

Referee 1 Comments – Authors’ responses:

This is an interesting study that I suspect will be of wide interest to ESD readers. The premise for the research is straightforward and clearly articulated, with a generally well-reasoned account of methods and presentation of findings. I have some relatively minor points to consider for a redrafted version of this manuscript, as follows:

R1 → We appreciate the positive comments regarding the overall manuscript and the constructive suggestions to improve the manuscript. Responses to individual comments are given below.

There are quite a large number of acronyms used in the manuscript. Not all of these seem essential, and in places they detract from the overall readability. Some are rarely used (e.g. COV, PCC), others are arguably not really necessary (PCP for precipitation). I recommend the authors rethink and reduce the number of acronyms accordingly.

R2 → We will seek to reduce the number of acronyms in the final manuscript including some of the suggested that are rarely used. Use of PCP is expected to remain as we apply this in the context of identifying both PCP and PCPA (anomalies) and consider these to be clear, sensible abbreviations. [Review undertaken and actioned but usage considered helpful in most instances where 3 letters replace 3 words and/or acronym used in figure or table captions]

Inconsistent capitalisation of nouns. Some terms that are converted to acronyms are not capitalised (e.g. global hydrological models, GHMs – p2, line 52), whereas others are (Groundwater Storage, ΔGWS).

R3 → This comment is helpful in identifying necessary corrections that we will make in the course of the review of acronyms. [Capitalisation corrected]

P2, line 61-62. Presumably this increase of 15%/decade is a global average?

R4 → We will amend to clarify the text as ‘...rapid escalation in global groundwater extraction at an average of ~15% per decade...’ [NB this has now been eliminated by the change agreed in response to Reviewer 2 comments L63-65]

P7, lines 168 and 189. It is potentially somewhat misleading to provide coefficient of variation statistics for sample sizes of 3 and 4. The mean, standard deviation (and thus the coefficient of variation too) are not particularly informative descriptors of central tendency or distribution of samples that are this small.

R5 → We acknowledge that the population (not sample) size is small for the use of coefficient of variation. We nevertheless consider this parameter has merit in explaining divergence among the respective datasets; we propose to include a phrase acknowledging the small population sizes. [Phrases added]

45

P7, line 192. Improper use of the term “significant” when discussing the magnitude of statistical terms. From my reading of this sentence, “substantial” would be a more appropriate descriptor.

50

R6 → We agree and will amend to read ‘substantial’ and check all other uses of the word, ‘significant’ (3 instances), are appropriate. [Done]

55

Figure 3. The different y-axis scales, and scaling of the data are noted in the caption, but this twin approach to scaling, plus the relatively small size of the panes and faint axis labels makes it somewhat difficult to compare scales of variability between the different aquifers.

60

R7 → We appreciate this comment and will look at ways to improve the clarity of what is being expressed in Figure 3 and improve axis labels. As the primary purpose of the figure is to illustrate correlation between change in GWS and precipitation anomaly for individual systems, we have scaled the data to make these correlations clear. Further information on the variability between aquifers is given in the final two columns of Table 1. [Spatial considerations make this difficult to improve]

65

P16, line 400. The reference for this statement is slightly misleading. It could be read that Zwiers and Von Storch (1995) found robust statistical significance, but actually I think the authors mean that robust statistical significance was found using the methods described by Zwiers and Von Storch.

70

R8 → We agree with that this clarification is necessary and will amend to state: ‘when tested using the methods described by Zwiers and Von Storch (ref.)’ [Done]

75

Authors’ responses to Bridget Scanlon’s Review:

80

We thank Bridget Scanlon for her review and support for the general intent of the paper. We are encouraged that, apart from minor comments, the methods, results and findings of the paper are considered robust and useful. We appreciate that the appended comments are helpful and intended to improve the manuscript.

85

This response is also provided in the form of a pdf supplement.

We provide the following responses to specific comments:

Abstract:

90 'In the abstract the authors refer to ENSO as seasonal; however, I think of ENSO as more interannual with 3 – 5 yr timescales'

95 **AR 1: We agree and propose simply to omit the word 'seasonal'. The original reason for the inclusion of this adjective was the known (inter-annual) periodicity in the control that ENSO exerts on seasonal precipitation (i.e. heavy rains, drought). [Done]**

L40 – L42:

100 'The authors refer to baseflow from groundwater sustaining rivers and wetlands being fundamentally important, especially in semiarid and arid regions; however, one should recognize that in many semiarid regions surface water recharges groundwater.'

105 **AR 2: This characterisation is misrepresentative and a regretted oversight on our part; we propose to amend the wording as follows:**

"Baseflow from groundwater sustains rivers and wetlands in the absence of rainfall and is therefore fundamentally important to the ecology of semi-arid and arid regions in particular..."

110 **to be replaced by:**

115 **"Bidirectional flows between surface water and groundwater are fundamentally important to the ecology of semi-arid and arid regions (drylands), where surface water often recharges groundwater and baseflow from groundwater can sustain rivers and wetlands in the absence of rainfall" [Done]**

L45:

120 'The authors suggest that future management of freshwater resources will be a critical issue linked to climate change but I think it is already a critical issue because of climate extremes (droughts and floods)'

AR 3: Agreed – we propose to amend this sentence:

125 **"Future climate change in which anthropogenic emissions of greenhouse gases transform patterns of natural variability, together with substantial socio-economic change, predicate that management of freshwater resources will become a critical task ((Famiglietti, 2014)."**

to read:

130 **"Climate change in which anthropogenic emissions of greenhouse gases transform patterns of natural variability, together with substantial socio-economic change, predicates that management of freshwater resources has and will increasingly become a critical task (Famiglietti, 2014)."** [Done]

135

L63 – 65:

'In contrast to the statement from Wada et al., 2014; recent reports suggest that water use has been stable or decreasing in past decades in the U.S. and China. Zhou, F., et al. (2020), Deceleration of China's human water use and its key drivers, Proceedings of the National Academy of Sciences, 117(14), 7702-7711. https://www.usgs.gov/special-topic/water-science-school/science/trends-water-use/united-states-1950-2015?qt-science_center_objects=0#qt-science_center_objects'

AR 4: We appreciate this insight and helpful suggestion. Zhou et al. (2020) specifically refer to a slowing in the rate of the increase in water use in China (Zhou et al., 2020) whereas the USGS online report clearly highlights a decline in total freshwater withdrawals in the US. To avoid confusion, we have revised the text here to refer simply to the estimate that relates to global groundwater withdrawals so that it is amended to read:

"In the context of an ~85% increase in global groundwater abstraction from 1979 to 2010 (Wada et al., 2014), an understanding..." [Done]

L175-180:

'GRACE uses a baseline from 2004 – 2009; however, I think it would be better to calculate anomalies for your data based on your entire record. It is not clear why you don't combine JPL and CSR mascons? JPL relies on models to process the GRACE data whereas CSR does not. The combination should be more robust.'

AR 5: We appreciate these comments and, with respect to the suggested baseline, have conducted the analysis as proposed. Our TWS anomalies are calculated based on the entire record which means they have been adjusted to a time mean baseline taken over the whole period of GRACE data. This text on lines 175 to 178 is intended to convey that methodology:

"GRACE Δ TWS is not a time-invariant measure (Wahr et al., 1998) and in the standard datasets all anomalies are given with respect to a baseline which is the mean over the period January 2004 to December 2009 (JPL NASA, 2019). Here, the completed available GRACE Δ TWS time series is examined with respect to climate anomalies in the same timeframe. Consequently, the JPL-Mascon Δ TWS dataset has been rescaled with respect to a time-mean taken over the whole period of GRACE operation (08.2002 – 07.2016), which is the study reference period (SRP) (JPL NASA, 2019)."

To clarify we will amend this passage to:

"GRACE Δ TWS is not a time-invariant measure (Wahr et al., 1998) and in the standard datasets all anomalies are given with respect to a baseline which is the mean over the period January 2004 to December 2009 (JPL NASA, 2019). However, we examine the completed available GRACE Δ TWS time series with respect to climate anomalies over the consistent timeframe of the entire series. Consequently, the employed JPL-Mascon Δ TWS

185 dataset has been rescaled with respect to a time-mean taken over the whole period of GRACE operation (08.2002 – 07.2016), which is the study reference period (SRP) (JPL NASA, 2019).” [Done]

190 On the use of CSR Mascons, we did consider this and we decided that for this particular use of the JPL Mascons, the strong correlation (rank correlation coefficients mostly >0.9) between the CSR Mascons and JPL Mascons Δ TWS time series, identified by Scanlon et al. (2016) (Scanlon et al., 2016), was sufficient to provide a robust result using JPL Mascons alone. Specifically, we relied on the following arguments:

- 195 • Since the data are adjusted with respect to the time mean baseline for the entire series, we are not concerned with long-term trends which are the major source of divergence between the JPL and CSR datasets.
- With the exception of California Central Valley, nine of the analysed aquifer systems are large (>500 km²) and 4 are medium scale (>100 km²), meaning that the correlation between JPL and CSR Mascon datasets is robust due to the scale of the aquifers studied.
- 200 • Our study is concerned with Tropical and Sub-Tropical latitudes where the resolution of JPL Mascons (~300 km) is equivalent to that of the raw GRACE data, minimising concerns of processing and again improving correlation between the two Mascon datasets.
- 205 • We consider that the greatest source of uncertainty in the GRACE data for this study is derived from the use of one or more LSM’s to derive Δ GWS from Δ TWS, as discussed in the paper (L90-93, L188-190)

210 L237:

‘It would be good to use a consistent time period for calculating all anomalies throughout the paper.’

215 AR 6: We agree and this is indeed what we have done in the paper. To make this point as clear as possible, we will amend the current text:

“PCPA = monthly precipitation anomalies with respect to the time-mean baseline for the study reference period (SRP – 2002-16)”

220 to read:

“PCPA = monthly precipitation anomalies with the respect to the consistently applied study reference period time-mean baseline, 2002 – 2016.” [Done]

225 L310

‘I think it is important to indicate the uncertainties in these global datasets. For example, the water table depth map developed by Fan differs markedly from that developed by the British Geological Survey for Africa’

230

AR 7: We agree that it is important to indicate the uncertainties in these global datasets and propose to amend the text as set out below. The divergence between the Fan and BGS datasets for Africa is primarily one of scaling in which Fan shows 5 divisions between 0 m and 25 m depth whereas the BGS shows only 2 divisions (Fan et al., 2013; MacDonald et al., 2012). Within the constraints of the BGS scaling, the two datasets apparently agree reasonably well for most of the continent. For this study, the relevant exception is the Karoo Basin where the two datasets disagree; BGS data show some shallower areas of water-table depth that do not appear in the Fan dataset. We recognise that the response of the Karoo Basin to recharge is to some extent anomalous in this study but we do not explore the reasons in any detail.

Proposed amendment:

235

L318 on: 'However, K is based on permeability mapping from hydrogeologies that have a standard deviation of ~2 orders of magnitude (Gleeson et al., 2011) and this variance underlies the uncertainty in each of these datasets used. WTD is 30 arc-second (~1km) resolution dataset compiled from available observational data extended by modelled interpolation with both these data sources being subject to considerable sampling bias and model uncertainty respectively (Fan et al., 2013). All of these datasets are global and derived from combinations of observations and modelled data.' [Done]

240

245

250

L423:

'Long and Mahler (2013) applied the analysis to karst aquifers, which are similar to surface water drainage systems. These systems differ markedly from many aquifer systems.'

255

AR 8: The original and current intent of this reference (Long and Mahler, 2013) is to illustrate the use of transfer functions in hydraulic analysis and efficiencies that have been derived in the use of metrics. However, we do not wish there to be any confusion arising from this contextual use of the reference and propose to omit it and to amend L420-426:

260

"HM ultimately derives from the physical properties of the saturated portion of the aquifer system (Townley, 1995) and system memory as measured by Eq. (5) is representative of the physical properties of an aquifer system and its climate. Long and Mahler (2013), for example, used a total of 16 metrics to describe particular North American Karst aquifer Systems in their IRF (θ in Eq. (5)). In contrast, von Asmuth and Knotters (2004) used 4 parameters to describe groundwater dynamics in their transfer function that they argue represents a more accurate description of the physical system than previously used parametric methods."

