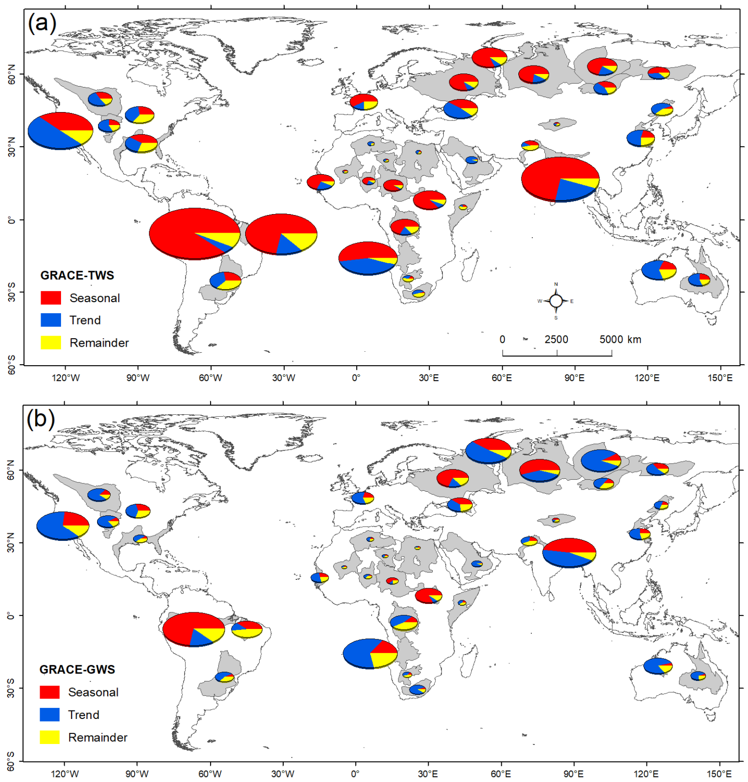

Fig. 3. Seasonal-Trend decomposition of (a) GRACE ΔTWS and (b) GRACE ΔGWS time-series data (2002 to 2016) for the world's 37 large aquifer systems using the STL decomposition method; seasonal, trend and remainder or irregular components of time-series data are decomposed and plotted as pie charts that are scaled by the variance of the time series in each aquifer system.

