How realistic are multi-decadal reconstructions of GRACE-like total water storage anomalies?

Charlotte Hacker[1] and Jürgen Kusche[1]

[1]University of Bonn

June 1, 2023

Abstract

The Gravity Recovery and Climate Experiment (GRACE) mission has monitored total water storage anomalies (TWSA) globally with unprecedented resolution and accuracy since 2002. However, many applications require a data-based, multi-decadal extended record of TWSA prior to the GRACE period and for bridging the eleven-months gap between GRACE and its successor GRACE Follow-On (GRACE-FO), that does not depend on hydrological modelling. Statistical and machine-learning ‘reconstruction’ approaches have been developed to this end, mostly via identifying relations of GRACE-derived TWSA to climate variables, and some regional or global land data sets are now publicly available.

In this contribution, we compare the two global reconstructions by Humphrey and Gudmundsson (2019) and Li et al.(2021) mutually and against output from the water Global Analysis and Prognosis (WaterGAP) hydrological model from 1979 onwards, against large-scale mass-change derived from geodetic satellite laser ranging (SLR) from 1992 onwards, and finally against differing GRACE/-FO solutions from 2002 onwards.

We find that the reconstructions agree surprisingly well in many regions at seasonal and sub-seasonal timescales, even in the pre-GRACE era. We find larger differences at inter-annual timescales which we speculate are in part due to the way reconstructions are trained, and in part on which specific GRACE solution they are trained as well as the climatological characteristics of the region. Our comparison against independent SLR data reveals that reconstructions (only) partially succeed in representing anomalous TWSA for regions that are influenced by large climate modes such as El Ni$\tilde{\text{n}}$o-Southern Oscillation (ENSO).

How realistic are multi-decadal reconstructions of GRACE-like total water storage anomalies?

C. Hacker[1], J.Kusche[1]

[1]Institute for Geodesy and Geoinformation, University of Bonn, Bonn, Germany.

Key Points:

- Analysis of signal components of pre-GRACE (1979-2002) and long-term (1979-2016) terrestrial water storage based on global GRACE like TWSA reconstructions
- The low frequency part of the reconstructions are evaluated by low degree gravity fields from satellite laser ranging
- The reconstructions reveal similar regions affected by water storage changes over the last four decades.

Corresponding author: Charlotte Hacker, hacker@geod.uni-bonn.de

Abstract

[The Gravity Recovery and Climate Experiment (GRACE) mission has monitored total water storage anomalies (TWSA) globally with unprecedented resolution and accuracy since 2002. However, many applications require a data-based, multi-decadal extended record of TWSA prior to the GRACE period and for bridging the eleven-months gap between GRACE and its successor GRACE Follow-On (GRACE-FO), that does not depend on hydrological modelling. Statistical and machine-learning 'reconstruction' approaches have been developed to this end, mostly via identifying relations of GRACE-derived TWSA to climate variables, and some regional or global land data sets are now publicly available.

In this contribution, we compare the two global reconstructions by Humphrey and Gudmundsson (2019) and F. Li et al. (2021) mutually and against output from the water Global Analysis and Prognosis (WaterGAP) hydrological model from 1979 onwards, against large-scale mass-change derived from geodetic satellite laser ranging (SLR) from 1992 onwards, and finally against differing GRACE/-FO solutions from 2002 onwards.

We find that the reconstructions agree surprisingly well in many regions at seasonal and sub-seasonal timescales, even in the pre-GRACE era. We find larger differences at inter-annual timescales which we speculate are in part due to the way reconstructions are trained, and in part on which specific GRACE solution they are trained as well as the climatological characteristics of the region. Our comparison against independent SLR data reveals that reconstructions (only) partially succeed in representing anomalous TWSA for regions that are influenced by large climate modes such as El Niño-Southern Oscillation (ENSO).]

Plain Language Summary

Water is a life sustaining resource, crucial for human survival, agricultural and economical proposes. Since 2002, the Gravity Recovery and Climate Experiment (GRACE) mission monitors total water storage anomalies (TWSA) on a global scale, allowing the analysis of temporal changes in the water cycle. However, the time series is limited to 20 years. Many data analysis applications require a data-based, multi-decadal extended record of TWSA prior to the GRACE period. Reconstructions are directly "build" based on the GRACE observation, finding a relationship between GRACE-derived TWSA and climate variables.

In this contribution, we compare the two global reconstructions by Humphrey and Gudmundsson (2019) and F. Li et al. (2021) mutually and against output from the water Global Analysis and Prognosis (WaterGAP) hydrological model from 1979 onwards, against large-scale mass-change derived from geodetic satellite laser ranging (SLR) from 1992 onwards, and finally against differing GRACE/-FO solutions from 2002 onwards.

We find the reconstructions reveal similar regions affected by water storage changes over the last four decades, especially for basins with strong TWSA signals like the Amazon. Our comparison against independent SLR data reveals that reconstructions (only) partially succeed in representing anomalous TWSA for regions that are influenced by large climate modes such as El Niño-Southern Oscillation (ENSO).

1 Introduction

The GRACE (Gravity Recovery And Climate Experiment) and GRACE-FO (GRACE Follow-On) satellite missions provide time variable gravity field models and estimates of total water storage anomalies (TWSA) with monthly resolution since 2002 (Tapley et al., 2019). GRACE and GRACE-FO (hereafter GRACE-/FO) data have enabled studying the natural variability of the terrestrial water cycle and its response to radiative forcing, land use, and water withdrawal and redirection (Rodell et al., 2018). TWSA maps have been used to quantify groundwater stress (Richey et al., 2015), large-scale droughts

(Zhao et al., 2017; Gerdener et al., 2020) and floods (Reager et al., 2014; Han et al., 2021), vegetation response (Geruo et al., 2017), and soil processes (Swenson & Lawrence, 2014). In general, the long-term temporal evolution of water storage observed by satellites can be linked to modifications of the land boundary conditions and the resulting climate forcing, the direct and indirect impacts of anthropogenic activities such as groundwater abstraction and land use change, and the hydrological response of the system (Eicker et al., 2016), and regions where trends and accelerations share the same sign may be seen as moving away from the long-term equilibrium of the water cycle. In addition, several studies (Zaitchik et al., 2008; Eicker et al., 2014; Tangdamrongsub et al., 2015; Girotto et al., 2016; Khaki et al., 2017; Schumacher et al., 2018; B. Li et al., 2019; Springer et al., 2019; Tangdamrongsub et al., 2020; Gerdener et al., 2020) assimilated GRACE/-FO TWSA data into hydrological and land surface models, to add realism to model simulations. Apart from these climate and hydrology applications, observation-based TWSA maps are increasingly employed in geodesy e.g. for loading computations for satellite altimetry (Ray et al., 2013) and GNSS (Chanard et al., 2018; Mémin et al., 2020), deriving geocenter estimates (Wu & Heflin, 2015), or simulating the hydrological angular momentum contribution to Earth rotation variations (Seoane et al., 2009; Jin et al., 2010).

However, the use of GRACE/-FO data for e.g. deriving changes in the frequency of water storage extremes (Kusche et al., 2016) or for climate model evaluation (Jensen et al., 2019), is severely hampered by the short duration of the time series. Similarly, it is well-known that trend and acceleration estimates derived within the present GRACE/-FO time period are affected by inter-annual variability (Eicker et al., 2016). In addition, there is an eleven-month gap between GRACE and GRACE-FO, and the GRACE record has several shorter gaps due to battery problems on the spacecraft. All this limits not only the use of GRACE/-FO TWSA data for confronting model simulations, but also the assimilation of TWSA data into models, as assimilation frameworks are usually ill-prepared for data sets with gaps. Unfortunately, no geodetic observing technique provides either time-variable gravity fields or any other variable that could be related to global land TWSA in the pre-GRACE era with a spatial resolution comparable to GRACE.

Data-based 'reconstructions' of gridded total water storage seek to provide substitute data and in this way to overcome many of the issues mentioned above, typically by deriving and training a relationship between the monthly GRACE TWSA maps and predictors for which multidecadal data records are available. The mathematical framework to derive a relationship is either based on regression techniques or on machine learning methods. Predictors are mostly chosen as climate variables that play a dominating role in the terrestrial water budget, such as precipitation or land surface temperature, but they can also include other hydrological or space-geodetic observations. Predictors can be sets of single time series (e.g. ENSO index, modes derived via empirical orthogonal function analysis) or spatial fields.

TWSA reconstruction methods can serve, at the same time, for predicting near real-time or even future total water storage variability (Forootan et al., 2014). This may be useful in diagnosing problems e.g. in quick-look data analysis, in identifying real anomalies, or in forecasting e.g. drought or flood conditions at the seasonal timescale (Reager et al., 2014).

Table Appendix B.1 provides an overview of different reconstructions of total water storage. Becker et al. (2011) were upon the first in reconstructing water storage, for the 1980−2008 time frame. In the Amazon basin, gridded TWSA maps are derived based on the most energetic spatial modes seen by GRACE, as identified via from singular value decomposition, and scaled with in-situ water level data from river gauges. This approach is essentially identical to how the sea level community reconstructs past sea surface height maps from a few decades of radar altimetry and long records from tide gauges (Church et al., 2004). Forootan et al. (2020) and Richter et al. (2021) applied similar methods later at the global scale in order to close the eleven-months gap between GRACE and GRACE-FO with data from the Swarm satellites. Both studies reported a notable decrease in the noise level and an improvement in the spatial resolution by combining Swarm and GRACE.

Forootan (2014) developed an approach for predicting water storage anomalies over West Africa based on low-degree autoregressive TWSA modelling driven by external variables, which were selected from independent or principal components of relevant climate fields. Autoregressive modelling was also used to close the gap between the two GRACE missions by Lenczuk et al. (2022). Their process model does not employ any external predictor data sets and thus uncertainty accumulates quickly; to mitigate this Lenczuk et al. (2022) suggest forward- and backward propagation.

Many recent studies turned to machine learning algorithms (see table Appendix B.1). Long et al. (2014) studied floods and droughts in a karst plateau in China using temporally extended GRACE data. They used an artificial neural network (ANN) to learn the relationship between monthly precipitation, mean temperature and soil moisture derived from the Global Land Data Assimilation System (GLDAS) (Rodell et al., 2004). A similar study was carried out by Zhang et al. (2016) for the Yangtze basin in China, where they used an ANN with precipitation and soil moisture inputs to extend the GRACE time series and characterize drought impacts on the Yangtze. A. Y. Sun et al. (2019) applied a deep convolutional neural network (CNN) to learn the relation between GLDAS and GRACE-derived TWSA for India. The focus of their study was laid on the success of the CNN training: CNNs are mostly used for image classification and trained on big data sets, whereas the GRACE data set with around 170 monthly maps (Tapley et al., 2019) is comparably small. Yu et al. (2021) also used deep learning to reconstruct GRACE-like TWSA for the Canadian landmass, while training with modelled TWSA. Apart from deep learning, Z. Sun et al. (2020) used multiple linear regression (MLR) and seasonal autoregressive integrated moving average with exogenous variable (SARIMAX) to predict water storage changes in around 60 global basins. They found good agreement across the three methods for the reconstructed signals for humid and low-intensity irrigated areas. For drier regions their results showed significant variations, indicating that the performance of an algorithm may depend on hydroclimatic characteristics of the basin. In a second study A. Y. Sun et al. (2021) focused on the GRACE– GRACE-FO data gap. They used six different machine learning algorithms and multiple groups of meteorological and climatic variables to fill the time series over conterminous U.S., and suggest to combine different ML models to provide a final robust estimate.

From a user perspective, there is no clear picture yet emerging as to what approach might show advantages and disadvantages at which temporal and spatial scales, and in which climate regimes. Some of the above mentioned studies were designed to fill the relatively short gap between the GRACE and GRACE-FO missions, while others aimed indeed at multidecadal time series. Except for the studies utilizing Swarm data, all the works mentioned above provide regionally restricted reconstructions.

In this contribution, we therefore focus on (in the following abbreviated as HG19) Humphrey and Gudmundsson (2019) and (Li21 in the following) F. Li et al. (2021), which are, to our knowledge, the only two published reconstructions of total water storage anomalies for the entire global land excluding ice caps and glaciated regions. Both HG19 and Li21 cover more than four decades and here we will look at 1979-2020, later on called the full time frame. Both HG19 and Li21 reconstruct TWSA variations from long-term climate variable records via relations trained in the GRACE time period, but they rely on very different mathematical approaches.

HG19 formulated a model with deterministic and stochastic components to describe the inter-annual variability of water storage change over time. While the deterministic part relates storage to precipitation and temperature via a simple 1D (i.e. grid-cell based) first-order decay model, Humphrey and Gudmundsson (2019) fitted a spatial autoregressive noise model (Cressie, 1993) to spatial and temporal correlation structure in the GRACE TWSA maps, in order to quantify the underlying stochastic process. Markov Chain Monte Carlo is used to achieve a representative error distribution, and the reconstruction is derived as an ensemble consisting of 100 realisations.

In a very different setting, the Li21 framework combines machine learning with time series analysis and statistical decomposition techniques. In a first step the dominant (both statistically independent and orthogonal) modes of GRACE and climate data are identified. Selected relevant modes are then decomposed into linear trend, seasonal, interannual part, and subseasonal using different approaches, and each temporal signal component is reconstructed by either simple neural network, autoregressive modelling with exogenous variables or multilinear regression. The global land is divided into basins, and for each basin the optimal combination of the methods is learned (F. Li et al., 2021).

For evaluating these two global multidecadal reconstructions in the pre-GRACE era, we use output from the hydrological model WaterGAP (Müller Schmied et al., 2020a) and, for the first time, low degree time-variable gravity fields from the geodetic satellite laser raging (SLR) technique (Löcher & Kusche, 2020).

WaterGAP simulates water flows and water storage in canopy, snow, soil, groundwater, lakes, man-made reservoirs, wetlands and rivers (Müller Schmied et al., 2020a), with total water storage being defined over the sum of these compartments. The model differs from many global hydrological or land surface models in its representation of anthropogenic processes; its global water use models determine the water use for irrigation, livestock, domestic, manufacturing and thermal power, then required net water abstraction is partitioned into net abstraction from surface water and groundwater. In the context of comparing to TWSA reconstructions trained on GRACE/-FO data, it is worth mentioning that WaterGAP is however limited in its representation of water abstraction, that it uses rather simple algorithms for simulating reservoir operations, and that it does not simulate at all glacier mass variability. The model has, however, quite often been compared to GRACE/-FO data with favorable results in particular at seasonal and sub-seasonal timescales (Müller Schmied et al., 2020a).

SLR contributes to the International Terrestrial Reference Frame (ITRF) as well as low degree gravity field models (Pearlman et al., 2019). The sensitivity of the SLR satellites to the time variable gravity field ranges from the Earth's center-of-mass to about degree ten (M. Cheng et al., 2011). To increase the spectral resolution while sacrificing some independence of the procedure from GRACE, Löcher and Kusche (2020) fitted, in addition to certain low-degree spherical harmonics, a few empirical orthogonal functions of the recent GRACE solutions to the SLR ranging data. This hybrid approach resulted, after back-transformation of the empirical orthogonal functions (EOF), in monthly sets of spherical harmonic coefficients complete up to degree and order 60.

The paper is organised as follows: In section 2 we present the data sets amd methods that will be employed in this study. Section 3 describes our analysis of the two global land water storage reconstructions L21 and HG19 with respect to each other, and when compared to WaterGAP. For each data set the trend, acceleration, annual amplitude and phase, the interannual signal and the subseasonal signal are discussed for the pre-GRACE era (1979-2002) and for the full reconstruction period (1979-2016). In section 4 the two reconstructions are compared to different GRACE solutions, i.e. not only those they were trained on, and after spectral truncation, to the SLR data by Löcher and Kusche (2020) for the years 1992−2002. An examination of the long term evolution of the water storage for nine major river basins closes this section. The findings of this study are briefly summarized in section 5.

2 Data and Methods

2.1 Data

2.1.1 GRACE/FO data

Three different GRACE/-FO level-2 (L2) solutions, the ITSG2018 series (Mayer-Gürr et al., 2018), the release 06 (RL06) from the Center for Space Research of the University of Texas (CSR) from the website CSR (2018) and the RL06 GFZ Potsdam data (Dahle et al., 2018, 2019), and the mass concentration (mascon) solution provided by

CSR (Save et al., 2016) were used to evaluate the reconstructions. All GRACE/-FO data sets cover the time from April 2002 to December 2020. As no interpolation was applied the GRACE/FO data contain gaps, including the nearly one year period between the two missions.

The spherical harmonic solutions were expanded up to degree and order 96. The degree-one and C_{20} coefficients were replaced as recommended in M. Cheng et al. (2011). The coefficients were smoothed with a DDK3 filter (Kusche, 2007). Time-variable signals due to glacial isostatic adjustment (GIA), i.e. the ongoing viscoelastic uplift of the Earth's crust in response to the melting of the former ice sheets, was corrected for using the model by Peltier et al. (2018).