265

270

to read:

"HM ultimately derives from the physical properties of the saturated portion of the aquifer system (Townley, 1995) and system memory as measured by Eq. (5) is representative of the physical properties of an aquifer system and its climate. Von Asmuth and Knotters (2004) use 4 parameters to describe groundwater dynamics in their transfer

275

function (θ in Eq. (5)) that they argue represents a more accurate description of the physical system than previously used parametric methods (von Asmuth and Knotters, 2004).” [Done]

References:

- Famiglietti, J.S., 2014. The global groundwater crisis. *Nature Climate Change* 4, 945–948. <https://doi.org/10.1038/nclimate2425>
- Fan, Y., Li, H., Miguez-Macho, G., 2013. Global Patterns of Groundwater Table Depth. *Science* 339, 940–943. <https://doi.org/10.1126/science.1229881>
- Gleeson, T., Smith, L., Moosdorf, N., Hartmann, J., Dürr, H.H., Manning, A.H., Beek, L.P.H. van, Jellinek, A.M., 2011. Mapping permeability over the surface of the Earth. *Geophysical Research Letters* 38. <https://doi.org/10.1029/2010GL045565>
- JPL NASA, 2019. Frequently Asked Questions | About [WWW Document]. GRACE Tellus. URL <https://grace.jpl.nasa.gov/about/faq> (accessed 8.20.19).
- Long, A.J., Mahler, B.J., 2013. Prediction, time variance, and classification of hydraulic response to recharge in two karst aquifers. *Hydrology and Earth System Sciences* 17, 281–294. <https://doi.org/10.5194/hess-17-281-2013>
- MacDonald, A.M., Bonsor, H.C., Dochartaigh, B.É.Ó., Taylor, R.G., 2012. Quantitative maps of groundwater resources in Africa. *Environ. Res. Lett.* 7, 024009. <https://doi.org/10.1088/1748-9326/7/2/024009>
- Scanlon, B.R., Zhang, Z., Save, H., Wiese, D.N., Landerer, F.W., Long, D., Longuevergne, L., Chen, J., 2016. Global evaluation of new GRACE mascon products for hydrologic applications: GLOBAL ANALYSIS OF GRACE MASCON PRODUCTS. *Water Resources Research* 52, 9412–9429. <https://doi.org/10.1002/2016WR019494>
- Townley, L.R., 1995. The response of aquifers to periodic forcing. *Advances in Water Resources* 18, 125–146. [https://doi.org/10.1016/0309-1708\(95\)00008-7](https://doi.org/10.1016/0309-1708(95)00008-7)
- von Asmuth, J.R., Knotters, M., 2004. Characterising groundwater dynamics based on a system identification approach. *Journal of Hydrology* 296, 118–134. <https://doi.org/10.1016/j.jhydrol.2004.03.015>
- Wada, Y., Wisser, D., Bierkens, M.F.P., 2014. Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. *Earth System Dynamics* 5, 15–15.
- Wahr, J., Molenaar, M., Bryan, F., 1998. Time variability of the Earth’s gravity field: Hydrological and oceanic effects and their possible detection using GRACE. *Journal of Geophysical Research: Solid Earth* 103, 30205–30229. <https://doi.org/10.1029/98JB02844>
- Zhou, F., Bo, Y., Ciais, P., Dumas, P., Tang, Q., Wang, X., Liu, J., Zheng, C., Polcher, J., Yin, Z., Guimberteau, M., Peng, S., Oettle, C., Zhao, X., Zhao, J., Tan, Q., Chen, L., Shen, H., Yang, H., Piao, S., Wang, H., Wada, Y., 2020. Deceleration of China’s human water use and its key drivers. *PNAS* 117, 7702–7711. <https://doi.org/10.1073/pnas.1909902117>

Climate-groundwater dynamics inferred from GRACE and the role of hydraulic memory

Simon Opie^{1,*}, Richard G. Taylor¹, Chris M. Brierley¹, Mohammad Shamsudduha² and Mark O. Cuthbert^{3,4}

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

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

³ School of Earth and Ocean Sciences, Cardiff University, Cardiff, UK

⁴ Connected Waters Initiative Research Centre, University of New South Wales, Sydney, New South Wales, Australia

* Corresponding author: Simon Opie (simon.opie.18@ucl.ac.uk)

Abstract

Groundwater is the largest store of freshwater on Earth after the cryosphere and provides a substantial proportion of the water used for domestic, irrigation and industrial purposes. Knowledge of this essential resource remains incomplete, in part, because of observational challenges of scale and accessibility. Here we examine a 14-year period (2002-2016) of GRACE observations to investigate climate-groundwater dynamics of 14 tropical and sub-tropical aquifers selected from WHYMAP's 37 large aquifer systems of the world. GRACE-derived changes in groundwater storage resolved using GRACE JPL Mascons and the CLM Land Surface Model are related to precipitation time series and regional-scale hydrogeology. We show that aquifers in dryland environments exhibit long-term hydraulic memory through a strong correlation between groundwater storage changes and annual precipitation anomalies integrated over the time series; aquifers in humid environments show short-term memory through strong correlation with monthly precipitation. This classification is consistent with estimates of Groundwater Response Times calculated from the hydrogeological properties of each system, with long (short) hydraulic memory associated with slow (rapid) response times. The results suggest that groundwater systems in dryland environments may be less sensitive to seasonal climate variability but vulnerable to long-term trends from which they will be slow to recover. In contrast, aquifers in humid regions may be more sensitive to climate disturbances such as ENSO-related drought but may also be relatively quick to recover. Exceptions to this general pattern are traced to human interventions through groundwater abstraction. Hydraulic memory is an important factor in the management of groundwater resources, particularly under climate change.

1.0. Introduction:

Formatted: Line spacing: single

Formatted: Font: 12 pt, Not Bold

Commented [so1]: Reviewer 2 AR 1

Deleted: seasonal

360

The availability of freshwater is essential for sustaining human life, economic security, and access to the benefits of a wide range of ecosystem services (Taylor et al., 2013a). After the cryosphere, groundwater is the second largest store of freshwater on the planet supplying 36% of domestic water, 42% of irrigation for agriculture and 27% of industrial water use (Döll et al., 2012).

365

Bidirectional flows between surface water and groundwater are fundamentally important to the ecology of semi-arid and arid regions (drylands), where surface water often recharges groundwater and baseflow from groundwater can sustain rivers and wetlands in the

Formatted: Font: Not Bold, Not Italic

absence of rainfall (Alley et al., 2002; Graaf et al., 2019). Climate change in which

Commented [so2]: Reviewer 2 AR 2

Deleted: Baseflow from groundwater sustains rivers and wetlands in the absence of rainfall and is therefore fundamentally important to the ecology of semi-arid and arid regions in particular

370

together with substantial socio-economic change, predicates that management of freshwater resources has and will increasingly become a critical task (Famiglietti, 2014). In a climate where it is broadly predicted that 'wet gets wetter, dry gets drier' (Trenberth, 2011), water storage at and below the land surface will be a vital tool in enabling successful adaptation to the changing global environment (Damkjaer and Taylor, 2017; Wada, 2016).

Formatted: Font: Not Bold

Commented [so3]: Reviewer 2 AR 3

Deleted: Future climate change in which anthropogenic emissions of greenhouse gases transform patterns of natural variability, together with substantial socio-economic change, predicate that management of freshwater resources will become a critical task

375

Despite the importance of groundwater there are considerable gaps in current knowledge and understanding (Güntner et al., 2007). Direct observations of groundwater are sparse in relation to its geographical scale so most global or regional groundwater data are based on output from large-scale models. These include Global Hydrological Models (GHMs) (Sood and Smakhtin, 2015) or

Land Surface Models (LSMs) (Bierkens, 2015; Overgaard et al., 2006; Wood et al., 2011) for which

Deleted: g

Deleted: h

Deleted: m

380

there are often insufficient data available to constrain or calibrate (Döll et al., 2016). Model simulation of key processes such as soil hydrodynamics and groundwater recharge is therefore based on theoretical frameworks rather than field data (Scanlon et al., 2002). As a result, there is also considerable uncertainty about climate-groundwater dynamics. Recent work in this area has either focused on localised observations of changes in Groundwater Storage (Δ GWS) from

Commented [so4]: Changes as per reviewer 1 AR 3

Deleted: l

Deleted: -

Deleted: s

Deleted: m

385

piezometry (Cuthbert et al., 2019b) or occurred adjacent to large centres of population where human intervention, through extraction of groundwater by pumping, can greatly influence observational measurements (Scanlon et al., 2018). In the context of an ~85% increase in global groundwater abstraction from 1979 to 2010 (Wada et al., 2014), an understanding of climate-

Commented [so5]: Reviewer 2 AR 4

Deleted: a recent, rapid escalation in groundwater abstraction of ~15% per decade from 1960 to 2010

390

sustainable access to groundwater resources (Taylor et al., 2009).

In response to the lack of in situ field observations, remote-sensing by satellite is increasingly being utilised to expand the scope of observational data available to Earth sciences (Acker and Leptoukh, 2007). An important advance in the quality of global data for hydrological

studies has come from the Gravity Recovery and Climate Experiment (GRACE), a collaboration between the National Aeronautics and Space Administration (NASA) in the USA and the German Aerospace Centre (DLR) launched in March 2002 (Tapley et al., 2004). Completed sets of ~monthly measurements are used to derive the changes in mass at the Earth's surface and from these data mass fluxes can be extracted that directly relate to the hydrosphere. Over land, the flux is expressed as a change in Total Water Storage (Δ TWS) at a spatial resolution of ~300km and with an expected accuracy of better than 2 cm equivalent water height (EWH) (Tapley et al., 2004). GRACE ceased operation due to battery failure in mid-2016 having created a record of 163 monthly gravity solutions (Tapley et al., 2019). Although GRACE operated for ten years longer than anticipated at its launch, it is a relatively brief dataset in relation to large-scale climate patterns impacting the global hydrological system with frequencies of several years or decades (e.g. Pacific Decadal Oscillation (PDO), Atlantic Multidecadal Oscillation (AMO)). Nevertheless, inter-annual periodicities associated with the El Niño Southern Oscillation (ENSO) and the Antarctic Circumpolar Wave (ACW) have been detected (Mémin et al., 2015; Ni et al., 2018; Phillips et al., 2012).

Intrinsic parameters of GRACE data effectively define the spatial and temporal dimensions of this study but there are additional constraints related to the derivation of Δ GWS data from GRACE Δ TWS that also need to be considered. The sub-division of GRACE Δ TWS into its component parts, including Δ GWS, requires the application of GHM or LSM output that is itself subject to associated uncertainty, as already noted (Döll et al., 2014). It has been demonstrated that there is relatively poor correlation between GRACE and GHMs/LSMs in the evaluation of Δ TWS, with significant discrepancies at the basic level of whether storage trends are increasing or decreasing (Scanlon et al., 2018). These findings have been confirmed with reference to regional piezometric groundwater measurements from tropical aquifers in Africa (Bonsor et al., 2018). Thus, the application of GRACE data to Δ GWS implies three distinct areas of uncertainty: in the processing of the GRACE signal, accuracy of GHM/LSM model projections and mutual consistency of the observed (GRACE) and modelled (GHM/LSM) data (Long et al., 2015).

This study investigates the spatio-temporal properties of climate-groundwater dynamics using a subset of the 37 Large Aquifer Systems of the World (LASW) as defined by the Worldwide Hydrogeological Mapping and Assessment Programme (WHYMAP) ("BGR - WHYMAP - Large Aquifers," 2008), and shown in **Figure S1**. This subset comprises aquifers that lie broadly within the tropics and sub-tropics climate variability is mostly defined by rainfall (Shepherd, 2014). The 14 aquifers selected are listed in **Table 1** together with their key characteristics including Aridity Index (AI) calculated from the Consultative Group for International Agriculture Research's Consortium for Spatial Information (CGIAR-CSI) Global-Aridity Dataset (Trabucco and Zomer, 2019), shown in **Figure**

1. Following the work of (Shamsudduha and Taylor, 2019), the groundwater storage response to regional climate variability for these 14 large scale aquifer systems is investigated using Δ GWS data extracted from the whole of the available GRACE Δ TWS time series (August 2002 – July 2016) together with climate data that are defined by the areal extent of each of the aquifer systems.

450 Several studies have used GRACE data to examine storage changes within a particular GW system e.g. (Becker et al., 2010; Bonsor et al., 2018; Chen et al., 2016, 2010; Henry et al., 2011; Z. Huang et al., 2015; Ramillien et al., 2014; Shamsudduha et al., 2017, 2012; Tiwari et al., 2009; Xavier et al., 2010; Yeh et al., 2006). Here, we examine the dynamics of climate-groundwater interactions inferred from the underlying patterns of large-scale Δ GWS in response to extremes of precipitation.