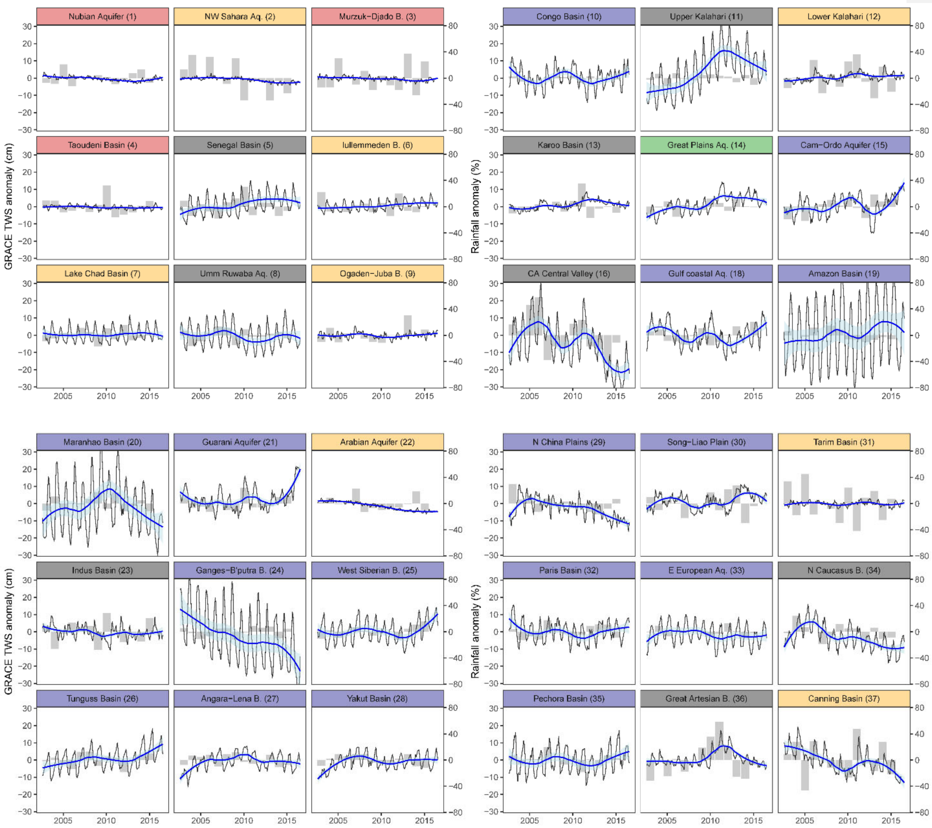

Fig. 4. Monthly time-series data (black) of ensemble GRACE ΔTWS for 36 large aquifer systems with a fitted non-linear trend line (Loess smoothing line in thick blue) through the time-series data; GRACE ΔTWS for the remaining large aquifer system (High Plains Aquifer System, (WHYMAP aquifer no. 17) is given in Fig. 2. Shaded area in semi-transparent cyan shows the range of 95% confidence interval of the fitted loess-based non-linear trends; light grey coloured bar diagrams behind the lines on each panel show annual precipitation anomaly (i.e., percentage deviation from the mean precipitation for the period of 1901 to 2016); banner colours indicate the dominant climate of each aquifer based on the mean aridity index shown in the legend on Fig. 1.