2.1.2 Global reconstructions of GRACE-type gridded total water storage

The F. Li et al. (2021) data set represents a reconstruction of monthly, total water storage anomalies from 1979−2020 on a 0.5° x 0.5° grid. Their framework combine machine learning techniques with time series and statistical decomposition techniques. F. Li et al. (2021) employed the RL06 monthly CSR GRACE solution as predictand and tested several meteorological variables (see table Appendix B.1) as predictors. In a first step, the dominant modes of the input data were identified via either the Principal Component Analysis (PCA) or Independent Component Analysis (ICA) technique. Selected modes were then partitioned into linear trend, seasonal, and inter-annual , and the residual variability signals using either Least-Squares (LS) or seasonal-trend decomposition based on loess (STL) (Cleveland et al., 1990) methods. Each signal component is reconstructed by either artificial neural network (ANN), autoregressive exogenous model (ARX) or multi-linear regression (MLR) approaches. The global land area is divided into hydrological basins, and for each basin the optimal combination of the framework described above was determined via evaluation against observed GRACE/-FO data.

The Humphrey and Gudmundsson (2019) global reconstruction covers the time period from 1979 − 2019 on a 0.5° x 0.5° grid. The reconstruction is build as consisting of a deterministic and a stochastic part. The first one was formulated as a first order decay model, linking the effect of temperature and precipitation on water storage in a simplified way. To identify and quantify a stochastic process, underlying the spatial and temporal correlation structure seen in the original GRACE solutions, Humphrey and Gudmundsson (2019) employed a spatial autoregressive (SAR) model as in Cressie (1993). A Markov chain Monte Carlo procedure was then used to generate representative sample distributions. Two GRACE solutions, three precipitation and two temperature datasets were used to generate six different reconstructions. Each solution consists of an ensemble of 100 realisations. Here, we used the reconstruction based on the JPL masconcs solution and the ERA5 forcing data (Hersbach et al., 2020). We chose this combination since, according to Humphrey and Gudmundsson (2019) it is the closest to the "true" GRACE solution.

2.1.3 Low degree gravity fields from SLR

SLR is commonly used to derive low degree spherical harmonic coefficients, the geocenter position, station coordinates, and thus significantly contributes to the International Terrestrial Reference Frame (Altamimi et al., 2016; M. K. Cheng et al., 2013). Within the post-processing of GRACE and GRACE-FO data sets the C_{20} oblateness coefficient from the level-1 analysis is regularly replaced by estimates based on SLR, and this has been recommended for C_{30} as well (M. Cheng & Ries, 2023)). SLR results have been also frequently used to validate hydrological angular momentum (HAM) estimates derived from hydrological modelling (Śliwińska et al., 2021; W. Chen et al., 2017). The sensitivity of the SLR technique to the time variable gravity field is however limited to spherical harmonic degrees of about $n = 5$ or 6, with certain coefficients of higher degree and order being observable due to the SLR orbital geometry while others cannot be separated (Sośnica et al., 2015).

To increase the spectral resolution of the SLR technique Löcher and Kusche (2020) and recently M. Cheng and Ries (2023) employed a parameterization based on the EOF approach in the SLR data reduction instead of spherical harmonics. In this method, the leading spatial patterns of mass variability, i.e. the EOF functions, were derived from the unfiltered spherical harmonics coefficients from GRACE (based on the ITSG-Grace2018 solution in Löcher and Kusche (2020)) in a preprocessing step.

Within a dynamic orbit improvement procedure, in the hybrid approach some low-degree spherical harmonic coefficients and a set of scaling factors fitting the GRACE EOFs were derived, fitting the original laser range observations. Spherical harmonic coefficients complete to higher degrees are finally derived via re-mapping the scaled EOFs. In this way, Löcher and Kusche (2020) derived monthly gravity fields from $1992-2020$ complete up to spherical harmonic degree $n = 60$. These solutions must be viewed as having inherited spatial constraints from GRACE in the sense that EOFs beyond some threshold in signal power, and thus combinations of spherical harmonics, were truncated and effectively set to zero. However, they include information on mass change in the pre-GRACE era of unprecedented spatial resolution (see validations in Löcher and Kusche (2020)), are publicly available, and we thus use them here to assess the reconstructions described earlier.

2.1.4 The Water Global Analysis and Prognosis (WaterGAP) model

The WaterGAP model (Müller Schmied et al., 2020a) consists of three major components, i.e. the global water use model, the linking model Groundwater-Surface Water Use (GWSWUSE) and the WaterGAP Global Hydrology Model (WGHM). The global water use model determines water use for irrigation, livestock, domestic and manufacturing use, and thermal power. GWSWUSE divides the net water abstraction determined by the global water use model into net abstraction for surface water and groundwater. Net abstraction together with climate forcing then form the input for the WGHM model. WGHM represents water flows and storage in ten compartments; i.e. canopy, snow, soil, groundwater, lakes, man-made reservoirs, wetlands and rivers.

From this we derived monthly total water storage for the years $1901-2016$ by accumulating over all storages. It is worth mentioning here that WaterGAP does not simulate glaciers.

2.2 Methods

For all comparisons in this study, total water storage trend, acceleration, annual amplitude and annual phase are estimated within a least squares approach.

The sub-seasonal signal is defined here as the signal with a period shorter than a year. After removing trend and annual signal from the time series, a high pass filter was used to suppress frequencies below 1 cycle per year. The inter-annual signal is defined as the signal with periods longer than one year. It is obtained here via low pass filtering, after the time series were reduced by a trend, annual and semiannual signal. For the inter-annual and sub-seasonal signals the power is represented via the root mean square (RMS) of the variability.

For the comparison with results from satellite laser raging in section 4, all data sets are expanded into spherical harmonics, truncated at the same degree and order as the SLR data, and expressed as mass change fields. This step is required, as the SLR fields used in this study have a coarse resolution, and we use them only until a spherical harmonic degree of $n = 12$.

We excluded Antarctica and Greenland signals from the reconstructions and WaterGAP, as we feel we do not have reliable modelling data here that could complement our approach. We also exclude those regions from the SLR data to facilitate the comparison.

3 Continental total water storage anomalies in the pre-GRACE era

The primary objective of global, data based reconstructions of total water storage is to better understand the collective response of the land system to long-variability in the rainfall and temperature while remaining independent of hydrological modelling, with applications ranging from space geodesy and sea level budgeting to providing constraints to climate modelling studies. It is important however to understand that reconstructions are by construction principle developed to reproduce or disregard certain temporal scales, and are bound to reproduce events and signatures that are represented in the used predictor variables.

We suggest therefore to distinguish internal and external consistency. In what follows, we will first present a systematic analysis of the HG19 and L21 global land water storage reconstructions with respect to each other, and when compared to the WaterGAP model, which also relies on meteorological forcing fields but in a very different way. For each data set the linear trend, acceleration, annual amplitude and phase, interannual signal, and finally the subseasonal signal are discussed separately for the pre-GRACE era (1979-2002) and for the full reconstruction period (1979-2016). In section 4 following thereafter, the focus will then be on an independent 'external' evaluation with GRACE/-FO and SLR data.

3.1 Continental total water storage anomalies in the years 1979-2016

3.1.1 Humphrey and Gudmundsson (2019) reconstruction based on GRACE data

The first column of figure 1 shows the linear trend, acceleration, average annual amplitude and phase, sub seasonal and inter-annual signal variations for HG19 for the years 1979-2016. The metrics reveal a "chessboard"pattern, that is most likely due to the model formulation of HG19. This reconstruction does not include a trend – Humphrey and Gudmundsson (2019) argue that trends seen by GRACE are mainly driven by anthropogenic effects, which cannot be explained by their statistical model. However, changes in temperature and precipitation include a trend, which can be estimated based on a least square adjustment from the reconstructions. Negative trends are visible for the Congo basin, the region around the Lake Victoria, the Mississippi basin, the Tocantins basin, the Parnaiba basin and the Sao Francisco basin. Positive trends are found for the Orinoco basin, the Zambezi basin and around the Thar desert.

Negative accelerations, i.e. increasing rates of water storage decline or 'accelerated drying' are found for the Mississippi, Colorado and Rio Grande basin in North America, along with the the large basins along South America's Atlantic coast, the Amazon basin, for the Congo basin in Africa, the Volga and Don basin, the Amur basin, large swathes in the eastern part of China, and the Murray-Darling basin in Australia. Positive accelerations, i.e. increased rates of water storage due to rainfall increases, are shown for the Zambezi basin in Africa. Increasing mass rates are also shown for Iceland, consistent with Wu and Heflin (2015), but it is unclear to what extent the HG19 method may be trusted over glacierated regions.

Large long-term average annual cycle of total water storage are found for the Amazon basin, with the dryness peak (i.e. minimum water storage, phase expressed as day-of-year DOY) around February to March. Large annual amplitudes are also found for some coastal basins in South America with dryness peaking at around DOY 250 for the northern part to DOY 120-150 for the southern part. Furthermore, large average annual amplitude occurs in the Mississippi basin, with a dryness peak in May/June, and around the Gulf of Alaska. For the latter, the phase shows values around December to January.

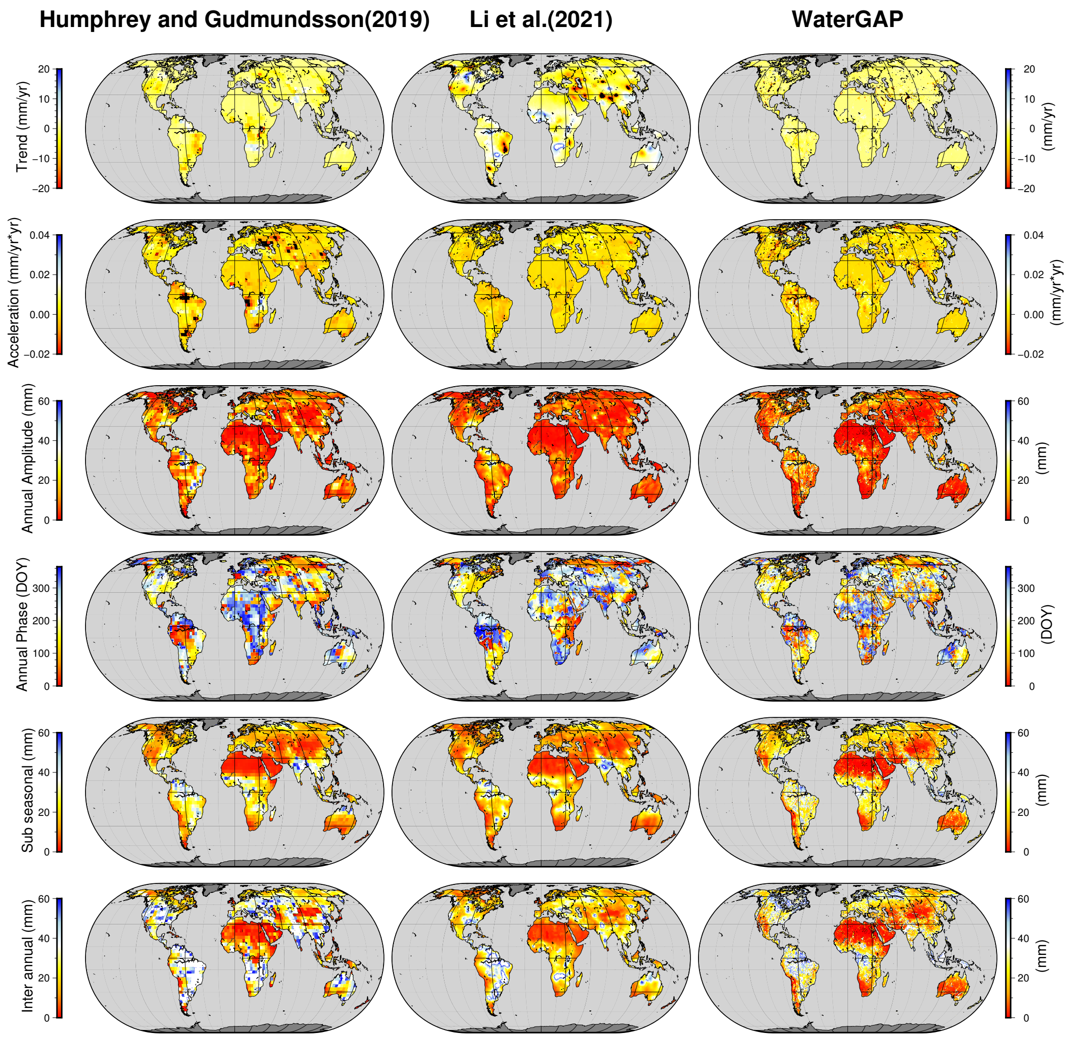

Figure 1. Trend, acceleration, annual amplitude and phase, subseasonal and interannual signal for the years 1979-2016 for the reconstructions and WaterGAP on a global scale

Regions with lower amplitudes (around 2.5 cm) are located in western and southern Europe, around the Caspian Sea, southern Africa, India, southern and eastern China and in the north-west of Australia. For Europe, in terms of peak water deficit, a gradient in magnitude between the north-eastern and south-western part is identified, ranging from DOY 50-160 for the south-western and DOY 250-360 for the north-eastern. For the region around the Caspian Sea, India, southern and eastern China and in north west Australia, minimum peak values for February to March are found, for southern Africa the values are around DOY 260-350. As expected, the annual signals disappear for deserts (Sahara, Arabian Peninsula, Mongolia, Australia) and arid regions (major parts of Canada).

Subseasonal signal power (Fig. 1, first column, fifth row) largely follows the Köppen-Geiger classification (Köppen, 1923; Kottek et al., 2006), with low variations in arid climate zones and large vaiability in tropical areas, e.g. the Amazon and La Plata basins, Niger, Congo and Lake Chad basins, India, southern China, Indonesia and tropical northern Australia. Also regions with polar/snow climate, e.g. northern Canada and Siberia and some associated with mild to warm climate with regular precipitation patterns, like the eastern part of the Mississippi basin and Japan, appear with stronger subseasonal signal variability in the long-term reconstruction.

At the inter-annual scale (bottom right panel) only northern Africa, Arabian Peninsula, Mongolia and south-eastern Australia, i.e. regions with arid climate, exhibit little variability; similar also for the northern hemisphere. Interestingly, large signals of an amplitude comparable to the annual cycle appear to be present (over nearly four decades) for the entire U.S., South America except the Andes, sub-Sahel and southern Africa, the north-eastern coast of Australia, Europe and most parts of Asia.

3.1.2 F. Li et al. (2021) reconstruction based on GRACE data

Even though the Li21 reconstruction does include a trend, this trend has not been reconstructed in the same way as the other signal components; rather it has been separated from the GRACE time series and added back to the finished reconstruction (F. Li et al., 2020, 2021). Therefore, the trends in figures 1 (first row, second column) and 2 (first row, second column) represent mean trends in the used GRACE data. Minor differences in the trend magnitudes can be attributed to the least squares approach used to generate the maps.

Trends in the observed total water storage are in fact due to a mixture of natural variability and anthropogenic effects (Humphrey et al., 2016; Rodell et al., 2018), and since not all human effects can be viewed as a response to natural variability (as e.g. increased water withdrawals in drought periods may be explained) it is unclear to what extent statistical reconstructions trained on such GRACE data sets would represent reality or artefacts. As mentioned above, visible negative trends represent extrapolations from the GRACE period aided by climate data. For the Middle East, they are generally attributed to a superposition of droughts and resulting groundwater depletion for agricultural use (Voss et al., 2013). The decrease in water storage in the Caspian Sea is mostly ascribed to reduced inflow of the Volga river (Loomis & Luthcke, 2017) in the GRACE period.

Accelerated water mass loss can be seen for the Amazon basin, around Victoria Lake and the Caspian Sea. Slightly positive accelerations are found for the Ob basin, around Alaska, the Patagonian glaciers and in Iceland, although for the latter it is not clear whether the Li21 approach is valid.

The Amazon basin, the Zambezi basin (both with a dryness peak around DOY 350) the southern part of South America, the Mississippi basin (both showing minimal phase around April/May), the Caspian Sea region with minimal peaks in the end of Spring, the Victoria lake (with minimal signal magnitudes around January/February) and some ice covered regions, i.e. Iceland, the Patagonian glaciers, Svalbard and the region around the Gulf of Alaska exhibit large amplitudes. For South America, Europe, India, north-eastern China and central Africa moderate annual amplitude of around 1.5−2 cm are shown. The corresponding annual phase, e.g. minimal signal magnitude for South America shows a notable shift between South and North, with minimal water storages around DOY 300-265 for the northern part and DOY 120-180 for the southern part. A similar phase difference can be found for Europe, with minimal signal magnitudes in early Spring / end of Winter for the south-western part and around end of Fall/ begin of Winter for the north-eastern part. North-eastern China and middle Africa both show minimal signal magnitudes for January/ February. The phase for India is around DOY 320 to 260.

Strong sub-seasonal variations are present for the equatorial belt, spanning the Orinoco, Amazon and La Plate basin in South America, the Zambezi, Congo and Niger basin in Africa, Madagascar, India, southern parts of China and the northern part of Australia. The Caspian Sea exhibits also high signal variability, as well as the region around the North China Plain and California, as well as the northern part of the Mississippi basin in North America. Also ice covered regions in Siberia, Canada, around the Gulf of Alaska, around the Patagonian glaciers, Iceland, Svalbard and the Franz-Joseph land depict high signal magnitudes. Most European regions show variability of around $1.5-2$ cm here.