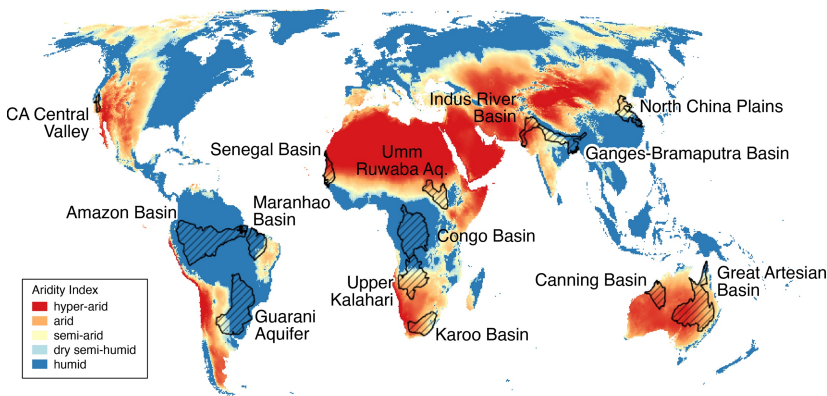
455 We find that Hydraulic Memory (HM) is a key component in the classification of groundwater response to climate variability. We then seek to reconcile the results with reference to the physical characteristics of individual aquifer systems (Cuthbert et al., 2019a) whilst accounting for anomalous responses in Δ GWS to climate variability.

Commented [so6]: Reviewer 1 AR 3

Deleted: h

Deleted: m

460



465 **Figure 1: 14 of the World's Large-scale Aquifers (the Study Aquifers) overlaid on CGIAR-CSI Global-Aridity dataset (Trabucco & Zomer, 2019).**

<u>WHYMAP Aquifer Number</u>	<u>Aquifer name</u>	<u>Continent</u>	<u>Population (millions)</u>	<u>Aquifer area (km²)</u>	<u>Proportion of Irrigation GW fed (%)</u>	<u>Climate zone based on aridity index</u>	<u>Mean (2002-16) annual precipitation (mm)</u>	<u>Rainfall variability (%)</u>
<u>5</u>	<u>Senegal-Mauritanian Basin</u>	<u>Africa</u>	<u>17.77</u>	<u>295k</u>	<u>1.0</u>	<u>Semi-Arid</u>	<u>540</u>	<u>14.6</u>
<u>8</u>	<u>Umm Ruwaba Aquifer</u>	<u>Africa</u>	<u>10.52</u>	<u>509k</u>	<u>0.0</u>	<u>Semi-Arid</u>	<u>789</u>	<u>10.7</u>
<u>10</u>	<u>Congo Basin</u>	<u>Africa</u>	<u>34.74</u>	<u>1.49m</u>	<u>0.0</u>	<u>Humid</u>	<u>1566</u>	<u>5.6</u>
<u>11</u>	<u>Upper Kalahari-Cavelai-Zambezi Basin</u>	<u>Africa</u>	<u>6.02</u>	<u>1.00m</u>	<u>0.1</u>	<u>Semi-Arid</u>	<u>819</u>	<u>10.0</u>
<u>13</u>	<u>Karoo Basin</u>	<u>Africa</u>	<u>14.53</u>	<u>568k</u>	<u>2.1</u>	<u>Semi-Arid</u>	<u>479</u>	<u>17.6</u>
<u>16</u>	<u>California Central Valley Aquifer System</u>	<u>North America</u>	<u>8.10</u>	<u>71k</u>	<u>57.8</u>	<u>Semi-Arid</u>	<u>515</u>	<u>32.0</u>
<u>19</u>	<u>Amazon Basin</u>	<u>South America</u>	<u>8.93</u>	<u>2.28m</u>	<u>1.0</u>	<u>Humid</u>	<u>2505</u>	<u>8.3</u>
<u>20</u>	<u>Maranhao Basin</u>	<u>South America</u>	<u>10.81</u>	<u>593k</u>	<u>32.6</u>	<u>Humid</u>	<u>1502</u>	<u>15.7</u>
<u>21</u>	<u>Guarani Aquifer (Parana Basin)</u>	<u>South America</u>	<u>47.84</u>	<u>1.83m</u>	<u>20.5</u>	<u>Humid</u>	<u>1450</u>	<u>10.6</u>
<u>23</u>	<u>Indus River Basin</u>	<u>Asia</u>	<u>155.85</u>	<u>308k</u>	<u>31.0</u>	<u>Arid</u>	<u>375</u>	<u>16.2</u>
<u>24</u>	<u>Ganges-Brahmaputra Basin</u>	<u>Asia</u>	<u>596.44</u>	<u>616k</u>	<u>55.8</u>	<u>Humid</u>	<u>1391</u>	<u>12.1</u>
<u>29</u>	<u>North China Plains Aquifer System</u>	<u>Asia</u>	<u>336.70</u>	<u>439k</u>	<u>37.1</u>	<u>Dry Sub-Humid</u>	<u>826</u>	<u>10.0</u>
<u>36</u>	<u>Great Artesian Basin</u>	<u>Australia</u>	<u>0.20</u>	<u>1.77m</u>	<u>0.9</u>	<u>Arid</u>	<u>444</u>	<u>28.9</u>
<u>37</u>	<u>Canning Basin</u>	<u>Australia</u>	<u>0.01</u>	<u>433k</u>	<u>0.4</u>	<u>Arid</u>	<u>443</u>	<u>21.2</u>

Table 1: Characteristics of the 14 Aquifer Systems selected for the study according to the WHYMAP and CGIAR-CSI databases with statistics giving (L to R): total number of resident population, aquifer area, proportion of irrigation GW-fed, mean aridity index classification (Trabucco and Zomer, 2019), mean annual rainfall and mean variability in annual rainfall.

475

Deleted: ¶
INSERT [Table 1], [Figure 1]¶

480 **Methods:**

2.1. GWS derived from GRACE data:

485 Mass fluxes relating to the hydrosphere contained in the GRACE land-signal measurement of changes in the Earth's gravitational field are defined as ΔTWS . In order to obtain information relating specifically to groundwater, this signal is separated into the component parts that comprise TWS, generally represented as:

$$\Delta TWS = \Delta GWS + \Delta SWS + \Delta SMS + \Delta SNS \quad (1)$$

490

where SWS is surface water storage, SMS is soil moisture storage and SNS is snow-water equivalent storage. ΔGWS is then derived from ΔTWS according to the following equation:

$$\Delta GWS = \Delta TWS - (\Delta SWS + \Delta SMS + \Delta SNS) \quad (2)$$

495

The locations of the 14 aquifers are outside areas where changes in snow-water equivalent substantially impact ΔTWS (Getirana et al., 2017). ΔSNS can consequently be omitted so that [Eq. \(2\)](#) can be rewritten for the purposes of this particular study as:

500

$$\Delta GWS = \Delta TWS - (\Delta SWS + \Delta SMS) \quad (3)$$

505 Since GRACE started transmitting, several solutions have been developed for analysing and producing GRACE ΔTWS data to increasing levels of accuracy, with the intention that the data be readily and freely available for research (Landerer and Swenson, 2012). In this instance, three different products were drawn from Shamsudduha and Taylor (2019), two of which are ~~Spherical~~ **Harmonics** (SH) solutions comprising CSR Land (version RL05.DSTvSCS1409) from the Jet Propulsion Laboratory (JPL) at NASA and CNES/GRGS (version RL03-v1) from the French Centre National d'Etudes Spatiales, and one JPL-Mascon (version RL05M 1.MSCNv01) from JPL-NASA. To derive ΔTWS , all GRACE solutions require additional processing that include corrections for glacial isostatic rebound and atmospheric mass variation (Landerer and Swenson, 2012). SH solutions also require spatial filtering (or 'de-striping') whereas JPL-Mascon does not as it directly converts the GRACE

Deleted: s

Commented [so7]: Reviewer 1 AR 3

Deleted: h

515 signal into mass concentration blocks (Mascons), rendering monthly gravitational fields directly as
3°x3° gridded spatial components to reduce errors (Watkins et al., 2015).

On inspection, the divergence between the 3 Δ TWS datasets was significant when summed
over the time series. The relatively large coefficient of variance, -104%, even though derived from a
small sample size, calls into question use of an ensemble mean for this study. Such an approach may
520 be appropriate for the use of SH products alone (Sakumura et al., 2014) but it is preferable not to
combine SH products and Mascons (Landerer, pers. comm.). Consequently we rely solely on the JPL-
Mascon dataset possessing a better signal-to-noise ratio and potentially less error (Scanlon et al.,
2016; Watkins et al., 2015; Xie et al., 2018). The employed JPL-Mascon dataset has been spatially
525 sampled at a 0.5° grid using dimensionless scaling factors provided as 0.5°x0.5° bins derived from
the CLM4.0 LSM (Long et al., 2015; Wiese et al., 2016). GRACE Δ TWS is not a time-invariant measure
(Wahr et al., 1998) and in the standard datasets all anomalies are given with respect to a baseline
which is the mean over the period January 2004 to December 2009 (JPL NASA, 2019). However, we
examine the completed available GRACE Δ TWS time series with respect to climate anomalies over
the consistent timeframe of the entire series. Consequently, the employed JPL-Mascon Δ TWS
530 dataset has been rescaled with respect to a time-mean taken over the whole period of GRACE
operation (08.2002 – 07.2016), which is the Study Reference Period (SRP) (JPL NASA, 2019).

As set out in Eq. (3), datasets for Δ SMS and Δ SWS derived from LSMs are required to
determine Δ GWS from Δ TWS since observational data at the spatio-temporal scales of this study do
not exist. Datasets for the 14 aquifer systems were drawn from NASA's Global Land Data
535 Assimilation System (GLDAS) (Rodell et al., 2004) comprising the output from four different LSMs
(Shamsudduha and Taylor, 2019): the Common [Community] Land Model (CLM, version 2.0), Noah
(version 2.7.1), the Variable Infiltration Capacity (VIC) model (version 1.0), and Mosaic (version 1.0)
(Rui and Beaudoin, 2019). As with Δ TWS, analysis of the four LSM datasets for Δ SWS+ Δ SMS shows
that their divergence summed over the entire time series is substantial, with a coefficient of
540 variance of 258%, suggesting that a LSM-ensemble mean approach may also not be appropriate for
this analysis, even given the restricted sample size. Further, the inter- and intra- model variability of
 Δ SWS in the LSM datasets, assessed as surface runoff (e.g. Shamsudduha and Taylor, 2019; Thomas
et al., 2017), is much less substantial than that of Δ SMS (inter-model coefficient of variance 378%).
545 In the absence of consideration of Δ SWS, groundwater recharge is primarily determined by the
effect of evapotranspiration on moisture in the soil zone (Long and Mahler, 2013). Therefore, for this
study, modelling of Δ SMS is a key determinant of the outcomes for Δ GWS computed using Eq. (3)
(de Vries and Simmers, 2002). Modelled soil profiles vary substantially in each of the 4 LSMs ranging
in depth from 3.5m (Mosaic) to 1.9m (VIC) and, in vertical layers, from 10 (CLM) to 3 (VIC & Mosaic)

Commented [so8]: Reviewer 1 AR 3

Deleted: c

Deleted: (COV)

Commented [so9]: Reviewer 1 AR 5

Deleted: Here

Deleted: is examined

Commented [so10]: Reviewer 2 AR 5

Deleted: in the same timeframe

Commented [so11]:

Commented [so12]: Reviewer 1 AR 3

Deleted: s

Deleted: r

Deleted: p

Deleted: COV

Commented [so13]: Reviewer 1 AR 5

Commented [so14]: Reviewer 1 AR 6

Commented [so15]: Reviewer 1 AR 3

Deleted: significant

Deleted: COV

560 (Rodell et al., 2004). CLM 2.0 (Bonan et al., 2002; Dai et al., 2003) with 3.4m depth and 10 vertical
layers features the most well developed soil model (Scanlon et al., 2018), has been shown to
perform well in comparative testing (Scanlon et al., 2018; Spennemann et al., 2014). In addition,
CLM has demonstrated appropriate variability in initial ensemble model runs undertaken here,
meaning that Δ SMS is almost always less than the magnitude of Δ TWS thereby ensuring that Δ GWS
565 estimates derived from Eq. (3) are not arbitrarily high or low (Shamsudduha and Taylor, 2019).
Therefore, this study employs a single model, CLM, for Δ SMS and Δ SWS rather than adopting a LSM
ensemble mean approach.

2.2. Climatology:

570

Individual aquifer system shapefiles from the WHYMAP LASW were prepared as ASCII files and
uploaded to KNMI Climate Explorer (KNMI Climate Explorer, 2018). This allowed a range of climate
data to be extracted for the precise spatial boundaries of each system. In particular, **P**recipitation
(PCP) data from the CRU TS4.03 dataset at 0.5° resolution (Climate Research Unit, University of East
575 Anglia, 2019) was obtained together with **A**nomalies (PCPA) normalised for the SRP (2002-16). The
CRU TS4.03 datasets together with the Δ GWS derived from JPL-Mascon Δ TWS and CLM 2.0 Δ SMS &
 Δ SWS, in accordance with Eq. (3), were used to create time series analyses to explore correlations
over different time and volume components through integration. In this respect the use of 'annual'
in this study implies the appropriate hydrological year.