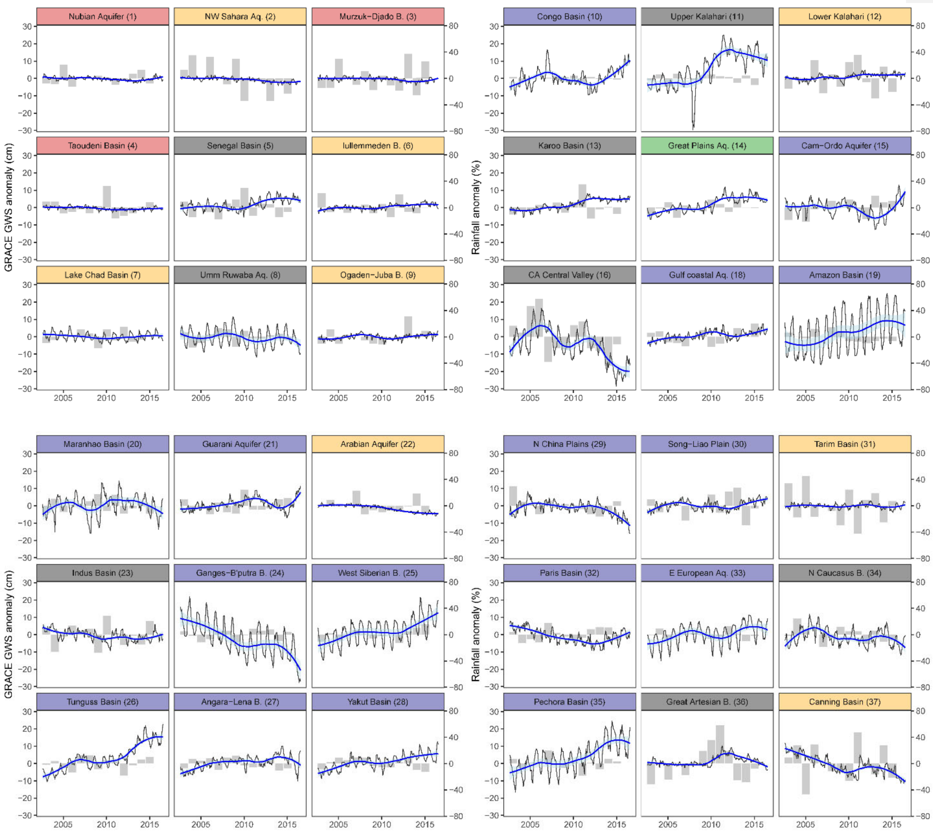

Fig. 5. Monthly time-series data (black) of ensemble GRACE ΔGWS for 36 large aquifer systems with a fitted non-linear trend line (Loess smoothing line in thick blue) through the time-series data; GRACE ΔGWS for the remaining large aquifer system (High Plains Aquifer System, (WHYMAP aquifer no. 17) is given in Fig. 2. Shaded area in semi-transparent cyan shows the range of 95% confidence interval of the fitted loess-based non-linear trends; light grey coloured bar diagrams behind the lines on each panel show annual precipitation anomaly (i.e., percentage deviation from the mean precipitation for the period of 1901 to 2016); banner colours indicate the dominant climate of each aquifer based on the mean aridity index shown in the legend on Fig. 1.

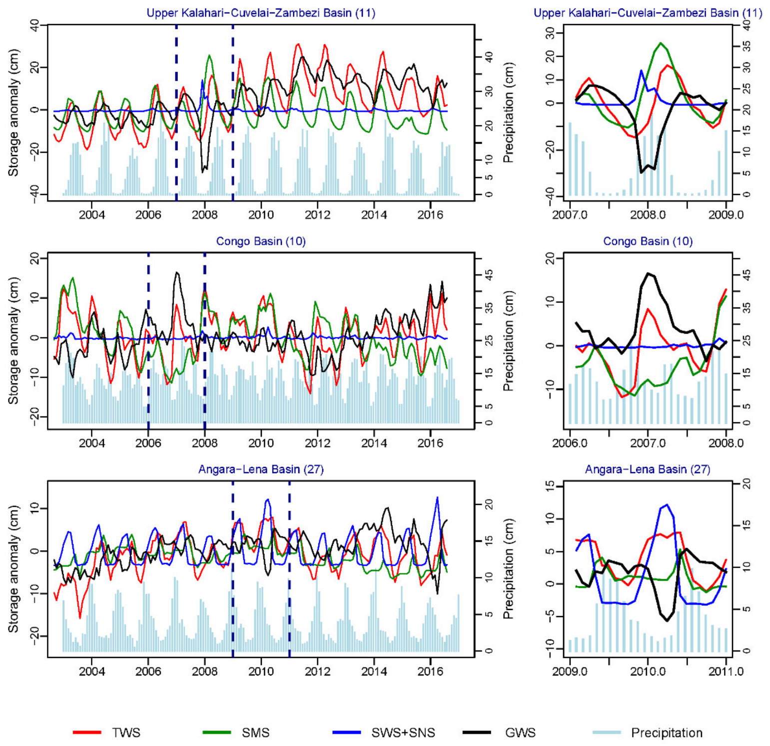

Fig. 6. Time series of ensemble mean GRACE ΔTWS (red), GLDAS ΔSMS (green), ΔSWS+ΔSNS (blue) and computed GRACE ΔGWS (black) showing the calculation of anomalously negative or positive values of GRACE ΔGWS that deviate substantially from underlying trends. Three examples include: (a) the Upper Kalahari-Cuvelai-Zambezi Basin (11) under a semi-arid climate; (b) the Congo Basin (10) under a tropical humid climate; and (c) the Angara-Lena Basin (27) under a temperate humid climate; examples from an additional five aquifer systems under semi-arid and arid climates are given in the supplementary material (Fig. S75).