At the inter-annual time scale, the equatorial band becomes even more pronounced, with strong inter-annual signals for most parts of South America, middle Africa, India, southern China and northern Australia likely caused by rainfall variability. Also ice covered regions, like Iceland, the Patagonian glaciers, the Gulf of Alaska, and China's High Mountains glaciers exhibit strong variations. Interestingly, regions where anthropogenic water use prevails, e.g. California, the North China Plain and parts of India exhibit large inter-annual signal magnitudes. The region around the Black Sea, including the Danube, Dnieper, Don, Volga and part of the Ob basin also reveal high signal variability.

3.1.3 WaterGAP Global Hydrological Model simulation

The majority of the trends derived from WaterGAP for $1979-2016$ range between $-0.5cm/yr$ and $0.5cm/yr$ (see figure 1, top left panel). Negative trends can be observed for large parts of America, Australia and Europe, whereas Africa and Asia present a mixed picture of negative and positive trends. Stronger negative trends can be found for the US, over the Arabian Peninsula and India, with patches of stronger trends located all over the world, some of which may be model artefacts. Trends for glacierated regions are not dominant in figure 1 (first row, third column) as WaterGAP does not include glaciers in the simulation (Müller Schmied et al., 2020a).

Trend rates generally tend towards negative accelerations (i.e. increased dryness), with some exceptions for the Congo, the Amazon, the Parana and Ob basin, around the southern region of the Hudson Bay and Indonesia. Notable are negative acceleration pattern over India. The same phenomenon can be observed for the Mississippi basin, where regions showing a strong negative trend also exhibit notable negative acceleration. As mentioned before, the Congo and Ob basin both show partially positive acceleration. The trends for these basins are slightly negative. The opposite sign of trend and acceleration may suggest that long-term reversal from drying or wetting is in progress. Striking is also the negative acceleration in river storage along the main Amazon system, in conjunction with a positive trend for the southern branch, i.e. Madeira and Jurua, and a negative for the northern one, Negro and Branco.

The most prominent annual signal in figure 1 is the Amazon basin. For this area the minimum signal, in terms of water deficit, has been around January to February in the four decades considered here. other regions of higher amplitude in South America can be found for the Orinoco, with minimal phase around DOY 250-300, and Parana basin (lowest storage in April/May). Large annual amplitudes are also found for the Mississippi basin, dryness peaking around DOY 120-160, around Hudson Bay and the Gulf of Alaska. Visible are also large amplitudes for the Congo and the Volga basins, both with minimal signal in February, in the Rhone basin, Norway and parts of Sweden, all showing minimal values around November/December, the Ob basin (March-May), and Indonesia (around DOY 300-350).Moderate annual amplitudes can be found for India, southern China and Central Europe (minmum around DOY 300-360).

In the model simulation, all equatorial regions show high subseasonal signal variations, this includes the majority of South America, Central Africa, India, southern China, Indonesia and the north-eastern coast of Australia. Also regions of polar climate exhibit bigger signal amplitudes. Among the regions with bigger variations is also the Mississippi basin. Europe shows a mixed picture with signal magnitudes around $1.5-2.5$ cm. River storage in WaterGAP appears pronounced in the subseasonal signals.

At inter-annual time scales, regions with an arid warm or cold climate exhibit small signal variations, like the Sahara and the Orange, Okavango and Limpopo basin, the Arabian Peninsula most parts of Australia, Central Asia, the Andes and parts of Mexico. Strong inter-annual signals can be found for the Amazon, Orinoco and La Plata basin and the Patagonian glaciers, the major part of North America, the Congon basin, the northern part of Europe and Siberia, along the Pacific coast of Asia, India, southern China and Indonesia.

3.2 Continental total water storage anomalies for the years 1979-2002

3.2.1 Humphrey and Gudmundsson (2019) reconstruction based on GRACE data

The first panel in the first row of figure 2 shows the trend for HG19 for the years $1979-2002$. HG19 exhibits positive trends for the coastal region of Alaska, the Nelson basin, parts of the Congo basin, the Limpopo basin, over India and the Amazon basin. The latter was not visible for the years 1979-2016. Negative trends are found for the Tocantins basin, the Parnaiba basin and the Sao Francisco basin, the region around the Victoria Lake and the Bandama basin.

Compared to the mean rate changes over the entire time frame (Fig. 1), the storage acceleration in two decades prior to GRACE was larger in magnitude, while also the spatial patterns were apparently different. In HG19, the increasing water storage loss over the US extend to the complete Mississippi basin during this period. Storage anomalies over Ellesmere Island are apparently characterized by positive rate changes, and similar behavior can be found for the Orinoco basin, the East of South America, Africa, Asia and even Europe. Quite notable changes in patterns and magnitude are found for Australia, with increasing wetness trends in the decades prior to GRACE while over the entire time frame (Fig. 1) (and of course over the GRACE period, (van Dijk et al., 2013)) this region experiences long-term drying trends.

The long-term average annual cycle of total water storage over two decades prior to GRACE revealed large amplitudes for the Amazon basin, with a minimum peaking around end of February and March. Large amplitudes were also present along the Pacific coast of North America with minimum storage in around DOY 250, in the Mississippi basin with a minimal peak in May-July and over northern Asia peaking towards the end of February to March. In total, compared to the long term average annual cycle for the entire time period, amplitudes seem bigger for the two decades prior to GRACE. In contrast, the sub-seasonal signal content seems not to depend overly on the time period, and we find the same for the inter-annual signal.

3.2.2 F. Li et al. (2021) reconstruction based on GRACE data

Mean rate changes for the twenty years prior to GRACE are shown in the second column and row of figure 2. Compared to the entire time frame (figure 1, second column and row), we identify a notable change in pattern for Australia – along the southern coasts negative acceleration and in tropical northern regions a positive one. Negative accelerations are also found for the Congo basin, the Parana basin, the Mississippi basin, the Amur basin. The Limpopo basin, the Zambezi basin, the north-western part of Australia,

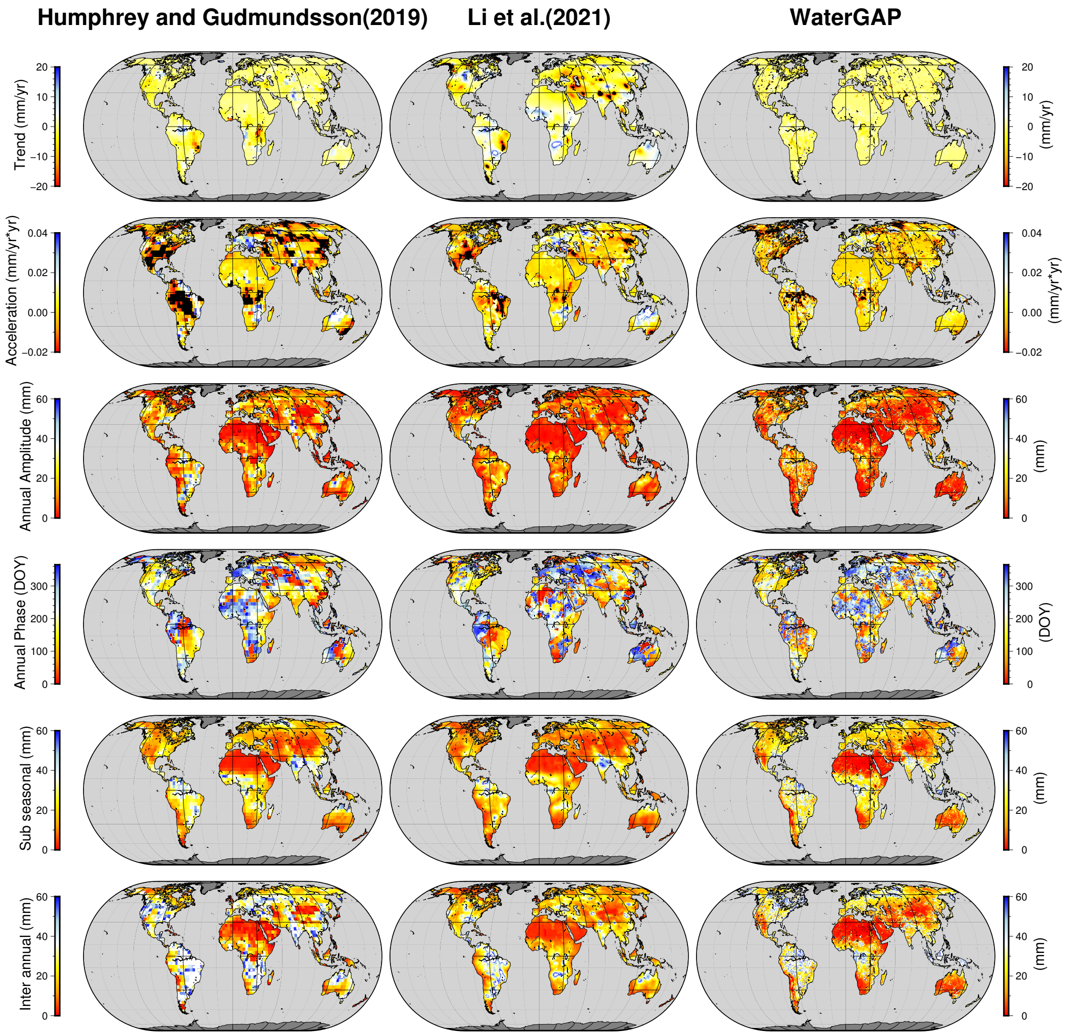

Figure 2. Trend, acceleration, annual amplitude and phase, subseasonal and interannual signal for the years 1979-2002 for the reconstructions and WaterGAP on a global scale

the Danube and Volga basin show positive range rate changes. All mentioned signals are not present in the accelerations derived for the full time frame.

The mean annual amplitude derived for the two decades prior to GRACE shows large annual amplitudes for all coastal basins in South America, with minimal values around January/February and October/November for the upper part of South America and minimal phases in Spring for the lower part. Furthermore large mean annual amplitudes are shown for the southern basins in Africa and the western basins in Siberia. Similar to South America the minimal peak in Africa is in November for the lower part begin of the year in the upper part. For the Northern Asian basins the minimal peak is in early spring, e.g around April/May. In comparison to the mean annual amplitude for the entire time frame the derived amplitude for the twenty years before GRACE has increased in magnitude.

The signal magnitude of the interannual and subseasonal signal is slightly smaller for the two decades prior to GRACE compared to the full time frame. The signal pattern, e.g. regions revealed by the data to experience signal variations, is the same for both time frames.

3.2.3 WaterGAP Global Hydrological Model simulation

Like the trend the mean range rate changes for the two decades prior to GRACE exhibit larger magnitudes compared to the full time frame. Regions with positive mean range rate changes are Indonesia, parts of the Ob basin, Central Europe, parts of the Congo, Zambezi basin in Africa, the Atlantic coast side of South America. The main river branch of the Amazon basin, Alaska, the Mackenzie basin, the Mississippi basin, the Neva basin and northern India reveal negative accelerations. For northern India a negative trend is found, indicating a water depletion in this region.

The average annual signal exhibits large amplitudes in the Amazon basin with minimum water storage ranging from October to March. Other regions with large average annual amplitudes are Indonesia, the region around the Golf of Alaska and Scandinavia. All mentioned regions show minimum storage around September-November. Notable are also the large annual amplitudes for the basins located around the Golf from Mexico (October-March).

3.3 Intercomparisons

In this section, we first briefly discuss general differences between the three data sets. Then, we focus on the Murray-Darling and Amazon basins, regions that are strongly influenced by the El-Niño-Southern Oscillation phenomenon (Nicholls et al., 1997; Trenberth, 1990; Forootan et al., 2016; García-García et al., 2011; Towner et al., 2021; J. L. Chen et al., 2010; Marengo & Espinoza, 2016). We expect the reconstructions to exhibit large storage signals, as both use precipitation in the model formulation. Then we will turn to Europe, where water storage signals (e.g. recent droughts) have large economic impact but are dwarfed by other regions. We suspect Europe to pose more of a challenge to the reconstructions.

Table 2 shows trend (without and with co-estimation of an acceleration term). We find that the trend from Li21 is by more than two times bigger compared to HG19 and WaterGAP. It seams like, WaterGAP (and HG19) underestimate trends in water storage changes compared to Li21. However, as discussed in section 4 the derived trends depend on the used GRACE solution. The trend derived from mascon based GRACE solution, like the one used by Li21 or in the study of Scanlon et al. (2018) exhibit higher water storage signal magnitudes compared to trends derived from spherical harmonic solution.

We note, that the accelerations found for the reconstruction of HG19 are stronger, then those for Li21 or WaterGAP. The acceleration derived from the data sets exhibits higher signal magnitude for the shorter time period. For the full time frame the derived average range rate changes are close to zero.

The data sets reveal similar regions with high average annual amplitudes, like the Amazon basin, the Mississippi basin or the Caspian Sea. Table Appendix C.1 presents values raging from 2.2 cm for HG19 to 2.4 cm for Li21 for the global land. The annual phase indicates similar values in the data sets, except for the Amazon basin, where a time lag of a month is found between HG19 and WaterGAP on the one side and Li21 on the other side.

The subseasonal signal of the two reconstructions exhibits similar signal magnitude and signal pattern and does not change notably for the pre GRACE era compared to the full time series. In comparison, the subseasonal signal of the hydrological model shows higher signal magnitudes, notable especially for North and South America and Europe.

On interannual scale all data sets show small signal variations for arid regions. The signal pattern of the data sets differ for North America, Europe, Asia and Australia. For Asia, Li21 and WaterGAP are in a good agreement, whereas HG19 portrays more region with high signal amplitudes. For Europe, Li21 exhibits smaller signal magnitudes compared to WaterGAP and HG19. For North America, Li21 and HG19 show moderate signal magnitude for the region around the Hudson Bay, whereas WaterGAP indicates signal variations, that are two times larger. The data sets show high signal magnitudes for the Amazon basin and the Orinoco basin. As for the accelerations, HG19 depicts the highest signal magnitudes compared to Li21 and WaterGAP. The signals for the full time period have slightly higher magnitudes compared to the two decades prior to GRACE. The signal pattern does not change.

3.3.1 Australia/ Murray-Darling basin

Australia's climate and rainfall is influenced by its geographic location, with the southwestern Pacific Ocean in the east and the Indian Ocean in the west (Trenberth, 1990; Nicholls et al., 1997). On interannual time scales the most dominant drivers for precipitation are the El-Niño-Southern Oscillation, ENSO, and the Indian Ocean Dipole, IOD (García-García et al., 2011; Risbey et al., 2009). Negative ENSO phases (El Niño) lead to reduced rainfall in the northern and eastern parts of Australia resulting in droughts especially in the center of Australia, while La Niña phases lead to strong precipitations and often flooding. Positive IOD events are linked to a decrease of precipitation in western Australia, where as negative events lead to an increase in rainfall (Trenberth, 1990; Nicholls et al., 1997) and `http://www.bom.gov.au/climate/`. At the beginning of the new century Australia experienced the severe "Millenium drought" (van Dijk et al., 2013). The strong La Niña event starting 2010 led to increased precipitations and caused floods all over the continent (`http://www.bom.gov.au/climate/enso/lnlist`).

Trends from WaterGAP have opposite sign and smaller magnitude as compared to Li21. The small trends for WaterGAP are also found by Schumacher et al. (2018) and Yang et al. (2020). The latter reported a mismatch between GRACE and WaterGAP supposedly due to a lack of ability of the model to represent ground water trends and soil moisture. Trends computed based on the reconstruction by Li21 are mostly positive indicating a shift from dry to wet. The trends from WaterGAP are mostly negative indicating a slightly increase in water mass loss.

The data sets show higher magnitudes in the average rate changes, when comparing the entire time frame with the twenty years before GRACE. For the entire time frame, the reconstructions by Li21 and WaterGAP both exhibit near-zero small acceleration, whereas HG19 suggests much higher rate range changes. The regions in the Murray-Darling basin exhibiting negative accelerations are associated with agricultural areas (van Dijk et al., 2013). The negative range rate changes might be related to a decline in the groundwater storage from 1993-2009 primary due to pumping for agricultural and domestic purpose (J. L. Chen et al., 2016). HG19 depicts positive accelerations for these regions, which might be due to the model formulation, that does not include any anthropogenic effects.

van Dijk et al. (2013) reported a positive trend in precipitation for the north and west of Australia, both regions show a positive acceleration for Li21, HG19 and WaterGAP in the twenty years prior to GRACE, reflecting the increase in water mass.

For the annual signal the data sets are in a good agreement except for one region, at the northwestern part of Australia. WaterGAP does not show any signal variation in this region. HG19 shows higher annual amplitude magnitudes compared to the other data sets for the entire time frame, whereas Li21 exhibits high magnitudes for the twenty years prior to GRACE. The minimal signal peaks for Australia range from October for the south-western part to February/ March for the southern part.

For the subseasonal signal the reconstruction displays low signal variability in the south and higher ones in the north of Australia. WaterGAP additionally depicts high signal variations for the eastern coast and the region around the Swan coastal area. Both regions are associated with a humid climate (Köppen, 1923) and are mostly affected by ENSO and IOD, both phenomena leading to changes in precipitation patterns and strength (Forootan et al., 2016).