580

In order to calibrate the time series for each aquifer system prior to further analysis, the lag
between monthly PCP, as the primary climate-groundwater index, and monthly GRACE Δ TWS was
set by maximising the Pearson Correlation Coefficient (PCC) between the two datasets, validated by
point-wise verification of alignment of the time series. In the majority of cases, this comparison
showed Δ TWS lagging PCP by two months. The lag for the PCPA time series were set in the same
585 way with relation to Δ GWS but with the already determined PCP time series lag set as a minimum. In
the case of all aquifer systems except for the Congo, Canning and Indus River Basins, this procedure
resulted in a consistent lag being applied to all of the time series investigations of each aquifer.

590

Initial investigations also established that only relatively weak first-order correlations exist between
 Δ TWS and other monthly observational climate data such as the self-calibrating Palmer Drought
Severity Index (PDSI-sc) (Wells et al., 2004) and Mean Temperature anomalies (CPC GHCN/CAMS
t2m analysis) (Fan and van den Dool, 2008). **B**y comparison **w**ith both these measures, it appeared
that PCPA carried a stronger climate variability signal due to the tropical/sub-tropical location of the
selected aquifers (Allan et al., 2010; Shepherd, 2014). An analysis was then conducted to test for

Commented [so16]: Reviewer 1 AR 3

Deleted: p

Deleted: a

correlations between ΔGWS and a series of measures of precipitation. Three separate time series of precipitation were developed to examine the temporal response of the study LASW with respect to the process by which precipitation at the land surface contributes to ΔGWS :

- 600
1. PCP = monthly precipitation
 2. PCPA = monthly precipitation anomalies with respect to the consistently applied study reference period time-mean baseline, 2002 - 2016
 3. JPCPA = cumulative monthly rainfall anomalies derived by integrating the PCPA time series

Commented [so17]: Reviewer 2 AR 6

Deleted: for the study reference period (SRP – 2002-16)

605 These monthly series were also summed to provide annual time series for each aquifer system. Correlation was measured using the PCC with statistical significance determined by a t-test with $\alpha=0.05$ (Spearman, 1904). In addition, as previously stated, the CGIAR-CSI Global-Aridity dataset (Trabucco and Zomer, 2019) was obtained and a numerical AI for each aquifer was extracted as a spatial mean value using QGIS. AI was used to place each aquifer into the climate zone classification specified by the dataset as set out in Table 1. Of the climate zones relating to the 14 aquifer systems, 610 3 are Arid, 5 are Semi-Arid, 1 is Dry Semi-Humid, giving 9 in total in dryland zones (Corvalán et al., 2005), and 5 are Humid.

2.3. Hydraulic Memory (HM):

615 In using cumulative rainfall anomalies, this study invokes the concept of system memory (Weber and Stewart, 2004). Several studies have considered the question of hydraulic or hydrologic memory, both as it impacts soil moisture including land/atmosphere dynamics (Castro et al., 2009; Lo and Famiglietti, 2010; Wu et al., 2002), and groundwater (Currell et al., 2016; Cuthbert et al., 2019a; 620 Güntner et al., 2007; Rodell and Famiglietti, 2001). Central to the definition of this 'memory' is that it represents the time taken for a system to re-equilibrate following a change in boundary conditions (Downing et al., 1974). In the case of an aquifer system, approximated to a one-dimensional flow of uniform diffusivity, the groundwater response time (GRT) is given by Eq. (4):

625
$$GRT = L^2 S / \beta T \quad (4)$$

where L is a measure of the scale of the system, S is the storativity, β is a dimensionless constant and T is transmissivity. Qualitatively Eq. (4) implies that long response times are characterised by large-scale systems and/or low hydraulic diffusivity (i.e. combination of high S and low T) (e.g. Kooi and

Groen, 2003). An alternative approach to quantifying memory may be needed in more complex - and realistic – multidimensional flow situations (see Cuthbert et al., 2019a). Nevertheless, Eq. (4) still provides a useful order of magnitude approximation. Here, it is helpful to consider the response time as a delay between system input and system output whereby the output state H , at time t , is given by:

$$H(t) = \int_{-\infty}^t p(\tau) \theta(t - \tau) d\tau \quad (5)$$

where $p(\tau)$ is the input state or function at time τ , $(t-\tau)$ is the delay between output and input, and θ is an Impulse Response Function (IRF), also known as a transfer function (Long and Mahler, 2013). The IRF is a multi-parameter function that is intended to model the properties of the system so that the output of the IRF determines the time, t , at which the state H is reached. The hydraulic memory is quantified by the length of time that the effect of the input persists in the system. As the IRF is commonly exponential, making the equilibrium state asymptotic, system memory can be defined as the time interval at which the IRF is 95% complete. This approach has been successfully applied to modelling aquifer responses to precipitation validated by piezometry in both the USA (Long and Mahler, 2013) and the Netherlands (von Asmuth and Knotters, 2004). Alternatively, system memory may be defined as the length of time taken for the effect of the anomalous input to decay to $1/e$ of its starting value where this can be explicitly measured (Cuthbert et al., 2019a; Lo and Famiglietti, 2010). In relation to Eq. (5), chosen precipitation measures are $p(\tau)$ input functions, and ΔGWS represents $H(t)$, the output measure. The timestep, τ , for each of the precipitation time series used is as shown in Table 2. Correlation between ΔGWS (output) and a particular precipitation dataset (input) can be considered to be a measure of the persistence of the effect of that input integrated over the timestep. The degree of correlation between ΔGWS and annual JPCPA is thus indicative of the duration of HM in the aquifer system.

Time series:	Timestep τ:
<u>PCP & PCPA</u>	<u>1 month</u>
<u>PCPA (HY)</u>	<u>1 year</u>
<u>JPCPA (HY)</u>	<u>1 year $\leq \tau \leq 14$ years (upper limit set by length of dataset)</u>

Table 2: The timestep, τ , for each of the precipitation time series investigated in the study

660

2.4. Regional-Scale Hydrogeology:

In an exploration of climate-groundwater dynamics using GRACE data, the lack of direct physical
665 observational data means that it is necessary to demonstrate that results are not simply artefacts of
modelling and signal processing (Rodell et al., 2009). The role of hydrogeology in determining
groundwater dynamics is widely acknowledged (Befus et al., 2017; Cuthbert et al., 2019a; de Vries
and Simmers, 2002; Lanen et al., 2013). Here, we seek to validate results inferred from GRACE data
670 with reference to the physical characteristics of specific aquifer systems. In order to categorise the
hydrogeology of each aquifer system, a number of available global datasets were sourced as raster
files and interrogated in QGIS using the aquifer vector files from WHYMAP LASW. Examined datasets
include:

1. Groundwater Response Time (GRT) (Cuthbert et al., 2019a)
- 675 2. Hydraulic Conductivity (K) and Porosity (Φ) GHLYMPS high resolution maps (Gleeson et al.,
2014)
3. Water Table Depth (WTD) (Fan et al., 2013)

As defined above, the GRT is a temporal measure of the latency of aquifer systems that is derived
680 from their scale and physical properties via [Eq. \(4\)](#). This measure relies on the other datasets listed
for its calculation (Cuthbert et al., 2019a). K and Φ are high-resolution datasets derived from
recently developed lithological maps of the Earth's surface (Hartmann and Moosdorf, 2012) and
their computation uses established geological parameters (Gleeson et al., 2014). However, K is
based on permeability mapping from hydrolithologies that have a standard deviation of ~ 2 orders of
685 magnitude (Gleeson et al., 2011) [and this variance underlies the uncertainty in each of these
datasets used](#). WTD is a 30 arc-second (~ 1 km.) resolution dataset compiled from available
observational data extended by modelled interpolation [with both of these data sources being
subject to considerable sampling bias and model uncertainty respectively](#) (Fan et al., 2013). All of
these datasets are global and derived from combinations of observations and modelled data.

690

3.0. Results:

Commented [so18]: Reviewer 2 AR 7

695 The main results for each aquifer system are given as a monthly time series of Δ TWS and Δ GWS vs. PCP and an annual time series of Δ GWS vs. PCPA and JPCPA, shown as **Figure 2 a-r** for dryland systems and **Figure 3 a-j** for humid systems. The outcomes are summarised in **Table 3**. As a general result, all time series plots show a qualitative relationship between Δ GWS and PCP that exhibits interesting and potentially important spatio-temporal variations. The quantitative results show that

700 for Δ GWS there is a strong correlation with annual JPCPA for aquifer systems in dryland environments whereas in humid environments, the strongest correlation is with monthly PCP. Three aquifers – Guarani Aquifer, Indus River Basin and Canning Basin - do not follow this general classification and anomalies are discussed further in section 4.3 below, and in the SI.

705

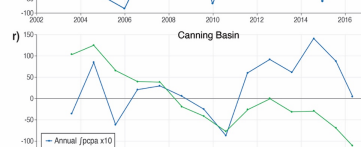
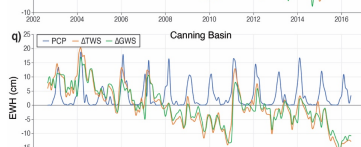
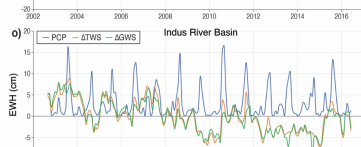
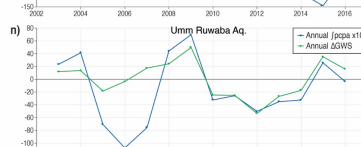
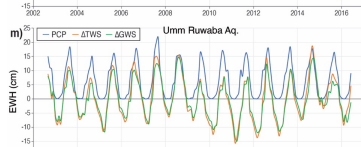
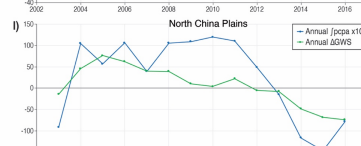
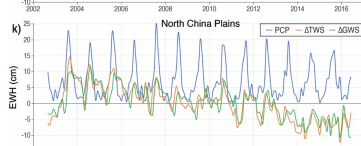
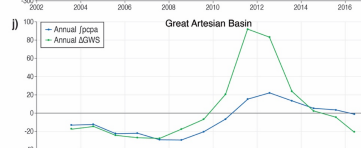
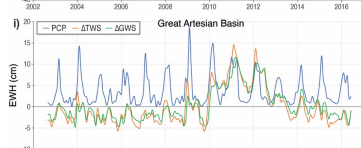
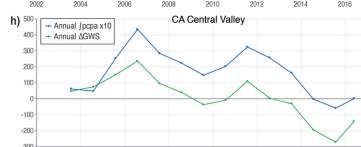
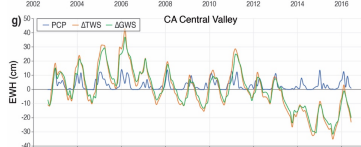
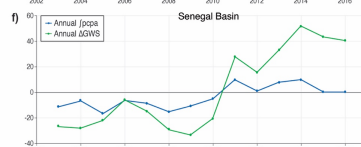
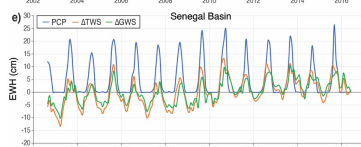
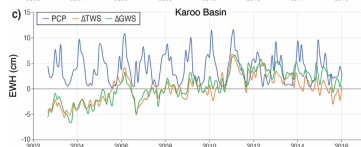
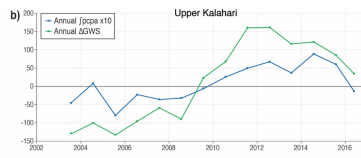
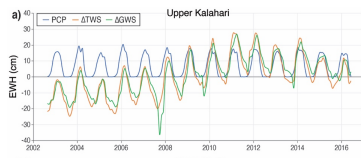
Aquifer System	Monthly PCP vs Δ TWS	Monthly PCP vs Δ GWS	Monthly PCPA vs Δ GWS	Annual PCPA vs Δ GWS	Monthly JPCPA vs Δ GWS	Annual JPCPA vs Δ GWS	Aridity Class	Aridity Index	GWS Net Change over SRP
Upper Kalahari	0.64 (2)	0.47 (2)	0.13 (2)	0.22 (2)	0.67 (2)	0.88 (2)	Semi-Arid	0.42	Increasing
Karoo	0.15 (7)	0.25 (7)	0.07 (7)	0.21 (7)	0.71 (7)	0.88 (7)	Semi-Arid	0.28	Increasing
Senegal	0.67 (2)	0.55 (2)	0.15 (2)	0.14 (2)	0.61 (2)	0.87 (2)	Semi-Arid	0.20	Increasing
California Central Valley	0.53 (2)	0.46 (2)	0.26 (2)	0.56 (2)	0.60 (2)	0.84 (2)	Semi-Arid	0.22	Decreasing
Great Artesian	0.45 (2)	0.33 (2)	0.34 (2)	0.67 (2)	0.61 (2)	0.80 (2)	Arid	0.18	Stable
North China Plains	0.34 (2)	0.22 (2)	0.18 (2)	0.26 (2)	0.65 (2)	0.80 (2)	Dry Sub-Humid	0.57	Decreasing
Umm Ruwaba	0.87 (2)	0.83 (2)	0.12 (2)	0.55 (2)	0.20 (2)	0.64 (2)	Semi-Arid	0.33	Stable
Congo	0.67 (2)	0.67 (2)	0.11 (3)	0.43 (3)	0.27 (3)	0.62 (3)	Humid	1.22	Stable
Maranhao	0.82 (2)	0.75 (2)	0.30 (2)	0.74 (2)	0.11 (2)	0.40 (2)	Humid	0.91	Decreasing
Indus River	0.30 (1)	0.11 (1)	0.19 (3)	0.37 (3)	0.15 (3)	0.34 (3)	Arid	0.16	Decreasing
Amazon	0.88 (2)	0.82 (2)	0.08 (2)	-0.12 (2)	0.13 (2)	0.33 (2)	Humid	1.99	Stable
Guarani	0.50 (3)	0.48 (3)	0.42 (3)	0.78 (3)	0.01 (3)	0.26 (3)	Humid	0.90	Increasing
Ganges-Brahmaputra	0.75 (2)	0.69 (2)	0.06 (2)	0.03 (2)	0.03 (2)	0.01 (2)	Humid	0.86	Decreasing
Canning	0.35 (2)	0.19 (2)	0.15 (3)	0.26 (3)	-0.15 (3)	-0.01 (3)	Arid	0.13	Decreasing
Indus River post '08	0.42 (1)	0.15 (1)	0.21 (3)	0.73 (3)	0.34 (3)	0.89 (3)	Arid	0.16	Decreasing
Canning post '06	0.41 (2)	0.24 (2)	0.22 (3)	0.61 (3)	-0.02 (3)	0.24 (3)	Arid	0.13	Decreasing