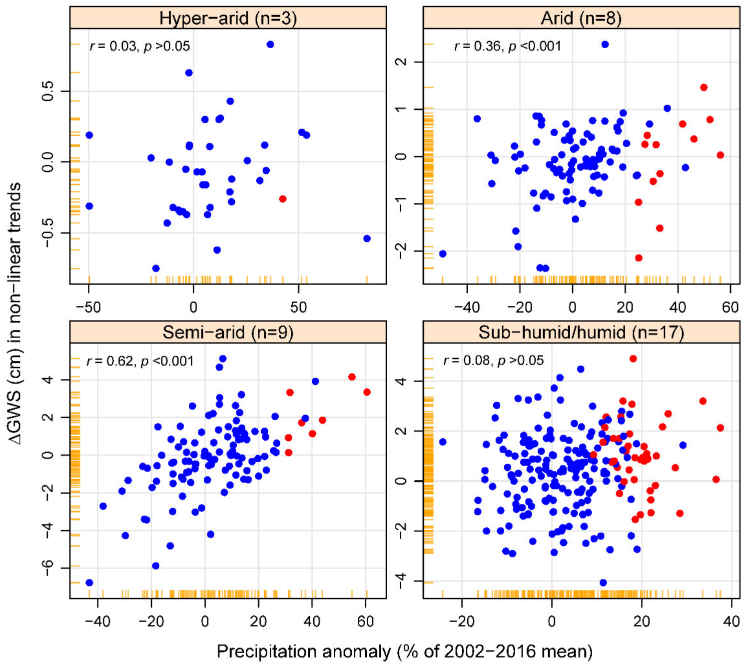

Fig. 7. Relationships between precipitation anomaly and annual changes in non-linear trends of GRACE ΔGWS in the 37 large aquifer systems grouped by aridity indices; annual precipitation is calculated based on hydrological year (August to July) for 12 of these aquifer systems and the rest 25 following the calendar year (January to December); the highlighted (red) circles on the scatterplots are the years of statistically extreme (>90th percentile; period: 1901 to 2016) precipitation.

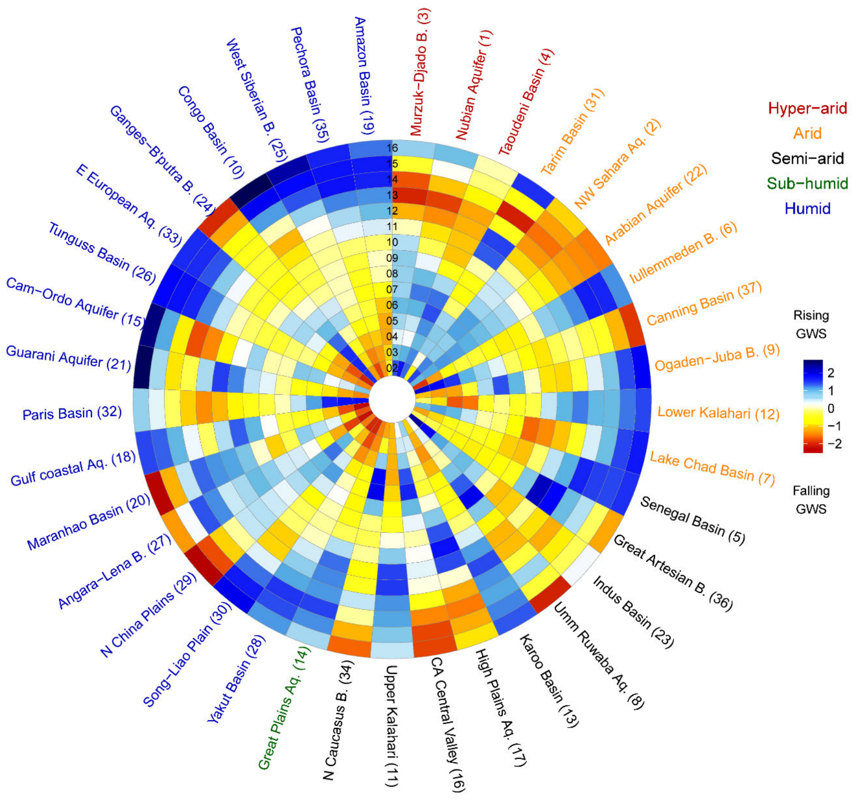

Fig. 8. Standardised monthly anomaly of non-linear trends of ensemble mean GRACE ΔGWS for the 37 large aquifer systems from 2002 to 2016. Colours yellow to red indicate progressively declining, short-term trends whereas colours cyan to navy blue indicate rising trends; aquifers are arranged clockwise according to the mean aridity index starting from the hyper-arid climate on top of the circular diagram to progressively humid. Legend colours indicate the climate of each aquifer based on the mean aridity index; time in year (2002 to 2016) is shown from the centre of the circle outwards to the periphery.