On interannual signal scale WaterGAP exhibits similar signal patterns as for the subseasonal signal, but with higher signal magnitudes. The reconstructions show extended signal patterns covering the north and the eastern and southern coast. The inter-annual signal variation derived from HG19 show higher magnitudes compared to Li21 and WaterGAP, especially for the eastern coast. This region is strongly affected by ENSO events (Forootan et al., 2016; van Dijk et al., 2013). The change in precipitation creates signal variabilities visible in the data sets.

On interannual signal scale WaterGAP exhibits similar signal patterns as for the subseasonal signal, but with higher signal magnitudes. The reconstructions show extended signal patterns covering the north and the eastern and southern coast. The inter-annual signal variation derived from HG19 show higher magnitudes compared to Li21 and WaterGAP, especially for the eastern coast. This region is strongly affected by ENSO events (Forootan et al., 2016; van Dijk et al., 2013). The change in precipitation creates signal variabilities visible in the data sets.

3.3.2 Amazon basin

The Amazon basin is Earth biggest drainage basin, with high seasonal rainfall variability (Costa & Foley, 1998; Marengo, Liebmann, et al., 2012). Frappart et al. (2013) found a high correlation between rainfall and GRACE derived TWS for the years 2003−2010, making it quite reasonable to assume, that rainfall is the main driver for water storage changes in the Amazon river basin. On inter-annual time scale the precipitation patterns are influenced by ENSO, leading to several droughts and floods (Marengo, Tomasella, et al., 2012; Towner et al., 2021; Marengo & Espinoza, 2016; Xavier et al., 2010; Phillips et al., 2012; J. L. Chen et al., 2010).

The trend from Li21 for the Amazon basin is mostly positive, indicating a water mass increase for this region. WaterGAP on the other side depicts a positive trend for the Negro and Madeira sub-basin and a negative one for the Solimoes sub-basin.For their reconstruction Becker et al. (2011) performed an empirical orthogonal function (EOF) analysis of precipitation data over the Amazon for the years 1980−2008, 1990−2008, 2003−2008. The leading EOFs show a strong positive signal (in terms of water accumulation) for the western part and a negative one for the eastern part. The derived dominant signals by Becker et al. (2011) correspond nicely to the trend pictured by WaterGAP, leading to the conclusion, that the derived trend is mainly driven by precipitation. Floods and droughts in the Amazon basin are known to be driven by large-scale climate variability, like ENSO (Towner et al., 2021). According to Rodell et al. (2018) the mostly positive trend derived from Li21 is due to an increase in precipitation and floods mainly driven by La Niña after an decrease in precipitation related to El Niño (J. L. Chen et al., 2010; Marengo & Espinoza, 2016; Towner et al., 2021).

For the entire time period the average rate range changes are mostly negative, except for the Madeira sub-basin. The trend derived for Li21 is mostly positive, the acceleration mostly negative. This suggests, that the Amazon basin may be in a near equilibrium state. We find the same in the WaterGAP simulations, even though the trends magnitude of WaterGAP is smaller compared to Li21. For the twenty years prior to GRACE the range rate change pattern of WaterGAP does not change. A change from positive to negative acceleration for the Negro sub-basin for Li21 is visible. The accelerations for HG19 show an opposite pattern with respect to the accelerations computed from WaterGAP. This mismatch means that at interannual timescales the water storage change in the Amazon cannot be well explained by temperature and precipitation trends; this may be due to processes not well represented in either reconstruction or model, or due to biases in the forcing data.

All three data sets show high annual amplitudes over the main stream. Due to the resolution of the underlying GRACE solution the derived signal patterns for the reconstructions are blurred, where as in WaterGAP the river stream is visible. Comparing the annual amplitude for the entire time period and the twenty years prior to GRACE reveals no change in signal magnitude. However, the signal pattern changes slightly. For Li21 it becomes narrower, mimicking the WaterGAP river routing. For HG19 the annual signal pattern shifts slightly towards the North Atlantic. The annual phase of the datasets displays minimal signal peaks for December to January.

Both reconstructions show high subseasonal signals for the Negro and Amazon subbasins. For the main branch the magnitude of the subseasonal signal of Li21 is smaller compared to HG19. WaterGAP displays high subseasonal signal variations for the whole Amazon basin. The signal magnitude increases from Li21 to HG19 to WaterGAP.

High inter-annual signal variability over most parts of the Amazon basin are visible in all three data sets on both time scales. HG19 and WaterGAP both show stronger inter-annual signals compared to Li21, with WaterGAP showing the biggest inter-annual signal variability, followed by HG19 and Li21.

To better understand the information content and 'effective' spatial resolution of the detrended and deseasonalized reconstructions, we decided to compute the spatial autocorrelation for a given point in the Silimoes sub basin of the Amazon. As expected, highest correlations are found for the Silimoes and Madeira (sub) basin. HG19 also exhibits correlations of around 0.5 for the Congo, Niger, Nil, Zambezi, Amur, Kem, Mississippi, Saint-Laurent, Nelson and Murray basin, as well as Lake Victoria. Simular correlation patterns for the Congo, Amur, Mississippi and Kem basin are also revealed in the Li21 data, however, the correlation itself is smaller compared to HG19. Moderate long-range correlations are indeed visible for both data sets; this can be explained by the irregularities (with respect to the annual cycle) in the precipitation driven by ENSO.

3.3.3 Europe

Europe is facing more and more severe droughts, leading to an decrease in water storage (Gerdener et al., 2020; Boergens et al., 2020; Gudmundsson & Seneviratne, 2015; Spinoni et al., 2015a, 2015b). The largest changes are found for the big European basins: Volga, Danube, Dnieper, Don and Rhine (Rodell et al., 2018; Humphrey et al., 2016).

Li21 shows positive trends for the Mediterranean, Boreal and Atlantic parts of Europe for the full time frame. For Central Europe the reconstructions indicate mass losses increasing from west to eastern Europe, the biggest negative trends allocated around the Black Sea. This findings are in a good agreement with Rodell et al. (2018); Eicker et al. (2016); Tapley et al. (2019). WaterGAP exhibits positive trends for the Boreal regions and western part of the Danube basin and negative ones elsewhere. For the two decades prior to GRACE, WaterGAP displays a shift from negative to positive trends. This is in a good agreement with the findings of Spinoni et al. (2015b) Similar behaviours are found for Li21. The change in trend patterns between the two time series suggest a decrease in water storage after 2002, so strong, that it influences the derived trends and, therefore, dominates the time series. Following WaterGAP, the water storage decrease affects all European countries, where as in Li21 the affected countries are located in the middle of Europe.

For the full time series, WaterGAP and Li21 exhibit similar slightly positive acceleration patterns. HG19 displays high signal magnitudes compared to the other data sets showing mostly negative accelerations, with maximum values around the Black See and parts of France. Positive accelerations are shown for the region around the Adriatic Sea. For the years 1979-2002 the acceleration patterns become more pronounced in all three data sets. HG19 exhibits mostly positive acceleration except for Siberia, parts of Norway and parts of Spain. In contrast to that, Li21 depicts negative acceleration for all countries located between the Adriatic and Black Sea and positive ones over Siberia. These patterns are reflected by WaterGAP. However, WaterGAP nevertheless displays negative accelerations for Sweden, whereas both reconstructions show positive ones.

The annual amplitude of HG19 does not change through the time frames. Regions with higher annual amplitudes are located around the Black Sea, the southern part of Spain and Portugal and the north-eastern part of France and Germany and Iceland. Low annual amplitudes are shown for the Boreas and Atlantic countries. Similar patterns are found for WaterGAP and Li21 for the years 1979-2002. For the full time frame the signal magnitude is smaller compared to the twenty years prior to GRACE, while the signal pattern stays the same. For the full time frame the data sets exhibit values around DOY 80-150 for south-eastern part, values around DOY 200-300 for the eastern part and values above DOY 300 for Central Europe. For the Pre-GRACE time frame, all data sets show, that the minimal day of the annual signal amplitude only changes for the eastern and middle part of Europe, so regions, that are more effected by water mass changes. For the shorter time period the data sets shows a good agreement for central and southern Europe, Scandinavia and the region around the Adriatic Sea. For the region around the Back Sea, Li21 and WaterGAP display values around 150 DOY. The values derived from HG19 are around 230 DOY and closer to the full time frame.

The derived subseasonal signal from the data sets is the same for both time periods. Li21 and HG19 identify similar regions in Central Europe and the coastal border of Norway with high subseasonal signal variations. However, the magnitude of the signal shown by Li21 is smaller compared to HG19. Low subseasonal signal variations are shown for England. WaterGAP exhibit higher signal magnitudes compared to the reconstructions, with high variations for the northern and southern part of Europe, with a small band of lower magnitudes covering the continual part of Europe.

The reconstructions show high interannual signal variations for the eastern part of Europe and southern part of Spain. Especially for eastern Europe the derived magnitude is higher for the full time series compared to the years 1979-2002. For Central Europe Li21 displays moderate interannual signal variations. The derived signal variations from HG19 and WaterGAP are notably higher, with WaterGAP displaying signal variations up to two times bigger compared to Li21.

We also compute the autocorrelation and time lag for the detrended and deseasonalized reconstructions for a point in Central Europe (figure 10). As expected for both reconstructions the correlation is high around the chosen point, and drops to near zero after a few thousand kilometres. High correlation are thus (in this example) shown for the Loire, Rhone, Seine, Garonne Elbe, Rhine and Po basin. We find that for Li21, the correlation pattern is stronger compared to HG19 and extends more to the west. The time lag for both data sets is in the range of a month. As expected, long-range correlations are nearly zero.

4 Evaluation of global TWSA reconstructions with satellite tracking data

4.1 GRACE and GRACE-FO data

We derived a trend, acceleration, annual amplitude and annual phase using a least square approach for four different GRACE solutions and the reconstructions for the years 2002−2020 for the global land. The computed values are shown in table Appendix C.3. The first column of table Appendix C.3 displays the trend, if only trend and annual amplitude and phase are estimated. The second column shows the trend in case an additional parameter for the acceleration is taken into account during the estimation process. As the annual amplitude and annual phase are derived based on orthogonal functions, the annual amplitudes and phase are less sensitive towards additional parameters. The according time series are shown in figure Appendix C.1. Clearly visible are the different slopes depicted by the time series and the slight differences in amplitude between the different data products.

The trends in the first column of table Appendix C.3 all show slightly negative values, indicating a water loss over land. For the spherical harmonic solution the computed trend varies from −0.15 mm/yr to −0.22 mm/yr. The trend from the mass concentration block solution (mascon) from Center of Space Research (CSR) is around ten times bigger compared to the spherical harmonic (SH) solution. Similar results were found by Jing et al. (2019) for the Tibetan Plateau. The reconstruction from Li21 is based on the CSR mascon solution, the magnitude of the trend of this specific solution is reflected by the reconstruction. The negative trends of the mascon based solutions are also clearly visible in figure Appendix C.1. As mentioned before, the HG19 reconstruction is not designed to reproduce a trend, e.g. as contained in the training TWSA data.However, it is nevertheless possible to derive a trend for any given time period and other reseachers have used such trends. The derived trend of −0.16 mm/yr is in the order of the spherical harmonic solutions.

Adding a parameter for the acceleration to the estimation process gives the model more flexibility in terms of model estimation, leading to changes in the estimated magnitude and sign of the trend. For the spherical harmonic solutions the values range from 0.59 mm/yr to 0.69 mm/yr. The negative trend from the mascon solution and Li21 decreases from −1.76 mm/yr to −1.13 mm/yr. Again, the trend for HG19 is within the range of the estimated trends of the spherical harmonic solutions.

The estimated acceleration for all datasets is negative, ranging from −0.04 mm/yr^2 for the spherical harmonic solutions and HG19 to −0.03 mm/yr^2 for the CSR mascon solution. The CSR mascon solution and Li21 are quite close to each other in terms of derived trend. However, the acceleration derived for Li21 is two times bigger compared to the one derived from the CSR mascon.

The annual amplitude varies from 13 mm for the German Research Center for Geoscience (GFZ) spherical harmonic solution to 24 mm for the CSR mascon solution and Li21. The two reconstructions together with the mascon solution depict the highest amplitude with values over 20 mm.

We conclude, that the reconstructions seem to follow the signal properties of the GRACE solution used to train them. This can be seen for the trend of Li21, which is close to the one of the CSR mascon solution, and for the annual amplitude, which is inherited for both data sets from a mascon based GRACE solution. We find, that the annual amplitudes of the reconstructions are closer to the mascon based compared to the spherical harmonic GRACE solution.

4.2 Pre-GRACE era comparison to satellite laser ranging data

For this section the reconstructions have been expanded into spherical harmonics and have been truncated at a spherical harmonic degree of $n = 12$, omitting most of the higher frequency signals. For a spherical harmonic degree of $n = 12$ no filtering is necessary for the SLR data and, therefore, for the reconstructions.

The strongest negative trends in Li21 are related to the mass loss of the Caspian Sea (Loomis & Luthcke, 2017; Rodell et al., 2018) and the groundwater depletion in India, Iraq, Iran and parts of Arabian Peninsula (Joodaki et al., 2014; J. Chen et al., 2014; Rodell et al., 2018). Furthermore, negative trends are found for Alaska, the Mackenzie basin, the Mississippi basin, parts of Mexico, the Sahara, the Congo basin, the Tocantins, the Parnaiba, the Sao Francisco basin, northern part of South America and western Europe. Positive trends are shown for Australia, the Limpopo, Orange, Okavango and Zambezi, the Orinoco basin, parts of the Amazon basin and parts of Canada. Like Li21, HG19 exhibits negative trends for the Mississippi basin and the region around the Caspian sea, the latter with smaller signal magnitude compared to Li21. Notable negative trends are also found for the Parana basin and the Congo basin. Regions with positive trends are Australia, except for the Murray-Darling basin, the Orange basin, Okavango basin, Limpopo basin, Zambezi basin and part of the Amazon basin. For SLR, the strongest negative trends are located over the Ellesmere and the Baffin Islands, the region of the Caspian, the Aral and the Black Sea. The trend derived for Africa is mostly positive, with the highest values depict for the Congo and Zambezi basin. A negative trend is visible for the west-southern part of Africa. Australia exhibits mostly positive trends, with negative trends at the southern western coast. In comparison, HG19 is missing signals around the Hudson Bay, Alaska and Fennoscandia, for which SLR displays the strongest trends. In Li21 these signals are only visible for Ellesmere and Baffin Islands. The data sets display negative trends for the US, the region around the Caspian, the Black and the Aral Sea, India and Siberia.

The accelerations are shown in the second row of figure 3, left to right: HG19, Li21, SLR. Strong acceleration patterns for SLR are located around the Hudson Bay, the Orinoco basin and spreading over the Black Sea, the Caspian Sea, India and China. Positive accelerations are derived for Alaska, the region under the Orinoco basin, the Congo and Zambezi basin, the Ob basin and northern Australia. Li21 exhibits negative range rate changes for the Parana basin, the Tocantins basin, the Parnaiba basin, the Sao Francisco basin, the Sierra Neveda, the Zambezi basin, the Black Sea, the Caspian Sea, northern Siberia, India and northern Australia. Positive accelerations are situated over the Ob basin, the Nil basin, the Japura and Solimoes basin (sub-basins of the Amazon basin), the Yukon basin, the southern land of the Hudson Bay and the southwestern part of Australia. HG19 exhibits positive accelerations for the Amazon basin and the northern part of South America, the Sahara and Arabian Peninsula, the Zambezi, Orange, Okagavo and Limpopo basin, Australia with the exception of the Murray-Darling basin, Alaska and Siberia. For the other regions negative range rate changes are shown. The magnitude as well as location of the accelerations differ widely across the data sets.