Table 3: Summary Table of Results from Monthly & Annual Time Series & Aridity Datasets.

Summary of all correlation results from time series datasets [Pearson Correlation Coefficient & (lag in months)] and the aridity indices derived from the CGIAR-CSI Global-Aridity dataset (Trabucco and Zomer, 2019). Δ GWS trend over SRP also shown. Results in italics fall below the t-test threshold. Aquifers are ranked in order of Pearson Correlation Coefficient for Annual [PCPA vs Δ GWS]. For each

710

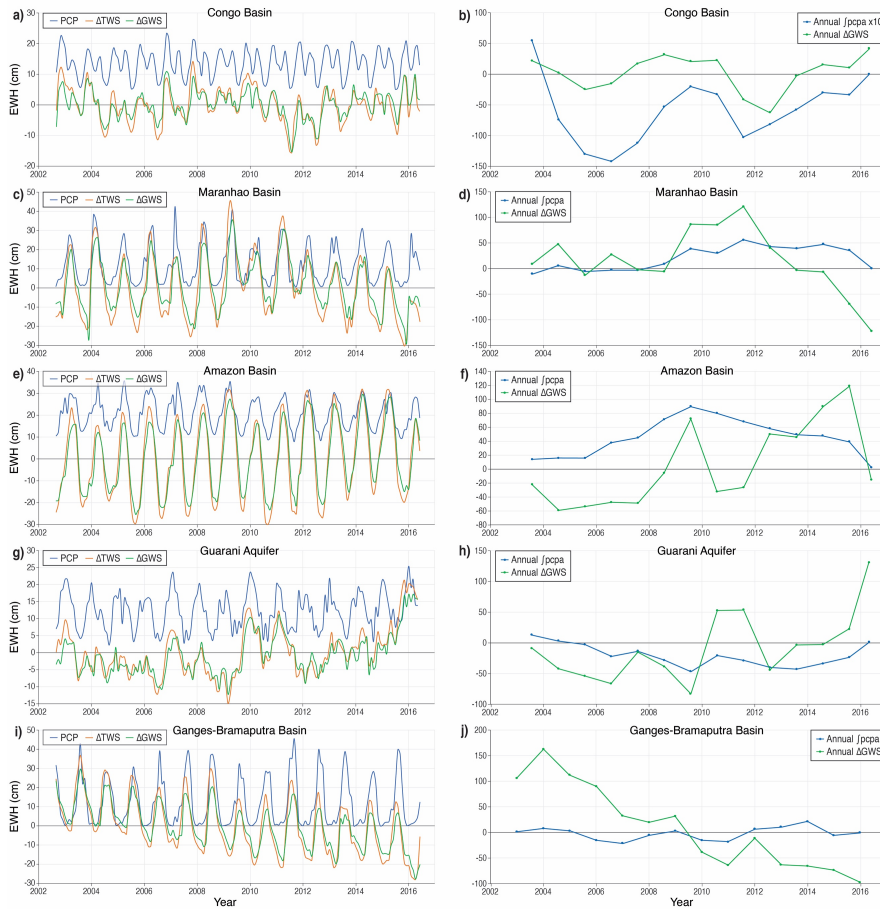
Aquifer system the strongest Δ GWS correlation with PCP or PCPA is shown in bold. Truncated time series results shown for 2 systems.



Formatted: Font: 10 pt, Bold

720

Figure 2: Monthly Δ TWS & Δ GWs vs PCP and Annual Δ GWs vs [PCPA Time Series for each of the dryland climate zone aquifer systems, as labelled. Systems are ordered by decreasing PCC for annual Δ GWs vs [PCPA. All time series are plotted to the aquifer system lag as set out in Table 3, where Δ TWS (Δ GWs) lags PCP (PCPA) by the specified number of months. Y-axis units are Equivalent Water Height (EWH) in cm. Note the variation in the y-axis scales. 7 of the annual [PCPA data series have been scaled x10 for clarity, where indicated.



725

Figure 3: Monthly Δ TWS & Δ GWs vs PCP and Annual Δ GWs vs [PCPA Time Series for each of the humid climate zone aquifer systems, as labelled. Systems are ordered by decreasing PCC for annual Δ GWs vs [PCPA. All time series are plotted to the aquifer system lag as set out in Table 3, where Δ TWS lags PCP by the specified number of months. Y-axis units are Equivalent Water Height (EWH) in cm. Note the variation in the y-axis scales. Congo Basin annual [PCPA data series has been scaled x10 for clarity.

730

GRT, shown in **Figure 4**, is a measure of the time it takes for an aquifer system to equilibrate after a change in boundary conditions, as discussed above. For the 14 studied aquifers, it extends from centennial to millennial timescales as indicated from median values reported in **Tables 4 and S1**. For humid aquifers, GRT ranges from 100 to 350 years whereas for dryland systems GRT then escalates to values well in excess of 1,000 years for semi-arid and arid basins; [the sub-humid](#) North China Plains Aquifer has a GRT of ~550 years. This order of magnitude point of transition can be identified as the threshold between sensitive (rapid) and insensitive (slow) aquifer response times (Cuthbert et al., 2019a), which show a broad global relationship with aridity. This observation helps to explain groundwater storage responses to climate variability through the memory of the aquifer system defined by both physical characteristics and geographical location. The role of HM is discussed further in section 4.1.

Aquifer System	Aridity Classification	Aridity Index	Annual [PCPA vs ΔGWS [PCC] (lag in months)]	GRT: log (GRT) (GRT in yrs)
Indus River post '08	Arid	0.16	0.89 (3)	3.96
Upper Kalahari	Semi-Arid	0.42	0.88 (2)	2.95
Karoo	Semi-Arid	0.28	0.88 (7)	5.74
Senegal	Semi-Arid	0.20	0.87 (2)	5.70
California Central Valley	Semi-Arid	0.22	0.84 (2)	3.01
Great Artesian	Arid	0.18	0.80 (2)	6.33
North China Plains	Dry Sub-Humid	0.57	0.80 (2)	2.74
Umm Ruwaba	Semi-Arid	0.33	0.64 (2)	4.42
Congo	Humid	1.22	0.62 (3)	2.12
Maranhao	Humid	0.91	0.40 (2)	2.55
Indus River	Arid	0.16	0.34 (3)	3.96
Amazon	Humid	1.99	0.33 (2)	2.03
Guarani	Humid	0.90	0.26 (3)	2.20
Ganges-Brahmaputra	Humid	0.86	0.01 (2)	2.10
Canning	Arid	0.13	-0.01 (3)	6.46

Table 4: Relationship between Aridity Index, Climate and Regional-Scale Hydrogeology: Data linking climate and regional-scale hydrogeology to GW dynamics. (Italicised results fall below t-test threshold.)

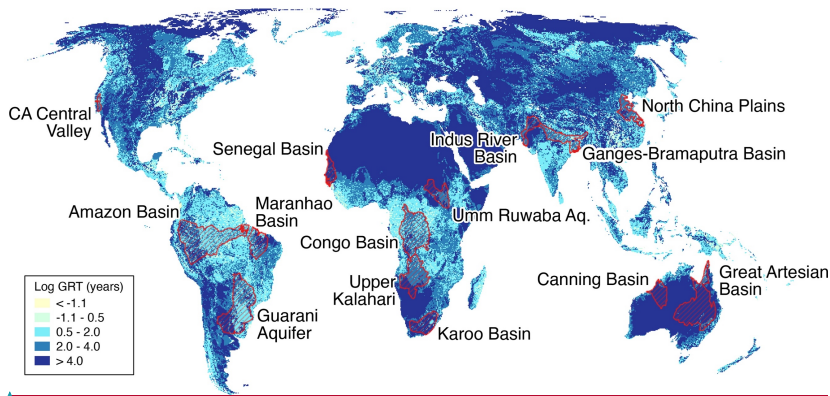


Figure 4: 14 of the World's Large-scale Aquifers (the Study Aquifers) overlaid on the GRT dataset [original dataset from: (Cuthbert et al., 2019a)]

Formatted: Font: Bold

755

Presented results represent the outcome of a detailed analysis of the available datasets and, as such, contain important assumptions that need to be acknowledged here. Firstly, the allocation of lag time has been done on a 'best-fit to the Δ TWS data' basis. It is therefore not derived from analysis of intrinsic physical characteristics of the aquifer systems but is consistent with the range of theoretical values derived from hydrodynamic first principles that anticipate a maximum lag time of 3 months for systems with a large GRT (Townley, 1995), as has been observed by Ahmed et al. (2011). Time lags have been tested for consistency through alignment of specific events in the various time series (Storch and Zwiers, 2001). The evident anomaly of a 7-month lag time for the Karoo Basin is discussed in the SI. Secondly, the restricted duration of the GRACE dataset should be acknowledged, particularly with regard to the annual time series. In mitigation, statistical significance appears to be robust when tested **using the methods described by Zwiers and Von Storch** (Zwiers and von Storch, 1995) and the use of PCPA and JPCPA datasets is designed to minimise the effect of seasonal climate and short-term trends in Δ GWs (Craddock, 1965). Thirdly, the use of **Eq. (3)** to derive Δ GWs from GRACE Δ TWS data represents a temporal and spatial approximation in representing sub-surface hydrological processes. Simply put, all water below the soil zone neither necessarily comprises GWS nor will it all eventually reach GWS due to lateral flow processes. However, on the scale of the aquifer systems considered here, the use of **Eq. (3)** is a reasonable approximation (de Vries and Simmers, 2002).

760

765

770

775

Commented [so19]: Reviewer 1 R8

4.0. Discussion:

780

4.1. Role of Hydraulic Memory (HM):

A key finding of this study is that GRACE-derived Δ GWS correlates most strongly with annual JPCPA for large-scale aquifers in dryland environments of the tropics and sub-tropics whereas GRACE-derived Δ GWS correlates most strongly with monthly PCP in humid environments at these latitudes. Further, we show that there is correspondence between the annual JPCPA vs Δ GWS correlations and GRTs of large-scale aquifer systems (Table 4); the latter is a measure derived in accordance with Eq. (4) (Cuthbert et al., 2019a). HM ultimately derives from the physical properties of the saturated portion of the aquifer system (Townley, 1995) and system memory as measured by Eq. (5) is representative of the physical properties of an aquifer system and its climate. Von Asmuth and Knotters (2004) use 4 parameters to describe groundwater dynamics in their transfer function (θ in Eq. (5)) that they argue represents a more accurate description of the physical system than previously used parametric methods (von Asmuth and Knotters, 2004). Further, their description of groundwater dynamics is capable of accommodating non-stationary elements such as climate change and groundwater abstraction (von Asmuth and Knotters, 2004). HM as measured by Eq. (5) is therefore representative of both spatial and temporal variability in aquifer systems but HM itself can vary spatio-temporally. Indeed the response time to a given boundary change can vary according to the physical circumstances, with persistence lasting from months to hundreds of thousands of years (Cuthbert et al., 2019a).