The annual amplitudes and corresponding phases are shown in the third and fourth row of figure 3. SLR depicts high values for the Amazon basin, the region around the Hudson Bay and the Gulf of Alaska, the Congo and Zambezi basin, the Niger and Lake Chad basin, the Ob basin, western India and southern China and north eastern Australia. For the Amazon basin, the Congo basin and north-eastern Australia the derived minimal signal magnitude is around January. The Mississippi basin and the region around the Hudson Bay display minimal phase values around DOY 120-150. For the Zambezi basin and ice covered regions the minimum is found for January. The Niger and Lake Chad basin, western India and southern China show minimal phase values around DOY 200-300. For the Ob basin it is around DOY 50-100. The average annual amplitudes in the Amazon river, around the Gulf of Alaska and in the Ob basin are also reflected by the reconstructions, with HG19 showing higher values compared to Li21. Li21 also shows a signal over the La Plata basin, the Mississippi basin and for the region between the Black and the Caspian Sea. The strong annual signal in the Congo and Zambezi basin detected by SLR is missing. HG19 pictures high annual signal amplitudes over Alaska, California, in the Mississippi basin, in the Amazon and La Plata basin, in the Zambezi

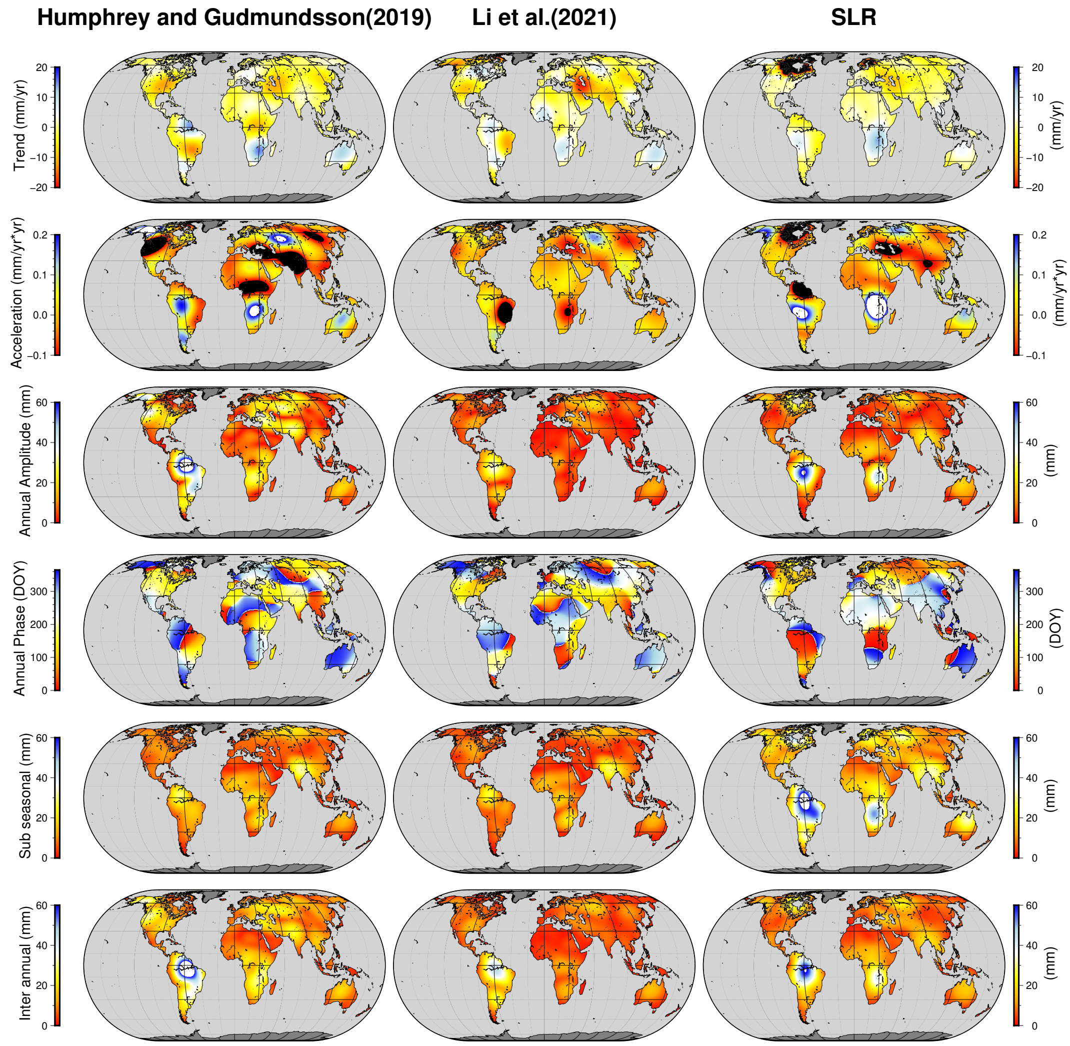

Figure 3. Trend, acceleration, annual amplitude and phase, subseasonal and interannual signal for the years 1992-2002 for the reconstructions and SLR on a global scale

basin, over the Victoria Sea, around the Black and the Caspian Sea, the Ob basin, north-western India and north-eastern Australia. The signal in the Mississippi basin, Australia and the signal spreading from the Black over the Caspian Sea to India can also be found for Li21 and partially for the SLR data, even though the signal magnitude is smaller. The annual amplitudes related to the La Plata basin and around the Victoria Sea are not reflected by Li21 or SLR. The signal in the Hudson Bay detected by SLR is not found in the reconstructions. For the Amazon basin and Alaska Li21 and HG19 depict minimal signal magnitudes around DOY 360. SLR exhibits values around zero, suggesting a small frequency shift between SLR and the reconstructions. A similar behaviour is found for Siberia, China and India.

The subseasonal signal is displayed in the fifth row of figure 3. Similar to the annual amplitude SLR exhibits high signal variations for the area around the Hudson Bay, Alaska, the Amazon basin and the Congo and Zambezi basin. Regions with a moderate subseasonal signal magnitudes are India and southern China, northern Australia, the region above the Black Sea and Alaska. Notable are also signals for the Mackenzie and Mississippi basin and the Caspian Sea. Both reconstructions show high subseasonal signal variation for the Orinoco basin, the Zambezi basin and over north-western India. Regions with lower signal variations are the northern part of Siberia, northern Australia, the Negro and Japura basin (both part of the Amazon basin), the Magdalena basin and in the Niger basin. HG19 exhibits higher signal magnitudes compared to Li21. The regions identified by the reconstructions are only partly found in the SLR data. The signal over South America in the SLR data is more pronounced and extends over the majority of the continent. The same holds for the signal over Africa and Australia. The high subseasonal signal variations over the Hudson Bay, and north-eastern Europe are not visible in the reconstructions.

The inter-annual signal variation is shown in the last row of figure 3. Similar regions compared to the annual amplitudes are found to have high inter-annual signal changes. In comparison to SLR and Li21, HG19 exhibits higher magnitudes for Siberia, Alaska and the region around the Caspian Sea. SLR shows bigger magnitudes for the Amazon basin, the Zambezi basin and the region around the Hudson Bay.

Except for region affected by GIA, the data sets exhibit similar signal patterns in all metrics. However, the magnitude of the signal differs. Strong water storage changes are, especially but not exclusively, found in the Amazon basin, the Congo, Zambezi, Limpopo, Orange and Okavango basin. As expected, the SLR data reveals larger trends and accelerations over regions affected by GIA, which are not present in the reconstructions.

4.3 Long-term evolution of water storage at river basin scale

We derived water storage changes on an interannual scale for the years 1979−2020 for the reconstructions for nine major river basins. The selected basin are the Amazon basin in South America, the Mississippi and Mackenzie basin in North America, the Nile, Niger and Congo basin in Africa and the Danube and Volga basin in Europe. The basins are visualised in figure Appendix D.2.

Several flood and drought events in the Amazon basin fall into the observation period (Marengo & Espinoza, 2016). The droughts in the years 1983 and 1997−1998 were linked to strong El Niño events (Marengo & Espinoza, 2016) and are clearly visible in the time series. Additionally, the droughts in the years 1980, 1995, 2005, the strong drought in 2015 and the floods from 1989, 1999 and 2005 are reflected by the data sets. Except for the strong drought in 2015, for which both data sets show similar signal magnitude, HG19 reveals stronger changes in the water storage compared to Li21.

The Congo basin is situated in about the same geographical latitude as the Amazon basin (Nicholson, 2022; Amarasekera et al., 1997) and is comparable to the Amazon basin in terms of climate and ecology (Nicholson, 2022). However, precipitation is lower in the Congo basin compared to the Amazon basin and, therefore, seasonal and inter-annual variability is smaller. Over the whole time series, HG19 shows a decrease in water storage for the basin, a negative trend is clearly visible. The time series from Li21 oscillates around zero. In comparison to HG19, Li21 seams to overestimate the water storage changes for the GRACE era and underestimates them for the pre GRACE era. Interestingly, the reconstructions show a good agreement for the years 1998−2002.

For the Danube basin both reconstructions reveal a loss of water mass from the year 1983 until 1990. After the year 1992 both reconstructions exhibit an increase of water storage followed by several peaks in the years 2006, 2011 and 2015.

For the Ganges basin the inter-annual time series for Li21 do not show any significant change in the water storage. HG19 on the other side, shows fluctuations of around 60 Gt over the whole time series. The precipitation over the Ganges and India depends highly on the Indian summer monsoon (MIRZA et al., 1998; Kumar et al., 2010). The moonson cycles are linked to ENSO and the IOD modes. The strong flood in 1998 partly due to the strong La Niña event is clearly visible in the reconstruction of HG19, as well as the decrease in water mass during the drought in 2002. Both events are only slightly to not visible in Li21.

The inter-annual time series for the Mackenzie river basin are in a partly agreement, for the years 1986−1989 or 2004−2015. For the years 1979−1984 and 2011−2015 the decrease in water storage shown by HG19 is more pronounced compared to Li21. The opposite is the case for the years 2002−2005 and 2016 onward. The time series for HG19 seam to be dominated by oscillating signal for the years 1992−199 with a wavelength of 3 years. This signal pattern is not reflected by the Li21 reconstruction.

The time series of the interannual signal of the Mississippi basin show a strong decrease in water mass for the years 1981, 1989 and 2013 and an increase in water storage for the years 1983 − 1987, 1994 and 2010. In comparison, HG19 exhibits stronger signal changes compared to Li21.

For the Nile basin a similar effect as for the Congo basin can be observed. HG19 shows a negative slope over the whole period, where as Li21 depicts no visible change in water storage.

During the GRACE era the reconstruction shows a strong decrease in water mass for the Niger basin for 2011. Before 2003 the reconstructions vary greatly, exhibiting partially inverse storage changes, like for the beginning of the time series. Compared to Li21, HG19 displays stronger signal oscillations for the years 1979 − 2003 and smaller ones for the GRACE era.

The reconstructions display a slight decrease in water storage for the Volga basin for the years 2010 to 2015 with a shift towards an increase for the years after 2015. The magnitude of the water mass loss for HG19 is higher compared to Li21. For the years 2000−2008 and the first years of the reconstructions, the data sets exhibit similar water mass changes. Notable, especially for HG19, but also for Li21 is an increase in water content for 1991.

We conclude, that the time series of the reconstructions are surprisingly close beyond the annual cycle for basins with signal amplitudes, dominated by seasonal precipitation, like the Amazon, Danube and Mississippi basin. For basins with low inter-annual signal variability, like the Congo and Nile basin we find differing long-term trends.

5 Conclusions

In this study we derived trend, acceleration, annual amplitude, annual phase, inter-annual signal parts and subseasonal TWSA signals from the two global reconstructions from Humphrey and Gudmundsson (2019) and F. Li et al. (2021) and from the hydrological model WaterGAP for the years 1979 − 2016 (full time frame) and 1979 − 2002 (pre GRACE time frame). Furthermore, we compared the reconstructions to the low degree gravity fields derived from SLR (Löcher & Kusche, 2020) for the pre GRACE time frame 1992 − 2002 for a spherical harmonic degree of expansion of $n = 12$. We also derived trend, acceleration, annual amplitude and phase from four different GRACE solutions and the reconstructions for the years 2002 − 2020.

We find similar sign and magnitude of water storage changes over the full time period and pre GRACE time period for South America and the large basins in southern Africa, like the Congo and Zambezi basin. These regions also stick out with highest TWSA in the SLR-derived maps of mass variability. For Europe, Siberia, India, southern China, Australia and North America the magnitude of TWSA over the four decades differs between the reconstructions and the hydrological model. The SLR observations exhibit higher magnitudes of water storage changes for India, northern China, northern Australia, Alaska and the north-eastern part of Siberia. However we caution that the SLR data also include gravity changes due to glacier mass variability, and a direct comparison is most meaningful in areas free of glaciers. Except for the acceleration (and trend), the signal pattern and magnitude of TWSA is similar for the full and pre GRACE time frame. The acceleration for the pre GRACE time period exhibit notably higher signal magnitudes and patterns compared to the full time frame. We recap here, that HG19 is not tuned to reproduce a trend and the one by Li21 is not reconstructed, but recovered from the GRACE period.

Given the good agreement of all data sets, we suggest, that the Congo basin has indeed suffered a prolonged loss of water storage over the last 40 years. The US show a similar picture, with negative accelerations and trends, also suggesting a water mass loss over the last fourty years. For the Amazon basin high TWSA variations are found. According to the trend and range rate changes the Amazon basin appears, over the last four decades in a-near equilibrium. For Europe and Australia no significant trend or acceleration pattern is found. The huge drying patterns of the Caspian Sea and the Aral Sea, clearly visible in GRACE observations, are not reflected by trend or accelerations over the fourty years. For this region moderate TWSA magnitudes are found on subseasonal and interannual scale. SLR on the other side, exhibits strong negative acceleration, with a slightly negative trend in this region.

Open Research Section

Data availability statement

CSR mascon (Save et al., 2016; Save, 2020): `https://www2.csr.utexas.edu/grace/RL06_mascons.html`; GRACE L2 data: GFZ (Dahle et al., 2018), ITSG2018 (Mayer-Gürr et al., 2018), CSR CSR (2018), all GRACE data sets were downloaded from `http://icgem.gfz-potsdam.de/series`; Reconstruction by F. Li (2021): `https://datadryad.org/stash/dataset/doi:10.5061/dryad.z612jm6bt`; Reconstruction by Humphrey (2019): `https://doi.org/10.6084/m9.figshare.7670849`; SLR hybrid solution by Löcher and Kusche (2020): `http://icgem.gfz-potsdam.de/series/04_SLR/IGG_SLR_HYBRID`; WaterGAP data (Müller Schmied et al., 2020b): `https://doi.pangaea.de/10.1594/PANGAEA.918447?format=html#download`

Acknowledgments

The authors acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—SFB 1502/1-2022—Project No. 450058266.

References

Altamimi, Z., Rebischung, P., Métivier, L., & Collilieux, X. (2016). ITRF2014: A new release of the International Terrestrial Reference Frame modeling nonlinear station motions. *Journal of Geophysical Research: Solid Earth*, *121*(8), 6109–6131. Retrieved 2023-03-13, from `https://onlinelibrary.wiley.com/doi/abs/10.1002/2016JB013098` doi: 10.1002/2016JB013098

Amarasekera, K. N., Lee, R. F., Williams, E. R., & Eltahir, E. A. B. (1997). ENSO and the natural variability in the flow of tropical rivers. *Jour-*

nal of Hydrology, *200*(1), 24–39. Retrieved 2022-12-02, from https://www.sciencedirect.com/science/article/pii/S0022169496033409 doi: 10.1016/S0022-1694(96)03340-9

Becker, M., Meyssignac, B., Xavier, L., Cazenave, A., Alkama, R., & Decharme, B. (2011). Past terrestrial water storage (1980–2008) in the Amazon Basin reconstructed from GRACE and in situ river gauging data. *Hydrology and Earth System Sciences*, *15*(2), 533–546. Retrieved 2022-07-08, from https://hess.copernicus.org/articles/15/533/2011/ doi: 10.5194/hess-15-533-2011

Boergens, E., Güntner, A., Dobslaw, H., & Dahle, C. (2020). Quantifying the Central European Droughts in 2018 and 2019 With GRACE Follow-On. *Geophysical Research Letters*, *47*(14). Retrieved 2023-01-09, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2020GL087285 doi: 10.1029/2020GL087285

Chanard, K., Fleitout, L., Calais, E., Rebischung, P., & Avouac, J.-P. (2018). Toward a Global Horizontal and Vertical Elastic Load Deformation Model Derived from GRACE and GNSS Station Position Time Series. *Journal of Geophysical Research: Solid Earth*, *123*(4), 3225–3237. Retrieved 2023-01-09, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2017JB015245 doi: 10.1002/2017JB015245

Chen, J., Li, J., Zhang, Z., & Ni, S. (2014). Long-term groundwater variations in Northwest India from satellite gravity measurements. *Global and Planetary Change*, *116*, 130–138. Retrieved 2022-11-25, from https://www.sciencedirect.com/science/article/pii/S0921818114000526 doi: 10.1016/j.gloplacha.2014.02.007

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*(12). Retrieved 2022-11-29, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2010WR009383 doi: 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 observations. *Global and Planetary Change*, *139*, 56–65. Retrieved 2022-11-25, from https://www.sciencedirect.com/science/article/pii/S0921818116300108 doi: 10.1016/j.gloplacha.2016.01.002

Chen, W., Li, J., Ray, J., & Cheng, M. (2017). Improved geophysical excitations constrained by polar motion observations and GRACE/SLR time-dependent gravity. *Geodesy and Geodynamics*, *8*(6), 377–388. Retrieved 2023-03-14, from https://www.sciencedirect.com/science/article/pii/S1674984717300800 doi: 10.1016/j.geog.2017.04.006

Cheng, M., & Ries, J. (2023). C20 and C30 Variations From SLR for GRACE/GRACE-FO Science Applications. *Journal of Geophysical Research: Solid Earth*, *128*(2). Retrieved 2023-03-13, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2022JB025459 doi: 10.1029/2022JB025459

Cheng, M., Ries, J. C., & Tapley, B. D. (2011). Variations of the Earth's figure axis from satellite laser ranging and GRACE. *Journal of Geophysical Research: Solid Earth*, *116*(B1). Retrieved 2023-03-13, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2010JB000850 doi: 10.1029/2010JB000850

Cheng, M. K., Ries, J. C., & Tapley, B. D. (2013). Geocenter Variations from Analysis of SLR Data. In Z. Altamimi & X. Collilieux (Eds.), *Reference Frames for Applications in Geosciences* (pp. 19–25). Berlin, Heidelberg: Springer. doi: 10.1007/978-3-642-32998-2_4

Church, J. A., White, N. J., Coleman, R., Lambeck, K., & Mitrovica, J. X. (2004). Estimates of the Regional Distribution of Sea Level Rise over the

1950–2000 Period. *Journal of Climate*, *17*(13), 2609–2625. Retrieved 2023-03-13, from https://journals.ametsoc.org/view/journals/clim/17/13/1520-0442_2004_017_2609_eotrdo_2.0.co_2.xml doi: 10.1175/1520-0442(2004)017⟨2609:EOTRDO⟩2.0.CO;2

Cleveland, R., Cleveland, W., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on Loess. *Journal of official statistics*, *6*, pp. 3–73.

Costa, M. H., & Foley, J. A. (1998). A comparison of precipitation datasets for the Amazon Basin. *Geophysical Research Letters*, *25*(2), 155–158. Retrieved 2022-11-29, from https://onlinelibrary.wiley.com/doi/abs/10.1029/97GL03502 doi: 10.1029/97GL03502

Cressie, N. (1993). *Statistics for Spatial Data.* Wiley-Interscience.

CSR. (2018). *GRACE FIELD GEOPOTENTIAL COEFFICIENTS CSR RELEASE 6.0| Physical Oceanography Distributed Active Archive Center (PO.DAAC).* Retrieved 2023-03-13, from https://podaac.jpl.nasa.gov/dataset/GRACE_GSM_L2_GRAV_CSR_RL06(lastaccess:23November2022)

Dahle, C., Flechtner, F., Murböck, M., Michalak, G., Neumayer, H., Abrykosov, O., ... König, R. (2018). GRACE Geopotential GSM Coefficients GFZ RL06. V. 6.0. GFZ Data Services.
doi: https://doi.org/10.5880/GFZ.GRACE_06_GSM

Dahle, C., Flechtner, F., Murböck, M., Michalak, G., Neumayer, K.-H., Abrykosov, O., ... König, R. (2019). GRACE-FO Geopotential GSM Coefficients GFZ RL06. V. 6.0. GFZ Data Services.
doi: https://doi.org/10.5880/GFZ.GRACEFO_06_GSM

Eicker, A., Forootan, E., Springer, A., Longuevergne, L., & Kusche, J. (2016). Does GRACE see the terrestrial water cycle "intensifying"? *Journal of Geophysical Research: Atmospheres*, *121*(2), 733–745. Retrieved 2022-09-23, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2015JD023808 doi: 10.1002/2015JD023808

Eicker, A., Schumacher, M., Kusche, J., Döll, P., & Schmied, H. M. (2014). Calibration/Data Assimilation Approach for Integrating GRACE Data into the WaterGAP Global Hydrology Model (WGHM) Using an Ensemble Kalman Filter: First Results. *Surveys in Geophysics*, *35*(6), 1285–1309. Retrieved 2022-09-08, from https://doi.org/10.1007/s10712-014-9309-8 doi: 10.1007/s10712-014-9309-8

Ferreira, V. G., Andam-Akorful, S. A., Dannouf, R., & Adu-Afari, E. (2019). A Multi-Sourced Data Retrodiction of Remotely Sensed Terrestrial Water Storage Changes for West Africa. *Water*, *11*(2), 401. Retrieved 2022-07-25, from https://www.mdpi.com/2073-4441/11/2/401 doi: 10.3390/w11020401

Forootan, E. (2014). *Statistical Signal Decomposition Techniques for Analyzing Time-Variable Satellite Gravimetry Data* (Thesis, Universitäts- und Landesbibliothek Bonn). Retrieved 2022-08-02, from https://bonndoc.ulb.uni-bonn.de/xmlui/handle/20.500.11811/5858

Forootan, E., Khandu, Awange, J. L., Schumacher, M., Anyah, R. O., van Dijk, A. I. J. M., & Kusche, J. (2016). Quantifying the impacts of ENSO and IOD on rain gauge and remotely sensed precipitation products over Australia. *Remote Sensing of Environment*, *172*, 50–66. Retrieved 2022-11-28, from https://www.sciencedirect.com/science/article/pii/S0034425715301796 doi: 10.1016/j.rse.2015.10.027

Forootan, E., Kusche, J., Loth, I., Schuh, W.-D., Eicker, A., Awange, J., ... Shum, C. K. (2014). Multivariate Prediction of Total Water Storage Changes Over West Africa from Multi-Satellite Data. *Surveys in Geophysics*, *35*(4), 913–940. Retrieved from https://doi.org/10.1007/s10712-014-9292-0 doi: 10.1007/s10712-014-9292-0

Forootan, E., Schumacher, M., Mehrnegar, N., Bezděk, A., Talpe, M. J., Farzaneh, S., ... Shum, C. K. (2020). An Iterative ICA-Based Reconstruction Method to Produce Consistent Time-Variable Total Water Storage Fields Using GRACE and Swarm Satellite Data. *Remote Sensing*, *12*(10), 1639. Retrieved 2022-07-07, from `https://www.mdpi.com/2072-4292/12/10/1639` doi: 10.3390/rs12101639

Frappart, F., Seoane, L., & Ramillien, G. (2013). Validation of GRACE-derived terrestrial water storage from a regional approach over South America. *Remote Sensing of Environment*, *137*, 69–83. Retrieved 2022-11-29, from `https://www.sciencedirect.com/science/article/pii/S0034425713002010` doi: 10.1016/j.rse.2013.06.008

García-García, D., Ummenhofer, C. C., & Zlotnicki, V. (2011). Australian water mass variations from GRACE data linked to Indo-Pacific climate variability. *Remote Sensing of Environment*, *115*(9), 2175–2183. Retrieved 2022-11-25, from `https://www.sciencedirect.com/science/article/pii/S0034425711001179` doi: 10.1016/j.rse.2011.04.007

Gerdener, H., Engels, O., & Kusche, J. (2020). A framework for deriving drought indicators from the Gravity Recovery and Climate Experiment (GRACE). *Hydrology and Earth System Sciences*, *24*(1), 227–248. Retrieved 2023-01-09, from `https://hess.copernicus.org/articles/24/227/2020/` doi: 10.5194/hess-24-227-2020

Geruo, A., Velicogna, I., Kimball, J. S., Du, J., Kim, Y., Colliander, A., & Njoku, E. (2017). Satellite-observed changes in vegetation sensitivities to surface soil moisture and total water storage variations since the 2011 Texas drought. *Environmental Research Letters*, *12*(5). Retrieved 2023-01-09, from `https://dx.doi.org/10.1088/1748-9326/aa6965` doi: 10.1088/1748-9326/aa6965

Girotto, M., De Lannoy, G. J. M., Reichle, R. H., & Rodell, M. (2016). Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model. *Water Resources Research*, *52*(5), 4164–4183. Retrieved 2023-03-13, from `https://onlinelibrary.wiley.com/doi/abs/10.1002/2015WR018417` doi: 10.1002/2015WR018417

Gudmundsson, L., & Seneviratne, S. I. (2015). European drought trends. *Proceedings of the International Association of Hydrological Sciences*, *369*, 75–79. Retrieved 2023-01-09, from `https://piahs.copernicus.org/articles/369/75/2015/` doi: 10.5194/piahs-369-75-2015

Han, S.-C., Ghobadi-Far, K., Yeo, I.-Y., McCullough, C. M., Lee, E., & Sauber, J. (2021). GRACE Follow-On revealed Bangladesh was flooded early in the 2020 monsoon season due to premature soil saturation. *Proceedings of the National Academy of Sciences*, *118*(47). Retrieved from `https://www.pnas.org/doi/abs/10.1073/pnas.2109086118` doi: 10.1073/pnas.2109086118

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. Retrieved 2022-05-18, from `https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803` doi: 10.1002/qj.3803

Humphrey, V. (2019). GRACE-REC: A reconstruction of climate-driven water storage changes over the last century [dataset]. doi: 10.6084/m9.figshare.7670849.v3

Humphrey, V., & Gudmundsson, L. (2019). GRACE-REC: a reconstruction of climate-driven water storage changes over the last century. *Earth System Science Data*, *11*(3), 1153–1170. Retrieved from `https://essd.copernicus.org/articles/11/1153/2019/` doi: 10.5194/essd-11-1153-2019

Humphrey, V., Gudmundsson, L., & Seneviratne, S. I. (2016). Assessing Global Water Storage Variability from GRACE: Trends, Seasonal Cycle, Subseasonal Anomalies and Extremes. *Surveys in Geophysics*, *37*(2), 357–395. Retrieved

2022-05-24, from http://link.springer.com/10.1007/s10712-016-9367-1 doi: 10.1007/s10712-016-9367-1

Jensen, L., Eicker, A., Dobslaw, H., Stacke, T., & Humphrey, V. (2019). Long-Term Wetting and Drying Trends in Land Water Storage Derived From GRACE and CMIP5 Models. *Journal of Geophysical Research: Atmospheres*, *124*(17-18), 9808–9823. Retrieved 2023-03-13, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2018JD029989 doi: 10.1029/2018JD029989

Jin, S., Chambers, D. P., & Tapley, B. D. (2010). Hydrological and oceanic effects on polar motion from GRACE and models. *Journal of Geophysical Research: Solid Earth*, *115*(B2). Retrieved 2023-01-09, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2009JB006635 doi: 10.1029/2009JB006635

Jing, W., Zhang, P., & Zhao, X. (2019). A comparison of different GRACE solutions in terrestrial water storage trend estimation over Tibetan Plateau. *Scientific Reports*, *9*(1). Retrieved 2022-12-14, from https://www.nature.com/articles/s41598-018-38337-1 doi: 10.1038/s41598-018-38337-1

Joodaki, G., Wahr, J., & Swenson, S. (2014). Estimating the human contribution to groundwater depletion in the Middle East, from GRACE data, land surface models, and well observations. *Water Resources Research*, *50*(3), 2679–2692. Retrieved 2022-09-23, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2013WR014633 doi: 10.1002/2013WR014633

Khaki, M., Hoteit, I., Kuhn, M., Awange, J., Forootan, E., van Dijk, A. I. J. M., ... Pattiaratchi, C. (2017). Assessing sequential data assimilation techniques for integrating GRACE data into a hydrological model. *Advances in Water Resources*, *107*, 301–316. Retrieved 2023-03-13, from https://www.sciencedirect.com/science/article/pii/S0309170816307564 doi: 10.1016/j.advwatres.2017.07.001

Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World Map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, *15*(3), 259–263. Retrieved 2022-07-12, from http://www.schweizerbart.de/papers/metz/detail/15/55034/World_Map_of_the_Koppen_Geiger_climate_classificat?af=crossref doi: 10.1127/0941-2948/2006/0130

Kumar, V., Jain, S. K., & Singh, Y. (2010). Analysis of long-term rainfall trends in India. *Hydrological Sciences Journal*, *55*(4), 484–496. Retrieved 2022-12-14, from https://doi.org/10.1080/02626667.2010.481373 doi: 10.1080/02626667.2010.481373

Kusche, J. (2007). Approximate decorrelation and non-isotropic smoothing of time-variable GRACE-type gravity field models. *Journal of Geodesy*, *81*(11), 733–749. Retrieved 2022-11-15, from https://doi.org/10.1007/s00190-007-0143-3 doi: 10.1007/s00190-007-0143-3

Kusche, J., Eicker, A., Forootan, E., Springer, A., & Longuevergne, L. (2016). Mapping probabilities of extreme continental water storage changes from space gravimetry. *Geophysical Research Letters*, *43*(15), 8026–8034. Retrieved 2023-01-09, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2016GL069538 doi: 10.1002/2016GL069538

Köppen, W. (1923). *Die Klimate der Erde: Grundriss der Klimakunde.* De Gruyter. Retrieved 2022-07-14, from https://www.degruyter.com/document/doi/10.1515/9783111491530/html?lang=de doi: 10.1515/9783111491530

Lenczuk, A., Weigelt, M., Kosek, W., & Mikocki, J. (2022). Autoregressive Reconstruction of Total Water Storage within GRACE and GRACE Follow-On Gap Period. *Energies*, *15*(13). Retrieved 2022-07-07, from https://www.mdpi.com/1996-1073/15/13/4827 doi: 10.3390/en15134827

Li, B., Rodell, M., Kumar, S., Beaudoing, H. K., Getirana, A., Zaitchik, B. F., ... Bettadpur, S. (2019). Global GRACE Data Assimilation for Groundwater and Drought Monitoring: Advances and Challenges. *Wa-*

ter Resources Research, *55*(9), 7564–7586. Retrieved 2023-01-09, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2018WR024618 doi: 10.1029/2018WR024618

Li, F. (2021). Long-term (1979-present) total water storage anomalies over the global land derived by reconstructing GRACE data [dataset]. Retrieved from http://datadryad.org/stash/dataset/doi:10.5061/dryad.z612jm6bt doi: 10.5061/DRYAD.Z612JM6BT

Li, F., Kusche, J., Chao, N., Wang, Z., & Löcher, A. (2021). Long-Term (1979-Present) Total Water Storage Anomalies Over the Global Land Derived by Reconstructing GRACE Data. *Geophysical Research Letters*, *48*(8). Retrieved 2022-07-18, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2021GL093492 doi: 10.1029/2021GL093492

Li, F., Kusche, J., Rietbroek, R., Wang, Z., Forootan, E., Schulze, K., & Lück, C. (2020). Comparison of data-driven techniques to reconstruct (1992-2002) and predict (2017-2018) GRACE-like gridded total water storage changes using climate inputs. *Water Resources Research, 56, e2019WR026551*. doi: 10.1029/2019wr026551

Long, D., Shen, Y., Sun, A., Hong, Y., Longuevergne, L., Yang, Y., … Chen, L. (2014). Drought and flood monitoring for a large karst plateau in Southwest China using extended GRACE data. *Remote Sensing of Environment*, *155*, 145–160. Retrieved from https://www.sciencedirect.com/science/article/pii/S0034425714003034 doi: https://doi.org/10.1016/j.rse.2014.08.006

Loomis, B. D., & Luthcke, S. B. (2017). Mass evolution of Mediterranean, Black, Red, and Caspian Seas from GRACE and altimetry: accuracy assessment and solution calibration. *Journal of Geodesy*, *91*, 195–206. Retrieved from https://doi.org/10.1007/s00190-016-0952-3 doi: 10.1007/s00190-016-0952-3

Löcher, A., & Kusche, J. (2020). A hybrid approach for recovering high-resolution temporal gravity fields from satellite laser ranging. *Journal of Geodesy*, *95*(1), 6. Retrieved from https://doi.org/10.1007/s00190-020-01460-x doi: 10.1007/s00190-020-01460-x

Marengo, J. A., & Espinoza, J. C. (2016). Extreme seasonal droughts and floods in Amazonia: causes, trends and impacts. *International Journal of Climatology*, *36*(3), 1033–1050. Retrieved 2022-11-30, from https://onlinelibrary.wiley.com/doi/abs/10.1002/joc.4420 doi: 10.1002/joc.4420

Marengo, J. A., Liebmann, B., Grimm, A. M., Misra, V., Silva Dias, P. L., Cavalcanti, I. F. A., … Alves, L. M. (2012). Recent developments on the South American monsoon system. *International Journal of Climatology*, *32*(1), 1–21. Retrieved 2022-11-30, from https://onlinelibrary.wiley.com/doi/abs/10.1002/joc.2254 doi: 10.1002/joc.2254

Marengo, J. A., Tomasella, J., Soares, W. R., Alves, L. M., & Nobre, C. A. (2012). Extreme climatic events in the Amazon basin. *Theoretical and Applied Climatology*, *107*(1), 73–85. Retrieved 2022-11-30, from https://doi.org/10.1007/s00704-011-0465-1 doi: 10.1007/s00704-011-0465-1

Mayer-Gürr, T., Behzadpur, S., Ellmer, M., Kvas, A., Klinger, B., Strasser, S., & Zehentner, N. (2018). *ITSG-Grace2018 - Monthly, Daily and Static Gravity Field Solutions from GRACE.* Retrieved 2022-12-14, from https://dataservices.gfz-potsdam.de/icgem/showshort.php?id=escidoc:3600910 doi: 10.5880/ICGEM.2018.003

MIRZA, M. Q., WARRICK, R. A., ERICKSEN, N. J., & KENNY, G. J. (1998). Trends and persistence in precipitation in the Ganges, Brahmaputra and Meghna river basins. *Hydrological Sciences Journal*, *43*(6), 845–858. Retrieved 2022-12-14, from https://doi.org/10.1080/02626669809492182 doi: 10.1080/02626669809492182

Mémin, A., Boy, J.-P., & Santamaría-Gómez, A. (2020). Correcting GPS measurements for non-tidal loading. *GPS Solutions*, *24*(2), 45. Retrieved 2023-01-09, from https://doi.org/10.1007/s10291-020-0959-3 doi: 10.1007/s10291-020-0959-3

Müller Schmied, H., Cáceres, D., Eisner, S., Flörke, M., Herbert, C., Niemann, C., ... Döll, P. (2020a). The global water resources and use model WaterGAP v2.2d - Standard model output. Retrieved 2022-07-27, from https://doi.pangaea.de/10.1594/PANGAEA.918447 doi: 10.1594/PANGAEA.918447

Müller Schmied, H., Cáceres, D., Eisner, S., Flörke, M., Herbert, C., Niemann, C., ... Döll, P. (2020b). The global water resources and use model WaterGAP v2.2d - Standard model output [dataset]. Retrieved from https://doi.pangaea.de/10.1594/PANGAEA.918447 doi: 10.1594/PANGAEA.918447