800 In this study, the GRACE dataset is not long enough to allow detailed IRF modelling of aquifer systems based on Δ GWS data, which would require an observational record longer than the system memory (Long and Mahler, 2013). An extended GRACE series together with reduced uncertainty in the permeability dataset from which GRT is derived, may generate closer numerical matches between GRT (Eq. (4)) and HM as measured by the method of this study (Eq. (5)).
805 Nevertheless, we show that aquifer responses to anomalous precipitation, discussed below, exhibit long HM in dryland environments and relatively short HM in humid environments. The correspondence with GRT extends the classification to two broad categories: dryland environment/long HM/slow GRT and humid environment/short HM/rapid GRT. Note that these categories represent a simplification of the correspondence between HM derived from the study
810 datasets and GRT, which in fact exhibits a spectrum in which Umm Ruwaba (dryland), Congo Basin

Deleted: Long and Mahler (2013), for example, used a total of 16 metrics to describe particular North American Karst aquifer systems in their IRF (θ in (5)). In contrast

Deleted: , v

Deleted: d

Commented [so20]: Reviewer 2 AR 8

and Maranhao (both humid) occupy an intermediate position in terms of the correlation between Δ GWS and annual JPCPA, as can be seen from **Table 3**. Aquifers in humid environments, with exception of the Congo Basin, generally exhibit less HM in this study than expected from GRT values. These humid aquifers, as can be seen from **Figure 4**, have some of their area with GRTs in the order of years to tens of years, perhaps meaning that a disproportionate amount of groundwater processes may be moving through these lower GRT areas. This may explain why humid regions have less HM overall than is implied by their median GRT.

825 **4.2 Aquifer Responses to Anomalous Precipitation:**

The annual time series of Δ GWS vs JPCPA for each aquifer have been examined to identify years in which the maximum annual increase in Δ GWS occurred, as identified by the steepest positive gradient of the Δ GWS line (**Table 5**). These years of extreme recharge, inferred from the increase in Δ GWS, are then further categorised by whether: (1) prior to the event JPCPA is negative, indicating anomalously dry conditions when Soil Moisture Deficits (SMDs) are likely to be widespread; and (2) the JPCPA is concurrently shifting from a negative to positive cumulative anomaly, associated with an extreme rainfall event. Finally, the NINO3.4 index for 2002-2016 (Huang et al., 2015) has been examined (KNMI Climate Explorer, 2018) to indicate the state of ENSO, the dominant control on equatorial precipitation, at the time of the recharge. Nearly all recharge events in dryland aquifer systems take place at a time of negative JPCPA (likely SMD), with most coinciding with extreme rainfall as recently observed in a pan-African study by Cuthbert et al., (2019b). Extreme recharge events also generally coincide with El Niño/La Niña events indicating an association with large-scale modes of climate variability identified previously in tropical Africa (Kolusu et al., 2019; Taylor et al., 2013b). In contrast, extreme recharge in humid aquifer systems is consistently associated with neither negative JPCPA (likely SMD), nor anomalous rainfall, though the latter is correlated with ENSO state.

Commented [so21]: Reviewer 1 AR 3

Deleted: s

Deleted: m

Deleted: d

855

Aquifer Systems grouped by AI: Dry	Year of Extreme Recharge	Negative [PCPA (likely SMD)]Y/N]	[PCPA Phase Change]Y/N]	ENSO State
Senegal	2010	Y	Y	La Nina
Umm Ruwaba	2014	Y	Y	Neutral
U. Kalahari	2008/9	Y	N	La Nina
Karoo	2010/11	N	N	La Nina
California CV	2015/16	Y	Y	El Nino
Indus River	2003	Y	Y	El Nino
Indus River	2015	Y	Y	El Nino
Great Artesian	2010/11	Y	Y	La Nina
Canning	2010/11	Y	Y	La Nina
North China Plains	2003	Y	Y	El Nino
Aquifer Systems by AI: Humid				
Ganges	2003	N	N	El Nino
Ganges	2011	Y	Y	Neutral
Amazon	2008/9	N	N	La Nina
Amazon	2011/12	N	N	La Nina
Maranhao	2008/9	N	N	La Nina
Guarani	2009/10	Y	N	El Nino
Guarani	2015/16	Y	Y	El Nino
Congo	2012/13	Y	N	Neutral

Table 5: Aquifer systems grouped by AI – Dry (Upper) and Humid (Lower). Extreme recharge years identified from annual time series by slope of Δ GWs plotted line. SMD status inferred by prior negative [PCPA and annual [PCPA phase change also derived from the same time series. ENSO state from NINO3.4 Index (Huang et al., 2015).

860

4.3 Anomalous Trends in Groundwater Storage:

865

Over the SRP determined by the availability of GRACE data, six aquifer systems show a net decline in groundwater storage: California Central Valley, North China Plains, Maranhao, Ganges, Indus & Canning Basins. Of these, two aquifer systems (Indus River and Canning Basins) do not show a strong correlation between Δ GWs and any of the precipitation data series. **Table 4** shows that these same two aquifers do not fit the general classification of the 14 aquifer systems into either

870 dryland/slow GRT/long HM or humid/rapid GRT/short HM systems. These anomalous characteristics
may reflect groundwater storage decline through the escalation of groundwater abstraction
referenced previously (Wada et al., 2014) [and this hypothesis was tested through further analysis as
follows below and in further detail in the SI.](#)

875 The Indus River and Canning Basins superficially present similar stories of groundwater
storage decline yet contextual analysis of their respective GRACE/CLM Δ GWS datasets reveals two
quite different realities. The Indus River Basin supports a population of ~210 million people
(Immerzeel et al., 2010) and its hydrology is strongly influenced by water supply from upstream of
the basin, much of it intended for irrigation (Immerzeel et al., 2010). Surface water is augmented by
groundwater abstraction, which supplies ~31% of the total irrigation demand, but it has been
880 estimated that ~84% of the groundwater abstracted returns to the aquifer system as leakage from
canals and intensively irrigated fields (Cheema et al., 2014). A net calculation of these effects on
 Δ GWS, which is detailed in the SI, shows that the underlying climate-groundwater dynamics are
consistent with the GRT derived from the regional-scale hydrogeology of the aquifer system. In
contrast, the Canning Basin is sparsely populated and is not a centre of agriculture (Richey et al.,
885 2015). It is, however, a source of freshwater for iron-ore extraction in adjacent areas (Western
Australia Department of Water, 2011) [and](#) very little of the abstracted groundwater is returned to
the aquifer system as its use in mining causes it to become contaminated (Western Australia
Department of Water, 2013). [This](#) contaminated groundwater is [subsequently](#) disposed in the sea or
evaporation ponds (Prosser et al., 2011). The Canning Basin has a very slow GRT and, situated in an
890 arid environment, is subject to low rates of groundwater recharge so that the physically sustainable
rate of groundwater abstraction is expected to be very low (Scanlon et al., 2006). The analysis of the
Indus and Canning Basins is evidence of [how](#) groundwater depletion, [which has been](#) reported
elsewhere (e.g. Famiglietti, 2014; Rodell et al., 2009), impacts relationships between precipitation
and Δ GWS.

895 **5.0 Conclusions:**

Strong correlations are found between GRACE-derived annual Δ GWS and JPCPA for large-scale
aquifer systems in dryland environments. This correlation is much weaker for large-scale aquifer
900 systems in humid zones where a stronger correlation generally exists between monthly Δ GWS and
monthly PCP. We propose that the correlation between annual Δ GWS and JPCPA demonstrates the
existence of hydraulic memory which is central to large-scale climate-groundwater dynamics. For the
studied aquifer systems, the measure of correlation between annual Δ GWS and JPCPA also shows

905 very good correspondence with the groundwater response time, a measure of the hydraulic memory
of an aquifer system derived from its regional-scale hydrogeological and catchment properties
(Cuthbert et al., 2019a). The 14 aquifer systems can be broadly categorised into two groups, with
each group listed in ascending order of groundwater response time:

- 910 • Group 1: Dryland/Long HM/slow GRT: North China Plains, Upper Kalahari, California Central
Valley, Indus River, Umm Ruwaba, Senegal-Mauritanian, Karoo, Great Artesian & [Canning]
Basins
- Group 2: Humid/Short HM/rapid GRT: Amazon, Ganges, Congo, Guarani, & Maranhao Basins

915 Aquifer systems in Group 1 may be less sensitive to seasonal climate variability but also
vulnerable to long-term trends from which they will be slow to recover. In contrast, aquifers in
Group 2 may be more sensitive to seasonal climate disturbances such as ENSO-related drought but
may also be relatively quick to recover. These characteristics can be applied to anticipate the
groundwater response to present conditions and to future pressures that can be expected from
anthropogenic climate change (Taylor et al., 2013a). The results from the analysis of GRACE data are
920 reconciled to regional-scale hydrogeological conditions, which gives confidence in their validity
(Beven and Germann, 2013), albeit with the caveat regarding the uncertainties inherent in all the
datasets used (Wilks, 2016).

The new GRACE follow on (GRACE-FO) project has now been launched (Frappart and
Ramillien, 2018; Tapley et al., 2019), providing an opportunity to augment the existing GRACE Δ TWS
925 dataset without recourse to modelling (Ahmed et al., 2019) and to give greater certainty in linking
climate-groundwater dynamics to decadal and longer timescale climate systems including the Pacific
Decadal Oscillation and Atlantic Multidecadal Oscillation (Wunsch, 1999). An extended dataset will
improve the calibration of HM as it relates to specific aquifer systems, providing a robust context for
monitoring Δ GWS, including groundwater decline, in real time and protecting fundamentally
930 important groundwater resources.

Competing Interests: The authors declare no competing interests.

935 *Authors' Contributions: SO led the analysis of datasets originally compiled by MS and MC,
supplemented by datasets developed by SO; RT, CB and MS contributed to the original design of
the study with key modifications made by SO; SO drafted the manuscript with input from RT; all
authors contributed to, and commented on, revisions to the submitted manuscript.*

940 **References:**

- 945 Acker, J.G., Leptoukh, G., 2007. Online analysis enhances use of NASA Earth science data. Eos, Transactions American Geophysical Union 88, 14–17. <https://doi.org/10.1029/2007EO020003>
- 950 Ahmed, M., Sultan, M., Elbayoumi, T., Tissot, P., 2019. Forecasting GRACE Data over the African Watersheds Using Artificial Neural Networks. Remote Sensing 11, 1769. <https://doi.org/10.3390/rs11151769>
- 955 Allan, R.P., Soden, B.J., John, V.O., Ingram, W., Good, P., 2010. Current changes in tropical precipitation. Environ. Res. Lett. 5, 025205. <https://doi.org/10.1088/1748-9326/5/2/025205>
- 960 Alley, W.M., Healy, R.W., LaBaugh, J.W., Reilly, T.E., 2002. Flow and Storage in Groundwater Systems. Science 296, 1985–1990. <https://doi.org/10.1126/science.1067123>
- 965 Becker, M., Llovel, W., Cazenave, A., Güntner, A., Crétaux, J.-F., 2010. Recent hydrological behavior of the East African great lakes region inferred from GRACE, satellite altimetry and rainfall observations. Comptes Rendus Geoscience 342, 223–233. <https://doi.org/10.1016/j.crte.2009.12.010>
- 970 Befus, K.M., Jasechko, S., Luijendijk, E., Gleeson, T., Bayani Cardenas, M., 2017. The rapid yet uneven turnover of Earth’s groundwater. Geophysical Research Letters 44, 5511–5520. <https://doi.org/10.1002/2017GL073322>
- 975 Beven, K., Germann, P., 2013. Macropores and water flow in soils revisited: REVIEW. Water Resources Research 49, 3071–3092. <https://doi.org/10.1002/wrcr.20156>
- 980 BGR - WHYMAP - Large Aquifers [WWW Document], 2008. URL https://www.whymap.org/whymap/EN/Maps_Data/Additional_maps/whymap_largeaquifers_g.html (accessed 8.19.19).
- 985 Bierkens, M.F.P., 2015. Global hydrology 2015: State, trends, and directions: Global Hydrology 2015. Water Resources Research 51, 4923–4947. <https://doi.org/10.1002/2015WR017173>
- Bonan, G.B., Oleson, K.W., Vertenstein, M., Levis, S., Zeng, X., Dai, Y., Dickinson, R.E., Yang, Z.-L., 2002. The Land Surface Climatology of the Community Land Model Coupled to the NCAR Community Climate Model. J. Climate 15, 3123–3149. [https://doi.org/10.1175/1520-0442\(2002\)015<3123:TLSCOT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<3123:TLSCOT>2.0.CO;2)
- Bonsor, H.C., Shamsudduha, M., Marchant, B.P., MacDonald, A.M., Taylor, R.G., 2018. Seasonal and Decadal Groundwater Changes in African Sedimentary Aquifers Estimated Using GRACE Products and LSMs. Remote Sensing 10, 904. <https://doi.org/10.3390/rs10060904>
- Castro, C.L., Beltrán-Przekurat, A.B., Pielke, R.A., 2009. Spatiotemporal Variability of Precipitation, Modeled Soil Moisture, and Vegetation Greenness in North America within the Recent Observational Record. J. Hydrometeor. 10, 1355–1378. <https://doi.org/10.1175/2009JHM1123.1>
- Cheema, M. j. m., Immerzeel, W. w., Bastiaanssen, W. g. m., 2014. Spatial Quantification of Groundwater Abstraction in the Irrigated Indus Basin. Groundwater 52, 25–36. <https://doi.org/10.1111/gwat.12027>
- Chen, J.L., Wilson, C.R., Tapley, B.D., 2010. The 2009 exceptional Amazon flood and interannual terrestrial water storage change observed by GRACE. Water Resources Research 46. <https://doi.org/10.1029/2010WR009383>
- Chen, J.L., Wilson, C.R., Tapley, B.D., Scanlon, B., Güntner, A., 2016. Long-term groundwater storage change in Victoria, Australia from satellite gravity and in situ