Nicholls, N., Drosdowsky, W., & Lavery, B. (1997). Australian rainfall variability and change. *Weather*, *52*(3), 66–72. Retrieved 2023-05-04, from https://onlinelibrary.wiley.com/doi/abs/10.1002/j.1477-8696.1997.tb06274.x doi: 10.1002/j.1477-8696.1997.tb06274.x

Nicholson, S. E. (2022). The Rainfall and Convective Regime over Equatorial Africa, with Emphasis on the Congo Basin. In *Congo Basin Hydrology, Climate, and Biogeochemistry* (pp. 25–48). Retrieved 2022-12-02, from https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119657002.ch3 doi: 10.1002/9781119657002.ch3

Pearlman, M., Arnold, D., Davis, M., Barlier, F., Biancale, R., Vasiliev, V., ... Bloßfeld, M. (2019). Laser geodetic satellites: a high-accuracy scientific tool. *Journal of Geodesy*, *93*(11), 2181–2194. Retrieved 2023-01-09, from https://doi.org/10.1007/s00190-019-01228-y doi: 10.1007/s00190-019-01228-y

Peltier, R. W., Argus, D. F., & Drummond, R. (2018). Comment on "An Assessment of the ICE-6G_c (VM5a) Glacial Isostatic Adjustment Model" by Purcell et al. *Journal of Geophysical Research: Solid Earth*, *123*(2), 2019–2028. Retrieved 2023-01-25, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2016JB013844 doi: 10.1002/2016JB013844

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*(16). Retrieved 2022-12-02, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2012GL052495 doi: 10.1029/2012GL052495

Ray, R. D., Luthcke, S. B., & Dam, T. v. (2013). Monthly Crustal Loading Corrections for Satellite Altimetry. *Journal of Atmospheric and Oceanic Technology*, *30*(5), 999–1005. Retrieved 2023-01-09, from https://journals.ametsoc.org/view/journals/atot/30/5/jtech-d-12-00152_1.xml doi: 10.1175/JTECH-D-12-00152.1

Reager, J. T., Thomas, B. F., & Famiglietti, J. S. (2014). River basin flood potential inferred using GRACE gravity observations at several months lead time. *Nature Geoscience*, *7*(8), 588–592. Retrieved 2023-01-09, from https://www.nature.com/articles/ngeo2203 doi: 10.1038/ngeo2203

Richey, A. S., Thomas, B. F., Lo, M.-H., Reager, J. T., Famiglietti, J. S., Voss, K., ... Rodell, M. (2015). Quantifying renewable groundwater stress with GRACE. *Water Resources Research*, *51*(7), 5217–5238. Retrieved 2022-09-23, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2015WR017349 doi: 10.1002/2015WR017349

Richter, H. M. P., Lück, C., Klos, A., Sideris, M. G., Rangelova, E., & Kusche, J. (2021). Reconstructing GRACE-type time-variable gravity from the Swarm satellites. *Scientific Reports*, *11*(1), 1117. Retrieved 2022-07-08, from https://www.nature.com/articles/s41598-020-80752-w doi: 10.1038/s41598-020-80752-w

Risbey, J., Pook, M., McIntosh, P., Wheeler, M., & Hendon, H. (2009). On the Remote Drivers of Rainfall Variability in Australia. *Monthly Weather Review - MON WEATHER REV*, *137*, 3233–3253. doi: 10.1175/2009MWR2861.1

Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., & Lo, M.-H. (2018). Emerging trends in global freshwater availability. *Nature*, *557*(7707), 651–659. Retrieved 2022-10-13, from https://www.nature.com/articles/s41586-018-0123-1 doi: 10.1038/s41586-018-0123-1

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., ... Toll, D. (2004). The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society*, *85*(3), 381–394. Retrieved 2023-01-09, from https://journals.ametsoc.org/view/journals/bams/85/3/bams-85-3-381.xml doi: 10.1175/BAMS-85-3-381

Save, H. (2020). *"CSR GRACE and GRACE-FO RL06 Mascon Solutions v02.* doi: doi:10.15781/cgq9-nh24

Save, H., Bettadpur, S., & Tapley, B. D. (2016). High-resolution CSR GRACE RL05 mascons. *Journal of Geophysical Research: Solid Earth*, *121*(10), 7547–7569. Retrieved 2022-12-14, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2016JB013007 doi: 10.1002/2016JB013007

Scanlon, B. R., Zhang, Z., Save, H., Sun, A. Y., Müller Schmied, H., van Beek, L. P. H., ... Bierkens, M. F. P. (2018). Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. *Proceedings of the National Academy of Sciences*, *115*(6). Retrieved 2022-11-29, from https://www.pnas.org/doi/full/10.1073/pnas.1704665115 doi: 10.1073/pnas.1704665115

Schumacher, M., Forootan, E., van Dijk, A. I. J. M., Müller Schmied, H., Crosbie, R. S., Kusche, J., & Döll, P. (2018). Improving drought simulations within the Murray-Darling Basin by combined calibration/assimilation of GRACE data into the WaterGAP Global Hydrology Model. *Remote Sensing of Environment*, *204*, 212–228. Retrieved 2022-11-28, from https://www.sciencedirect.com/science/article/pii/S0034425717304923 doi: 10.1016/j.rse.2017.10.029

Seoane, L., Nastula, J., Bizouard, C., & Gambis, D. (2009). The use of gravimetric data from GRACE mission in the understanding of polar motion variations. *Geophysical Journal International*, *178*(2), 614–622. Retrieved 2023-01-09, from https://doi.org/10.1111/j.1365-246X.2009.04181.x doi: 10.1111/j.1365-246X.2009.04181.x

Sośnica, K., Jäggi, A., Meyer, U., Thaller, D., Beutler, G., Arnold, D., & Dach, R. (2015). Time variable Earth's gravity field from SLR satellites. *Journal of Geodesy*, *89*(10), 945–960. Retrieved 2023-01-09, from https://doi.org/10.1007/s00190-015-0825-1 doi: 10.1007/s00190-015-0825-1

Spinoni, J., Naumann, G., Vogt, J., & Barbosa, P. (2015b). European drought climatologies and trends based on a multi-indicator approach. *Global and Planetary Change*, *127*, 50–57. Retrieved 2023-01-09, from https://www.sciencedirect.com/science/article/pii/S0921818115000284 doi: 10.1016/j.gloplacha.2015.01.012

Spinoni, J., Naumann, G., Vogt, J. V., & Barbosa, P. (2015a). The biggest drought events in Europe from 1950 to 2012. *Journal of Hydrology: Regional Studies*, *3*, 509–524. Retrieved 2023-01-09, from https://www.sciencedirect.com/science/article/pii/S2214581815000026 doi: 10.1016/j.ejrh.2015.01.001

Springer, A., Karegar, M. A., Kusche, J., Keune, J., Kurtz, W., & Kollet, S. (2019). Evidence of daily hydrological loading in GPS time series over Europe. *Journal of Geodesy*, *93*(10), 2145–2153. Retrieved 2023-01-09, from https://doi.org/10.1007/s00190-019-01295-1 doi: 10.1007/s00190-019-01295-1

Sun, A. Y., Scanlon, B. R., Save, H., & Rateb, A. (2021). Reconstruction of GRACE Total Water Storage Through Automated Machine Learning. *Water Resources Research*, *57*(2). Retrieved 2022-07-07, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2020WR028666 doi: 10.1029/2020WR028666

Sun, A. Y., Scanlon, B. R., Zhang, Z., Walling, D., Bhanja, S. N., Mukherjee, A., & Zhong, Z. (2019). Combining Physically Based Modeling and Deep Learning for Fusing GRACE Satellite Data: Can We Learn From Mismatch? *Water Resources Research*, *55*(2), 1179–1195. Retrieved 2022-07-25, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2018WR023333 doi: 10.1029/2018WR023333

Sun, Z., Long, D., Yang, W., Li, X., & Pan, Y. (2020). Reconstruction of GRACE Data on Changes in Total Water Storage Over the Global Land Surface and 60 Basins. *Water Resources Research*, *56*(4). Retrieved 2022-07-08, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2019WR026250 doi: 10.1029/2019WR026250

Swenson, S. C., & Lawrence, D. M. (2014). Assessing a dry surface layer-based soil resistance parameterization for the Community Land Model using GRACE and FLUXNET-MTE data. *Journal of Geophysical Research: Atmospheres*, *119*(17), 10,299–10,312. Retrieved 2023-01-09, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2014JD022314 doi: 10.1002/2014JD022314

Tang, S., Wang, H., Feng, Y., Liu, Q., Wang, T., Liu, W., & Sun, F. (2021). Random Forest-Based Reconstruction and Application of the GRACE Terrestrial Water Storage Estimates for the Lancang-Mekong River Basin. *Remote Sensing*, *13*(23), 4831. Retrieved 2022-07-07, from https://www.mdpi.com/2072-4292/13/23/4831 doi: 10.3390/rs13234831

Tangdamrongsub, N., Han, S.-C., Yeo, I.-Y., Dong, J., Steele-Dunne, S. C., Willgoose, G., & Walker, J. P. (2020). Multivariate data assimilation of GRACE, SMOS, SMAP measurements for improved regional soil moisture and groundwater storage estimates. *Advances in Water Resources*, *135*, 103477. Retrieved 2023-01-09, from https://www.sciencedirect.com/science/article/pii/S0309170819302970 doi: 10.1016/j.advwatres.2019.103477

Tangdamrongsub, N., Steele-Dunne, S. C., Gunter, B. C., Ditmar, P. G., & Weerts, A. H. (2015). Data assimilation of GRACE terrestrial water storage estimates into a regional hydrological model of the Rhine River basin. *Hydrology and Earth System Sciences*, *19*(4), 2079–2100. Retrieved 2023-01-09, from https://hess.copernicus.org/articles/19/2079/2015/ doi: 10.5194/hess-19-2079-2015

Tapley, B. D., Watkins, M. M., Flechtner, F., Reigber, C., Bettadpur, S., Rodell, M., … Velicogna, I. (2019). Contributions of GRACE to understanding climate change. *Nature Climate Change*, *9*(5), 358–369. Retrieved 2022-07-25, from http://www.nature.com/articles/s41558-019-0456-2 doi: 10.1038/s41558-019-0456-2

Towner, J., Ficchí, A., Cloke, H. L., Bazo, J., Coughlan de Perez, E., & Stephens, E. M. (2021). Influence of ENSO and tropical Atlantic climate variability on flood characteristics in the Amazon basin. *Hydrology and Earth System Sciences*, *25*(7), 3875–3895. Retrieved 2022-11-30, from https://hess.copernicus.org/articles/25/3875/2021/ doi: 10.5194/hess-25-3875-2021

Trenberth, K. E. (1990). Recent Observed Interdecadal Climate Changes in the Northern Hemisphere. *Bulletin of the American Meteorological Society*, *71*(7), 988–993. Retrieved 2023-05-04, from https://journals.ametsoc.org/view/journals/bams/71/7/1520-0477_1990_071_0988_roicci_2_0_co_2.xml doi: 10.1175/1520-0477(1990)071⟨0988:ROICCI⟩2.0.CO;2

van Dijk, A. I. J. M., Beck, H. E., Crosbie, R. S., de Jeu, R. A. M., Liu, Y. Y., Podger, G. M., ... Viney, N. R. (2013). The Millennium Drought in southeast Australia (2001–2009): Natural and human causes and implications for water resources, ecosystems, economy, and society. *Water Resources Research*, *49*(2), 1040–1057. Retrieved 2022-11-28, from `https://onlinelibrary.wiley.com/doi/abs/10.1002/wrcr.20123` doi: 10.1002/wrcr.20123

Voss, K. A., Famiglietti, J. S., Lo, M., de Linage, C., Rodell, M., & Swenson, S. C. (2013). Groundwater depletion in the Middle East from GRACE with implications for transboundary water management in the Tigris-Euphrates-Western Iran region. *Water Resources Research*, *49*(2), 904–914. Retrieved 2022-11-29, from `https://onlinelibrary.wiley.com/doi/abs/10.1002/wrcr.20078` doi: 10.1002/wrcr.20078

Wu, X., & Heflin, M. B. (2015). A global assessment of accelerations in surface mass transport. *Geophysical Research Letters*, *42*(16), 6716–6723. Retrieved 2022-11-22, from `https://onlinelibrary.wiley.com/doi/abs/10.1002/2015GL064941` doi: 10.1002/2015GL064941

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*(8), 1629–1637. Retrieved 2022-11-30, from `https://www.sciencedirect.com/science/article/pii/S0034425710000647` doi: 10.1016/j.rse.2010.02.005

Yang, X., Tian, S., Feng, W., Ran, J., You, W., Jiang, Z., & Gong, X. (2020). Spatio-Temporal Evaluation of Water Storage Trends from Hydrological Models over Australia Using GRACE Mascon Solutions. *Remote Sensing*, *12*(21). Retrieved 2022-11-28, from `https://www.mdpi.com/2072-4292/12/21/3578` doi: 10.3390/rs12213578

Yu, Q., Wang, S., He, H., Yang, K., Ma, L., & Li, J. (2021). Reconstructing GRACE-like TWS anomalies for the Canadian landmass using deep learning and land surface model. *International Journal of Applied Earth Observation and Geoinformation*, *102*. Retrieved 2022-07-07, from `https://www.sciencedirect.com/science/article/pii/S0303243421001112` doi: 10.1016/j.jag.2021.102404

Zaitchik, B. F., Rodell, M., & Reichle, R. H. (2008). Assimilation of GRACE Terrestrial Water Storage Data into a Land Surface Model: Results for the Mississippi River Basin. *Journal of Hydrometeorology*, *9*(3), 535–548. Retrieved 2023-01-09, from `https://journals.ametsoc.org/view/journals/hydr/9/3/2007jhm951_1.xml` doi: 10.1175/2007JHM951.1

Zhang, D., Zhang, Q., Werner, A. D., & Liu, X. (2016). GRACE-Based Hydrological Drought Evaluation of the Yangtze River Basin, China. *Journal of Hydrometeorology*, *17*(3), 811–828. Retrieved 2022-07-25, from `https://journals.ametsoc.org/view/journals/hydr/17/3/jhm-d-15-0084_1.xml` doi: 10.1175/JHM-D-15-0084.1

Zhao, M., A, G., Velicogna, I., & Kimball, J. S. (2017). A Global Gridded Dataset of GRACE Drought Severity Index for 2002–14: Comparison with PDSI and SPEI and a Case Study of the Australia Millennium Drought. *Journal of Hydrometeorology*, *18*(8), 2117–2129. Retrieved 2023-01-09, from `https://journals.ametsoc.org/view/journals/hydr/18/8/jhm-d-16-0182_1.xml` doi: 10.1175/JHM-D-16-0182.1

Śliwińska, J., Nastula, J., & Wińska, M. (2021). Evaluation of hydrological and cryospheric angular momentum estimates based on GRACE, GRACE-FO and SLR data for their contributions to polar motion excitation. *Earth, Planets and Space*, *73*(1), 71. Retrieved 2023-03-14, from `https://doi.org/10.1186/s40623-021-01393-5` doi: 10.1186/s40623-021-01393-5

Appendix A Data processing

The functional relationship to derive trend, acceleration, annual amplitude and phase reads

$$\vec{f}(\vec{t}) = a_0 + a_1(\vec{t} - t_0) + a_2(\vec{t} - t_0)^2 + c_1 \cos(2\pi(\vec{t} - t_0)) + s_1 \sin(2\pi(\vec{t} - t_0)), \tag{A1}$$

where $\vec{x} = [a_0, a_1, a_2, c_1, s_1]^T$ are the parameters estimated in the least square process and t_0 is a reference epoch. For this study the data has been referenced to the time of the first observation. As HG19 does not include a trend the equation above reduces to

$$\vec{f}(\vec{t}) = a_0 + a_2(\vec{t} - t_0)^2 + c_1 \cos(2\pi(\vec{t} - t_0)) + s_1 \sin(2\pi(\vec{t} - t_0)). \tag{A2}$$

As trend and acceleration estimated by the least squares approach differ, depending on whether parameters are estimated separately or joinly, we first estimated the trend for the other data sets, reduced them by the estimated trend and then derived acceleration, annual amplitude and phase. The annual amplitude and phase are computed based on the estimated coefficients c_1, s_1. For better interpretability the annual phase is expressed as the day of the signal minimum via

$$D_{\mathrm{Min}} = \frac{\arctan(s_1, c_1) + \pi}{2\pi} 365, 25. \tag{A3}$$

The annual amplitude is computed as

$$A = \sqrt{s_1^2 + c_1^2} \tag{A4}$$

We define interannual signal as all signal components with a period longer than a year. The interannual signal was derived based on a combination of filtering and least squares adjustment. In a first step the annual and semi-annual periods are computed and reduced from the data sets based on a least square adjustment fitting annual and semi-annual coefficients to the time series. A low pass filter is applied to reduce all frequencies greater than 1. The filter weights are derived for an ideal lowpass filter, which is defined as a filter that allows all frequencies greater than a given cutting frequency (in our case > 1 year) to pass and set all frequencies smaller then the cutting frequency (< 1 year) to zero. The derived coefficients, $c_{|k|}$, reads

$$c_{|k|} = c_{|k|}\sigma_{|k|}^N = \sum_{k=0}^{N} \frac{\nu_{\mathrm{stop}}}{\nu_N} \mathrm{sinc}\left(k\frac{\nu_{\mathrm{stop}}}{\nu_N}\right) \mathrm{sinc}\left(\frac{k}{N+1}\right) \tag{A5}$$

The filter length, N, was set to 13 months and the cutting frequency ν_{stop} to 1 year. The derived filter coefficients are finite. A Lanczos smoothing hamming$_{d}igital_{1}989, heredenotedas$ $\sigma_{|k|}^N$ is used to smooth the signal around the cutting frequency ν_{stop}.