- 990 [observations. *Global and Planetary Change* 139, 56–65.](#)
<https://doi.org/10.1016/j.gloplacha.2016.01.002>
- Climate Research Unit, University of East Anglia, 2019. CRU TS Version 4.01 [WWW Document]. CRU TS Version 4.01. URL https://crudata.uea.ac.uk/cru/data/hrg/cru_ts_4.01/ (accessed 8.20.19).
- 995 [Corvalán, C., Hales, S., McMichael, A.J., Millennium Ecosystem Assessment \(Program\), World Health Organization \(Eds.\), 2005. *Ecosystems and human well-being: health synthesis, Millennium ecosystem assessment*. World Health Organization, Geneva, Switzerland.](#)
- 1000 [Craddock, J.M., 1965. The Analysis of Meteorological Time Series for Use in Forecasting. *Journal of the Royal Statistical Society. Series D \(The Statistician\)* 15, 167–190.](#)
<https://doi.org/10.2307/2987390>
- [Currell, M., Gleeson, T., Dahlhaus, P., 2016. A New Assessment Framework for Transience in Hydrogeological Systems. *Groundwater* 54, 4–14.](#)
<https://doi.org/10.1111/gwat.12300>
- 1005 [Cuthbert, M.O., Gleeson, T., Moosdorf, N., Befus, K.M., Schneider, A., Hartmann, J., Lehner, B., 2019a. Global patterns and dynamics of climate–groundwater interactions. *Nature Climate Change* 1. <https://doi.org/10.1038/s41558-018-0386-4>](#)
- [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., Kukuric, N., 2019b. 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>
- 1015 [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., Yang, Z.-L., 2003. The Common Land Model. *Bull. Amer. Meteor. Soc.* 84, 1013–1024.](#)
<https://doi.org/10.1175/BAMS-84-8-1013>
- 1020 [Damkjaer, S., Taylor, R., 2017. The measurement of water scarcity: Defining a meaningful indicator. *Ambio* 46, 513–531. <https://doi.org/10.1007/s13280-017-0912-z>](#)
- [de Vries, J.J., Simmers, I., 2002. Groundwater recharge: an overview of processes and challenges. *Hydrogeology Journal* 10, 5–17. <https://doi.org/10.1007/s10040-001-0171-7>](#)
- 1025 [Döll, P., Douville, H., Güntner, A., Müller Schmied, H., Wada, Y., 2016. Modelling Freshwater Resources at the Global Scale: Challenges and Prospects. *Surv Geophys* 37, 195–221. <https://doi.org/10.1007/s10712-015-9343-1>](#)
- [Döll, P., Hoffmann-Dobrev, H., Portmann, F.T., Siebert, S., Eicker, A., Rodell, M., Strassberg, G., Scanlon, B.R., 2012. Impact of water withdrawals from groundwater and surface water on continental water storage variations. *Journal of Geodynamics, Mass Transport and Mass Distribution in the System Earth* 59–60, 143–156.](#)
<https://doi.org/10.1016/j.jog.2011.05.001>
- 1030 [Döll, P., Müller Schmied, H., Schuh, C., Portmann, F.T., Eicker, A., 2014. Global-scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites. *Water Resources Research* 50, 5698–5720.](#)
<https://doi.org/10.1002/2014WR015595>

- 1040 [Downing, R.A., Oakes, D.B., Wilkinson, W.B., Wright, C.E., 1974. Regional development of groundwater resources in combination with surface water. *Journal of Hydrology* 22, 155–177. \[https://doi.org/10.1016/0022-1694\\(74\\)90102-4\]\(https://doi.org/10.1016/0022-1694\(74\)90102-4\)](#)
- [Famiglietti, J.S., 2014. The global groundwater crisis. *Nature Climate Change* 4, 945–948. <https://doi.org/10.1038/nclimate2425>](#)
- [Fan, Y., Li, H., Miguez-Macho, G., 2013. Global Patterns of Groundwater Table Depth. *Science* 339, 940–943. <https://doi.org/10.1126/science.1229881>](#)
- 1045 [Fan, Y., van den Dool, H., 2008. A global monthly land surface air temperature analysis for 1948–present. *Journal of Geophysical Research: Atmospheres* 113. <https://doi.org/10.1029/2007JD008470>](#)
- [Frappart, F., Ramillien, G., 2018. Monitoring Groundwater Storage Changes Using the Gravity Recovery and Climate Experiment \(GRACE\) Satellite Mission: A Review. *Remote Sensing* 10, 829. <https://doi.org/10.3390/rs10060829>](#)
- 1050 [Getirana, A., Kumar, S., Girotto, M., Rodell, M., 2017. Rivers and Floodplains as Key Components of Global Terrestrial Water Storage Variability. *Geophysical Research Letters* 44, 10,359–10,368. <https://doi.org/10.1002/2017GL074684>](#)
- [Gleeson, T., Moosdorf, N., Hartmann, J., van Beek, L.P.H., 2014. A glimpse beneath earth's surface: GLobal HYdrogeology MaPS \(GLHYMPS\) of permeability and porosity. *Geophysical Research Letters* 41, 3891–3898. <https://doi.org/10.1002/2014GL059856>](#)
- 1055 [Gleeson, T., Smith, L., Moosdorf, N., Hartmann, J., Dürr, H.H., Manning, A.H., Beek, L.P.H., van, Jellinek, A.M., 2011. Mapping permeability over the surface of the Earth. *Geophysical Research Letters* 38. <https://doi.org/10.1029/2010GL045565>](#)
- [Graaf, I.E.M. de, Gleeson, T., Beek, L.P.H. \(Rens\) van, Sutanudjaja, E.H., Bierkens, M.F.P., 2019. Environmental flow limits to global groundwater pumping. *Nature* 574, 90–94. <https://doi.org/10.1038/s41586-019-1594-4>](#)
- 1060 [Güntner, A., Stuck, J., Werth, S., Döll, P., Verzano, K., Merz, B., 2007. A global analysis of temporal and spatial variations in continental water storage. *Water Resources Research* 43. <https://doi.org/10.1029/2006WR005247>](#)
- 1065 [Hartmann, J., Moosdorf, N., 2012. The new global lithological map database GLiM: A representation of rock properties at the Earth surface. *Geochemistry, Geophysics, Geosystems* 13. <https://doi.org/10.1029/2012GC004370>](#)
- [Henry, C.M., Allen, D.M., Huang, J., 2011. Groundwater storage variability and annual recharge using well-hydrograph and GRACE satellite data. *Hydrogeol J* 19, 741–755. <https://doi.org/10.1007/s10040-011-0724-3>](#)
- 1070 [Huang, B., Banzon, V.F., Freeman, E., Lawrimore, J., Liu, W., Peterson, T.C., Smith, T.M., Thorne, P.W., Woodruff, S.D., Zhang, H.-M., 2015. Extended Reconstructed Sea Surface Temperature \(ERSST\), Version 4. <https://doi.org/10.7289/v5kd1vvf>](#)
- 1075 [Huang, Z., Pan, Y., Gong, H., Yeh, P.J.-F., Li, X., Zhou, D., Zhao, W., 2015. Subregional-scale groundwater depletion detected by GRACE for both shallow and deep aquifers in North China Plain. *Geophysical Research Letters* 42, 1791–1799. <https://doi.org/10.1002/2014GL062498>](#)
- [Immerzeel, W.W., van Beek, L.P.H., Bierkens, M.F.P., 2010. Climate Change Will Affect the Asian Water Towers. *Science* 328, 1382–1385. <https://doi.org/10.1126/science.1183188>](#)
- 1080 [JPL NASA, 2019. Frequently Asked Questions | About \[WWW Document\]. GRACE Tellus. URL <https://grace.jpl.nasa.gov/about/faq> \(accessed 8.20.19\).](#)
- 1085 [KNMI Climate Explorer, 2018. Climate Explorer: Starting point \[WWW Document\]. KNMI Climate Explorer. URL <https://climexp.knmi.nl/start.cgi?id=b08c094a879f19247ae5839cc6377977> \(accessed 12.22.18\).](#)

- 1090 Kolusu, S.R., Shamsudduha, M., Todd, M.C., Taylor, R.G., Seddon, D., Kashaigili, J.J., Ebrahim, G.Y., Cuthbert, M.O., Sorensen, J.P.R., Villholth, K.G., MacDonald, A.M., MacLeod, D.A., 2019. The El Niño event of 2015–2016: climate anomalies and their impact on groundwater resources in East and Southern Africa. *Hydrology and Earth System Sciences* 23, 1751–1762. <https://doi.org/10.5194/hess-23-1751-2019>
- 1095 Kooi, H., Groen, J., 2003. Geological processes and the management of groundwater resources in coastal areas. *Netherlands Journal of Geosciences* 82, 31–40. <https://doi.org/10.1017/S0016774600022770>
- Landerer, F.W., 2019. Personal correspondence [WWW Document]. email. URL <https://outlook.office.com/owa/?realm=acl.ac.uk&path=/attachmentlightbox> (accessed 6.14.19).
- Landerer, F.W., Swenson, S.C., 2012. Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research* 48. <https://doi.org/10.1029/2011WR011453>
- 1100 Lanen, H.A.J. van, Wanders, N., Tallaksen, L.M., Loon, A.F. van, 2013. Hydrological drought across the world: impact of climate and physical catchment structure [WWW Document]. *Hydrology and Earth System Sciences*. <http://dx.doi.org/10.5194/hess-17-1715-2013>
- 1105 Lo, M.-H., Famiglietti, J.S., 2010. Effect of water table dynamics on land surface hydrologic memory. *Journal of Geophysical Research: Atmospheres* 115. <https://doi.org/10.1029/2010JD014191>
- Long, A.J., Mahler, B.J., 2013. Prediction, time variance, and classification of hydraulic response to recharge in two karst aquifers. *Hydrology and Earth System Sciences* 17, 281–294. <https://doi.org/10.5194/hess-17-281-2013>
- 1110 Long, D., Longuevergne, L., Scanlon, B.R., 2015. Global analysis of approaches for deriving total water storage changes from GRACE satellites. *Water Resources Research* 51, 2574–2594. <https://doi.org/10.1002/2014WR016853>
- 1115 Mémin, A., Flament, T., Alizier, B., Watson, C., Rémy, F., 2015. Interannual variation of the Antarctic Ice Sheet from a combined analysis of satellite gravimetry and altimetry data. *Earth and Planetary Science Letters C*, 150–156. <https://doi.org/10.1016/j.epsl.2015.03.045>
- Ni, S., Chen, J., Wilson, C.R., Li, J., Hu, X., Fu, R., 2018. Global Terrestrial Water Storage Changes and Connections to ENSO Events. *Surv Geophys* 39, 1–22. <https://doi.org/10.1007/s10712-017-9421-7>
- 1120 Overgaard, J., Rosbjerg, D., Butts, M.B., 2006. Land-surface modelling in hydrological perspective - a review. *Biogeosciences* 3, 229–241.
- Phillips, T., Nerem, R.S., Fox-Kemper, B., Famiglietti, J.S., Rajagopalan, B., 2012. The influence of ENSO on global terrestrial water storage using GRACE. *Geophysical Research Letters* 39, n/a-n/a. <https://doi.org/10.1029/2012GL052495>
- 1125 Prosser, I., Wolf, L., Littleboy, A., 2011. Water in mining and industry 12.
- Ramillien, G., Frappart, F., Seoane, L., 2014. Application of the Regional Water Mass Variations from GRACE Satellite Gravimetry to Large-Scale Water Management in Africa. *Remote Sensing* 6, 7379–7405. <https://doi.org/10.3390/rs6087379>
- 1130 Richey, A.S., Thomas, B.F., Lo, M.-H., Reager, J.T., Famiglietti, J.S., Voss, K., Swenson, S., Rodell, M., 2015. Quantifying renewable groundwater stress with GRACE. *Water Resources Research* 51, 5217–5238. <https://doi.org/10.1002/2015WR017349>
- 1135 Rodell, M., Famiglietti, J.S., 2001. An analysis of terrestrial water storage variations in Illinois with implications for the Gravity Recovery and Climate Experiment (GRACE). *Water Resources Research* 37, 1327–1339. <https://doi.org/10.1029/2000WR900306>

1140 [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., Toll, D., 2004. The Global Land Data Assimilation System. *Bull. Amer. Meteor. Soc.* 85, 381–394. <https://doi.org/10.1175/BAMS-85-3-381>](#)

1145 [Rodell, M., Velicogna, I., Famiglietti, J.S., 2009. Satellite-based estimates of groundwater depletion in India. *Nature* 460, 999–1002. <https://doi.org/10.1038/nature08238>](#)

[Rui, H., Beaudoin, H., 2019. Global Land Data Assimilation System \(GLDAS-1\) Products README 32.](#)

[Sakumura, C., Bettadpur, S., Bruinsma, S., 2014. Ensemble prediction and intercomparison analysis of GRACE time-variable gravity field models. *Geophysical Research Letters* 41, 1389–1397. <https://doi.org/10.1002/2013GL058632>](#)

[Scanlon, B.R., Healy, R.W., Cook, P.G., 2002. Choosing appropriate techniques for quantifying groundwater recharge. *Hydrogeology Journal* 10, 18–39. <https://doi.org/10.1007/s10040-001-0176-2>](#)

1150 [Scanlon, B.R., Keese, K.E., Flint, A.L., Flint, L.E., Gaye, C.B., Edmunds, W.M., Simmers, I., 2006. Global synthesis of groundwater recharge in semiarid and arid regions. *Hydrological Processes* 20, 3335–3370. <https://doi.org/10.1002/hyp.6335>](#)

1155 [Scanlon, B.R., Zhang, Z., Save, H., Sun, A.Y., Schmied, H.M., Beek, L.P.H. van, Wiese, D.N., Wada, Y., Long, D., Reedy, R.C., Longuevergne, L., Döll, P., Bierkens, M.F.P., 2018. Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. *PNAS* 115, E1080–E1089. <https://doi.org/10.1073/pnas.1704665115>](#)

1160 [Scanlon, B.R., Zhang, Z., Save, H., Wiese, D.N., Landerer, F.W., Long, D., Longuevergne, L., Chen, J., 2016. Global evaluation of new GRACE mascon products for hydrologic applications: GLOBAL ANALYSIS OF GRACE MASCON PRODUCTS. *Water Resources Research* 52, 9412–9429. <https://doi.org/10.1002/2016WR019494>](#)

[Shamsudduha, M., Taylor, R.G., 2019. Changes in global groundwater storage from GRACE: uncertainty and the role of extreme precipitation 85.](#)

1165 [Shamsudduha, M., Taylor, R.G., Jones, D., Longuevergne, L., Owor, M., Tindimugaya, C., 2017. Recent changes in terrestrial water storage in the Upper Nile Basin: an evaluation of commonly used gridded GRACE products.](#)

[Shamsudduha, M., Taylor, R.G., Longuevergne, L., 2012. Monitoring groundwater storage changes in the highly seasonal humid tropics: Validation of GRACE measurements in the Bengal Basin. *Water Resources Research* 48. <https://doi.org/10.1029/2011WR010993>](#)

1170 [Shepherd, T.G., 2014. Atmospheric circulation as a source of uncertainty in climate change projections. *Nature Geoscience* 7, 703–708. <https://doi.org/10.1038/ngeo2253>](#)

[Sood, A., Smakhtin, V., 2015. Global hydrological models: a review. *Hydrological Sciences Journal* 60, 549–565. <https://doi.org/10.1080/02626667.2014.950580>](#)

1175 [Spearman, C., 1904. The Proof and Measurement of Association between Two Things. *The American Journal of Psychology* 15, 72–101. <https://doi.org/10.2307/1412159>](#)

[Spennemann, P.C., Rivera, J.A., Saulo, A.C., Penalba, O.C., 2014. A Comparison of GLDAS Soil Moisture Anomalies against Standardized Precipitation Index and Multisatellite Estimations over South America. *J. Hydrometeor.* 16, 158–171. <https://doi.org/10.1175/JHM-D-13-0190.1>](#)

1180 [Storch, H. von, Zwiers, F.W., 2001. *Statistical Analysis in Climate Research*. Cambridge University Press.](#)

1185 [Tapley, B.D., Bettadpur, S., Watkins, M., Reigber, C., 2004. The gravity recovery and climate experiment: Mission overview and early results. *Geophysical Research Letters*. \[https://doi.org/10.1029/2004GL019920@10.1002/\\(ISSN\\)1944-8007.GRL40\]\(https://doi.org/10.1029/2004GL019920@10.1002/\(ISSN\)1944-8007.GRL40\)](#)

- 1190 [Tapley, B.D., Watkins, M.M., Flechtner, F., Reigber, C., Bettadpur, S., Rodell, M., Sasgen, I., Famiglietti, J.S., Landerer, F.W., Chambers, D.P., Reager, J.T., Gardner, A.S., Save, H., Ivins, E.R., Swenson, S.C., Boening, C., Dahle, C., Wiese, D.N., Dobslaw, H., Tamisiea, M.E., Velicogna, I., 2019. Contributions of GRACE to understanding climate change. *Nature Climate Change* 9, 358. <https://doi.org/10.1038/s41558-019-0456-2>](#)
- 1195 [Taylor, R.G., Koussis, A.D., Tindimugaya, C., 2009. Groundwater and climate in Africa—a review. *Hydrological Sciences Journal* 54, 655–664. <https://doi.org/10.1623/hysj.54.4.655>](#)
- 1200 [Taylor, R.G., Scanlon, B., Döll, 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., Treidel, H., 2013a. Ground water and climate change. *Nature Climate Change* 3, 322–329. <https://doi.org/10.1038/nclimate1744>](#)
- 1205 [Taylor, R.G., Todd, M.C., Kongola, L., Maurice, L., Nahozya, E., Sanga, H., MacDonald, A.M., 2013b. Evidence of the dependence of groundwater resources on extreme rainfall in East Africa. *Nature Climate Change* 3, 374–378. <https://doi.org/10.1038/nclimate1731>](#)
- 1210 [Thomas, B.F., Caineta, J., Nanteza, J., 2017. Global Assessment of Groundwater Sustainability Based On Storage Anomalies. *Geophysical Research Letters* 44, 11,445–11,455. <https://doi.org/10.1002/2017GL076005>](#)
- 1215 [Tiwari, V.M., Wahr, J., Swenson, S., 2009. Dwindling groundwater resources in northern India, from satellite gravity observations. *Geophysical Research Letters* 36. <https://doi.org/10.1029/2009GL039401>](#)
- 1220 [Townley, L.R., 1995. The response of aquifers to periodic forcing. *Advances in Water Resources* 18, 125–146. \[https://doi.org/10.1016/0309-1708\\(95\\)00008-7\]\(https://doi.org/10.1016/0309-1708\(95\)00008-7\)](#)
- 1225 [Trabucco, A., Zomer, R., 2019. Global Aridity Index and Potential Evapotranspiration \(ET0\) Climate Database v2. <https://doi.org/10.6084/m9.figshare.7504448.v3>](#)
- 1230 [Trenberth, K.E., 2011. Changes in precipitation with climate change. *Climate Research* 47, 123–138. <https://doi.org/10.3354/cr00953>](#)
- [von Asmuth, J.R., Knotters, M., 2004. Characterising groundwater dynamics based on a system identification approach. *Journal of Hydrology* 296, 118–134. <https://doi.org/10.1016/j.jhydrol.2004.03.015>](#)
- [Wada, Y., 2016. Modeling Groundwater Depletion at Regional and Global Scales: Present State and Future Prospects. *Surveys in Geophysics* 37, 419–451. <https://doi.org/10.1007/s10712-015-9347-x>](#)
- [Wada, Y., Wisser, D., Bierkens, M.F.P., 2014. Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources 27.](#)
- [Wahr, J., Molenaar, M., Bryan, F., 1998. Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. *Journal of Geophysical Research: Solid Earth* 103, 30205–30229. <https://doi.org/10.1029/98JB02844>](#)
- [Watkins, M.M., Wiese, D.N., Yuan, D.-N., Boening, C., Landerer, F.W., 2015. Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *Journal of Geophysical Research: Solid Earth* 120, 2648–2671. <https://doi.org/10.1002/2014JB011547>](#)
- [Weber, K., Stewart, M., 2004. A Critical Analysis of the Cumulative Rainfall Departure Concept. *Ground Water; Dublin* 42, 935–938.](#)

- 1235 [Wells, N., Goddard, S., Hayes, M.J., 2004. A Self-Calibrating Palmer Drought Severity Index. *J. Climate* 17, 2335–2351. \[https://doi.org/10.1175/1520-0442\\(2004\\)017<2335:ASPDSI>2.0.CO;2\]\(https://doi.org/10.1175/1520-0442\(2004\)017<2335:ASPDSI>2.0.CO;2\)](#)
- [Western Australia Department of Water, 2013. Western Australian water in mining guideline. Western Australia Department of Water, 2011. Capacity of water resources in the Mid west to meet mining and industrial growth: a status report. Dept. of Water, Perth, W.A.](#)
- 1240 [Wiese, D.N., Landerer, F.W., Watkins, M.M., 2016. Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution. *Water Resources Research* 52, 7490–7502. <https://doi.org/10.1002/2016WR019344>](#)
- [Wilks, D.S., 2016. “The Stippling Shows Statistically Significant Grid Points”: How Research Results are Routinely Overstated and Overinterpreted, and What to Do about It. *Bull. Amer. Meteor. Soc.* 97, 2263–2273. <https://doi.org/10.1175/BAMS-D-15-00267.1>](#)
- [Wood, E.F., Roundy, J.K., Troy, T.J., Beek, L.P.H. van, Bierkens, M.F.P., Blyth, E., Roo, A. de, Döll, P., Ek, M., Famiglietti, J., Gochis, D., Giesen, N. van de, Houser, P., Jaffé, P.R., Kollet, S., Lehner, B., Lettenmaier, D.P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., Whitehead, P., 2011. Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth’s terrestrial water. *Water Resources Research* 47. <https://doi.org/10.1029/2010WR010090>](#)
- 1250 [Wu, W., Geller, M.A., Dickinson, R.E., 2002. The Response of Soil Moisture to Long-Term Variability of Precipitation. *J. Hydrometeor.* 3, 604–613. \[https://doi.org/10.1175/1525-7541\\(2002\\)003<0604:TROSMT>2.0.CO;2\]\(https://doi.org/10.1175/1525-7541\(2002\)003<0604:TROSMT>2.0.CO;2\)](#)
- [Wunsch, C., 1999. The Interpretation of Short Climate Records, with Comments on the North Atlantic and Southern Oscillations. *Bulletin of the American Meteorological Society* 80, 245–255. \[https://doi.org/10.1175/1520-0477\\(1999\\)080<0245:TIOSCR>2.0.CO;2\]\(https://doi.org/10.1175/1520-0477\(1999\)080<0245:TIOSCR>2.0.CO;2\)](#)
- 1260 [Xavier, L., Becker, M., Cazenave, A., Longuevergne, L., Llovel, W., Filho, O.C.R., 2010. Interannual variability in water storage over 2003–2008 in the Amazon Basin from GRACE space gravimetry, in situ river level and precipitation data. *Remote Sensing of Environment* 114, 1629–1637. <https://doi.org/10.1016/j.rse.2010.02.005>](#)
- [Xie, X., Xu, C., Wen, Y., Li, W., 2018. Monitoring Groundwater Storage Changes in the Loess Plateau Using GRACE Satellite Gravity Data, Hydrological Models and Coal Mining Data. *Remote Sensing* 10, 605. <https://doi.org/10.3390/rs10040605>](#)
- 1265 [Yeh, P.J.-F., Swenson, S.C., Famiglietti, J.S., Rodell, M., 2006. Remote sensing of groundwater storage changes in Illinois using the Gravity Recovery and Climate Experiment \(GRACE\). *Water Resources Research* 42. <https://doi.org/10.1029/2006WR005374>](#)
- 1270 [Zwiers, F.W., von Storch, H., 1995. Taking Serial Correlation into Account in Tests of the Mean. *J. Climate* 8, 336–351. \[https://doi.org/10.1175/1520-0442\\(1995\\)008<0336:TSCIAI>2.0.CO;2\]\(https://doi.org/10.1175/1520-0442\(1995\)008<0336:TSCIAI>2.0.CO;2\)](#)

1275 **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.](#)

1280