We define subseasonal to be all signal parts < 1 year. The dominant annual signal is removed using a least square adjustment. All signals > 1 year are reduced using a high pass filter. Assuming, the input signal,$\{u_n\}_{\Delta x}$, can be written as $\{u_n\}_{\Delta x} = \{y_n\}_{\Delta x} + \{z_n\}_{\Delta x}$ with $\{y_n\}_{\Delta x}$ containing the low frequencies, i.e the output of the low pass filter, and $\{z_n\}_{\Delta x}$ the high frequencies, the high pass filter can be derived as $\{z_n\}_{\Delta x} = \{u_n\}_{\Delta x} - \{y_n\}_{\Delta x}$.

It should be mentioned, that the filtering operation, even though effectively reducing all unwanted frequencies also damps signal we are interested in. The frequency operation might also cause leakage, so a smearing of frequencies in others.

Appendix B Different reconstructions of GRACE like TWSA

Table Appendix B.1 gives an overview of different GRACE like TWSA reconstructions, including reconstruction period, region and employed methods.

Authors	Predictors	employed methods	Time period	Area	Data access
Humphrey and Gudmundsson (2019)	P, T	exponential first order decay function	01.1979 – 07.2019	global	`https://doi.org/10.6084/m9.figshare.7670849`
F. Li et al. (2021)	P, T, SST, climate indices	PCA,ICA,STL, LS, ANN; MLR, ARX	07.1979 – 06.2020	global	`https://doi.org/10.5061/dryad.z612jm6bt`
A. Y. Sun et al. (2021)	P, T, SST, NAO, MEI,GLDAS TWS	AutoML	06.2017 – 12.2019	conterminous U.S (CONUS)	on request
Forootan et al. (2020)	Swarm	ICA	2002 – 2020	global	`https://www.mdpi.com/2072-4292/12/10/1639/s1`
Z. Sun et al. (2020)	P, T, GLDAS Noah TWS	DNN, MLR, SARIMAX	04.2002 – 06.2018	60 basins	on request
Tang et al. (2021)	GLDAS Noah TWS, P, T, meteorological data	RF	1980 – 2014	Lancang-Mekong River Basin	on request
Yu et al. (2021)	EALCO TWSA	CNN, cGAN, DCAE, ConvL-STM	1979 – 2002	Canadian landmass	on request
Lenczuk et al. (2022)		forward and backward AR process	07.2017 – 05.2018	global	on request
A. Y. Sun et al. (2019)	GLDAS Noah TWS	CNN	04.2002 – 06.2017	India	on request
Ferreira et al. (2019)	T, P, E, SM, R, climate indices	NARX	1979 – 2013	West Africa	on request
Becker et al. (2011)	in situ river level records	PCA / EOF	1980 – 2008	Amazon basin	on request
Long et al. (2014)	SMS, P, T	MLP ANN	02.1979 – 12.2002	karst Plateau basin	on request
Richter et al. (2021)	SWARM	PCA	07.2017 – 05.2018	global	on request
Zhang et al. (2016)	SM, P	ANN	1979 – 2012	Yangtze river basin	on request

Table Appendix B.1: Overview: Different reconstructions of GRACE like TWSA

Note: P precipation, T land surface temperature, SST sea surface temperature, NAO North Atlantic Oscillation, MEI Multivariate ENSO index, E evaporation, SM soil moisture, R rainfall, NDVI normalized difference vegetation index, M mascon, SH spherical harmonics, PCA principal component analysis, ICA indipentend component analysis, STL seasonal-trend decomposition based on loess, LS least square, ANN artificial neural network, MLR multi linear regression, ARX autoregressive exogenous model, AutoML automated machine learning, DNN deep neural network, SARIMAX seasonal ARIMA (autoregressiv integrated moving average) with exogenous variables, RF random forest, CNN convolutional neural network, cGAN conditional generative adversarial network, DCAE deep convolutional autoencoder, ConvLSTM convolutional long short term memory, AR (process) auto regressive (process), NARX nonlinear autoregressive with exogenous input, MLP multi-layer percepton

Appendix C Trend, acceleration and annual amplitude and phase for the reconstructions, WaterGAP and GRACE

Table Appendix C.1 shows the derived average trend, acceleration, average annual phase and minimal day of the seasonal cycle for the reocntructions and the hydrological model WaterGAP for 1979-2016 for the global land. Table Appendix C.2 reveal the same, but for the Pre GRACE time frame. Average trend, average rate range change, annual amplitude and minimal day of the seasonal cycle for the GRACE time frame for four different GRACE solutions and the reconstructions for the global land are shown in table Appendix C.3. Figure Appendix C.1 present the respective time series.

Appendix D Long term evolution of TWS

Figure Appendix D.1 reveal the water mass change for the detrended and deseasonalized reconstructions for 1979-2016 for nine major river basins. The time series of HG19 is presented in red, Li21 in black. The according location of the basins are shown in figure Appendix D.2. The time series are discussed in section 4.3.

Appendix E Supplementary results

The average trend, acceleration, average annual amplitude and phase, subseasonal and interannual signal variability of water storage for Europe is shown in figure Appendix E.1, for the years 1979-2016, and in figure Appendix E.2 for the pre GRACE era. The results are discussed in section 3.3.3. Figure Appendix E.3 reveals the autocorrelation pattern of the detrended and deseasonalized reconstructions for a point in the Amazon basin. Figure Appendix E.4 shows the same, but for a point in Europe. The figures are discussed in section 3.3.2 and 3.3.2.

	Trend* $[mm/yr]$	Trend$^{\diamond}$ $[mm/yr]$	Acceleration$^{\diamond}$ $[mm/yr^2]$	annual amplitude$^{*\diamond}$ $[mm]$	annual phase$^{*\diamond}$ $[d]$
Li21	-1.59	-1.63	0.001	24.29	83
HG19	-0.57	-0.0135	-0.0137	21.95	266
WaterGAP	-0.69	-0.56	-0.003	23.85	246

Table Appendix C.1. Trend, acceleration and annual amplitude and annual phase derived from the two reconstructions and the hydrological model WaterGAP for the global land for the time frame 1979 – 2016, without (1st column) and with (2nd column) estimating acceleration

	Trend* $[mm/yr]$	Trend$^{\diamond}$ $[mm/yr]$	Acceleration$^{\diamond}$ $[mm/yr^2]$	annual amplitude$^{*\diamond}$ $[mm]$	annual phase$^{*\diamond}$ $[d]$
Li21	-1.62	-2.61	0.04	24.29	83
HG19	-0.04	0.57	-0.03	21.73	267
WaterGAP	-0.65	-1.22	0.03	23.85	246

Table Appendix C.2. Trend, acceleration and annual amplitude and annual phase derived from the two reconstructions and the hydrological model WaterGAP for the global land for the time frame 1979 – 2002, without (1st column) and with (2nd column) estimating acceleration

	Trend* $[mm/yr]$	Trend$^{\diamond}$ $[mm/yr]$	Acceleration$^{\diamond}$ $[mm/yr^2]$	annual amplitude$^{*\diamond}$ $[mm]$	annual phase$^{*\diamond}$ $[d]$
ITSG SH 2018	-0.22	0.63	-0.04	18.58	180
CSR SH RL06	-0.15	0.69	-0.04	18.5	181
GFZ SH RL06	-0.15	0.59	-0.04	13.31	180
CSR M RL06, v.2	-1.76	-1.13	-0.03	24.30	173
Li21	-1.66	-0.62	-0.06	24.30	174
HG19	-0.16	0.61	-0.04	22.30	205

Table Appendix C.3. Trend, acceleration and annual amplitude and annual phase derived from different GRACE/GRACE-FO solutions and the reconstructions for the global land for the time frame 2002 – 2020, without (1st column) and with (2nd column) estimating acceleration

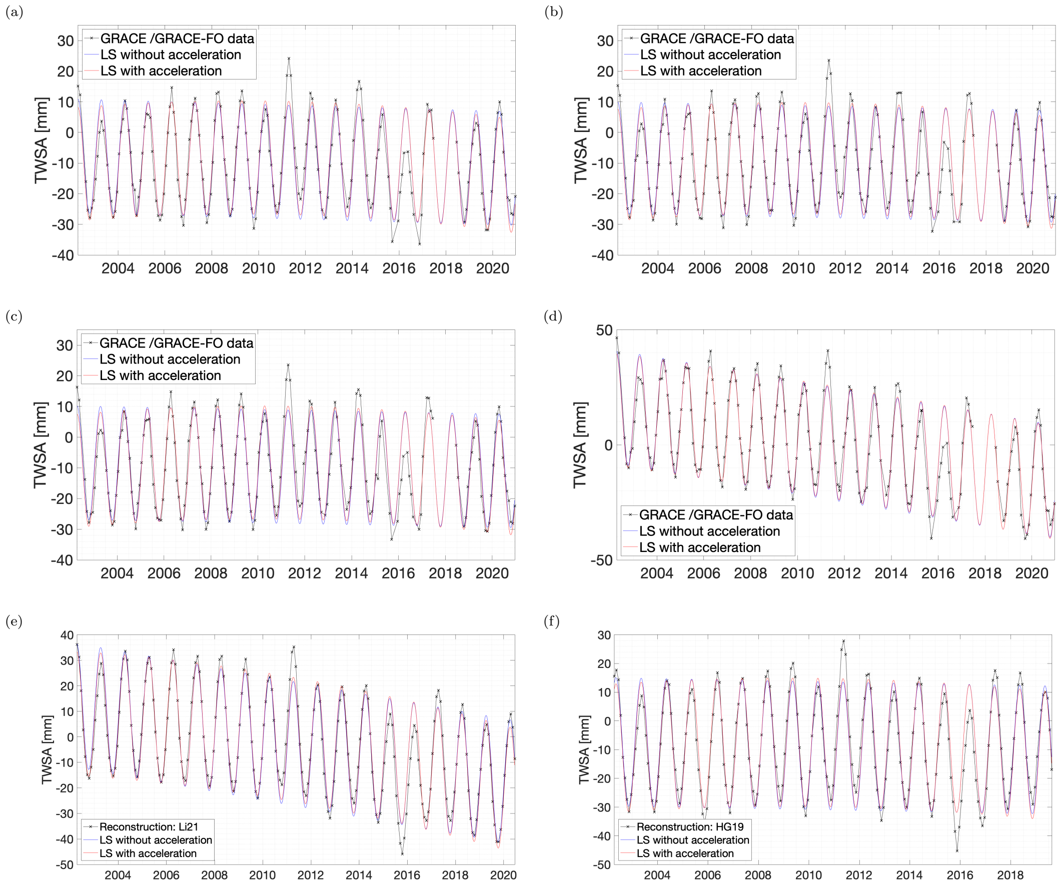

Figure Appendix C.1. Time series for four different GRACE solutions and the two global reconstructions for the global land for the time frame 2002 – 2020

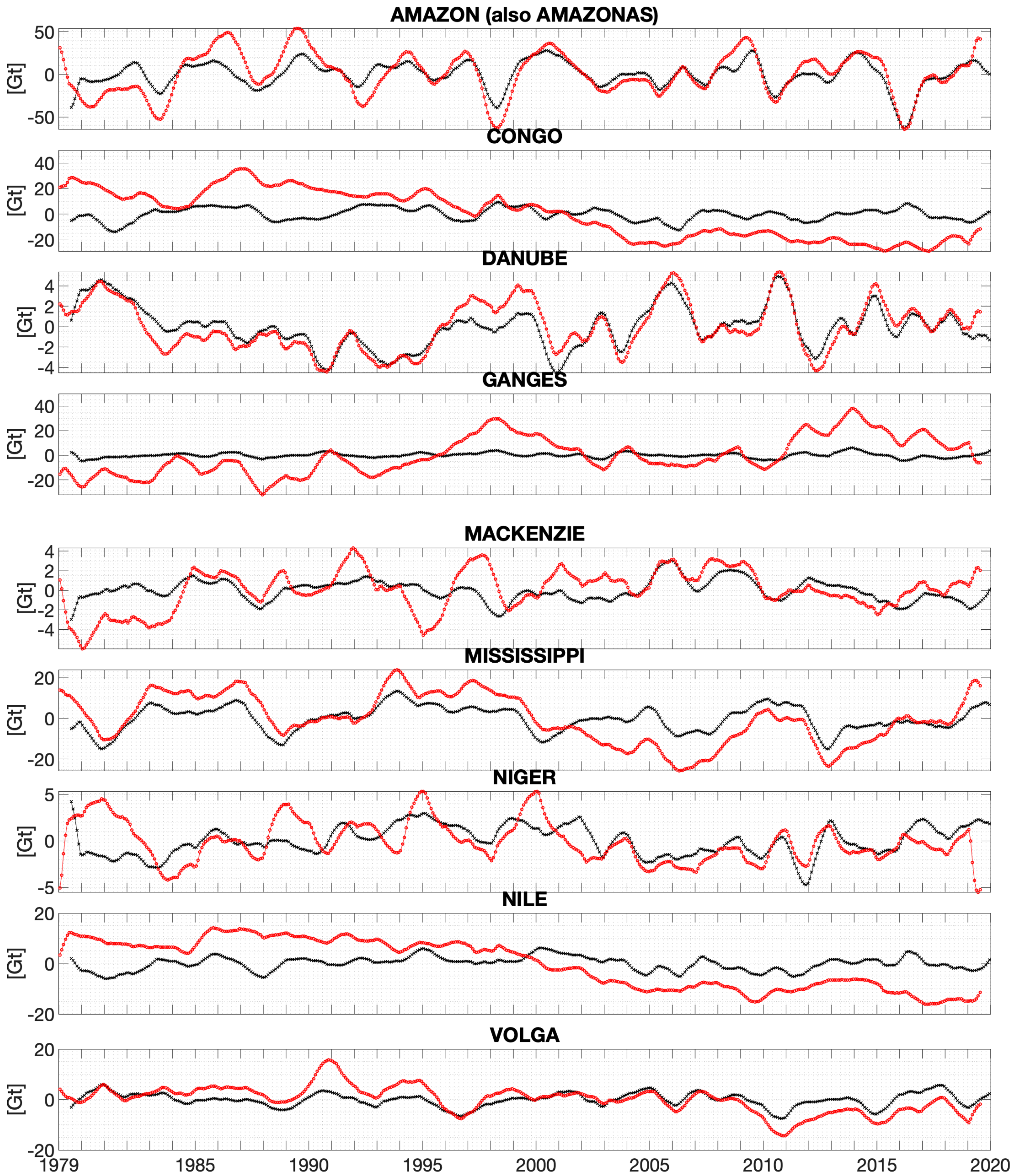

Figure Appendix D.1. Detrended time series for Li21 (black crosses) and HG19 (red circles) for the Amazon, Mississippi, Mackenzie, Volga, Danube, Ganges, Nile, Niger, Congo basin. The seasonal signal is reduced using a 12 month moving average filter.

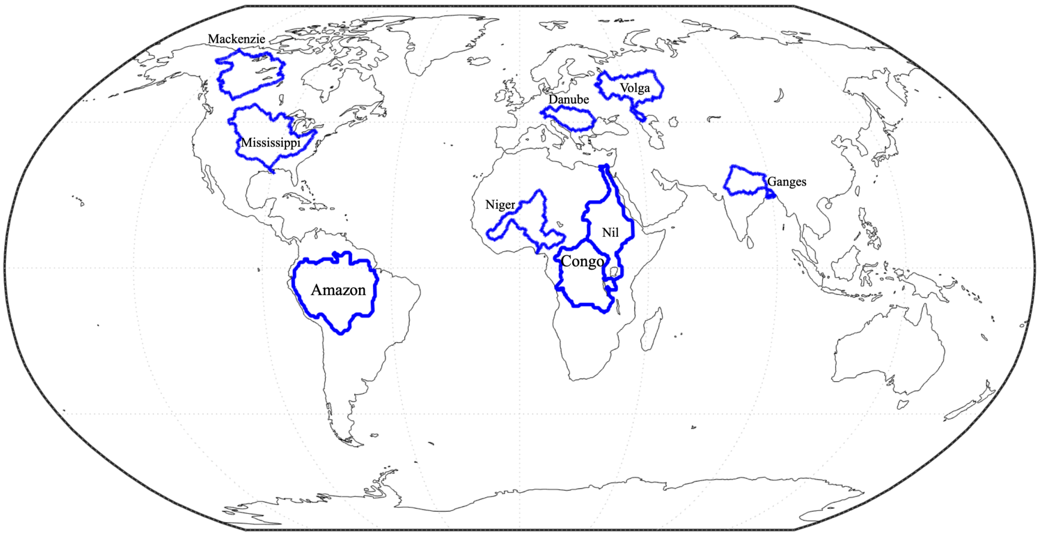

Figure Appendix D.2. Overview: Selected basins

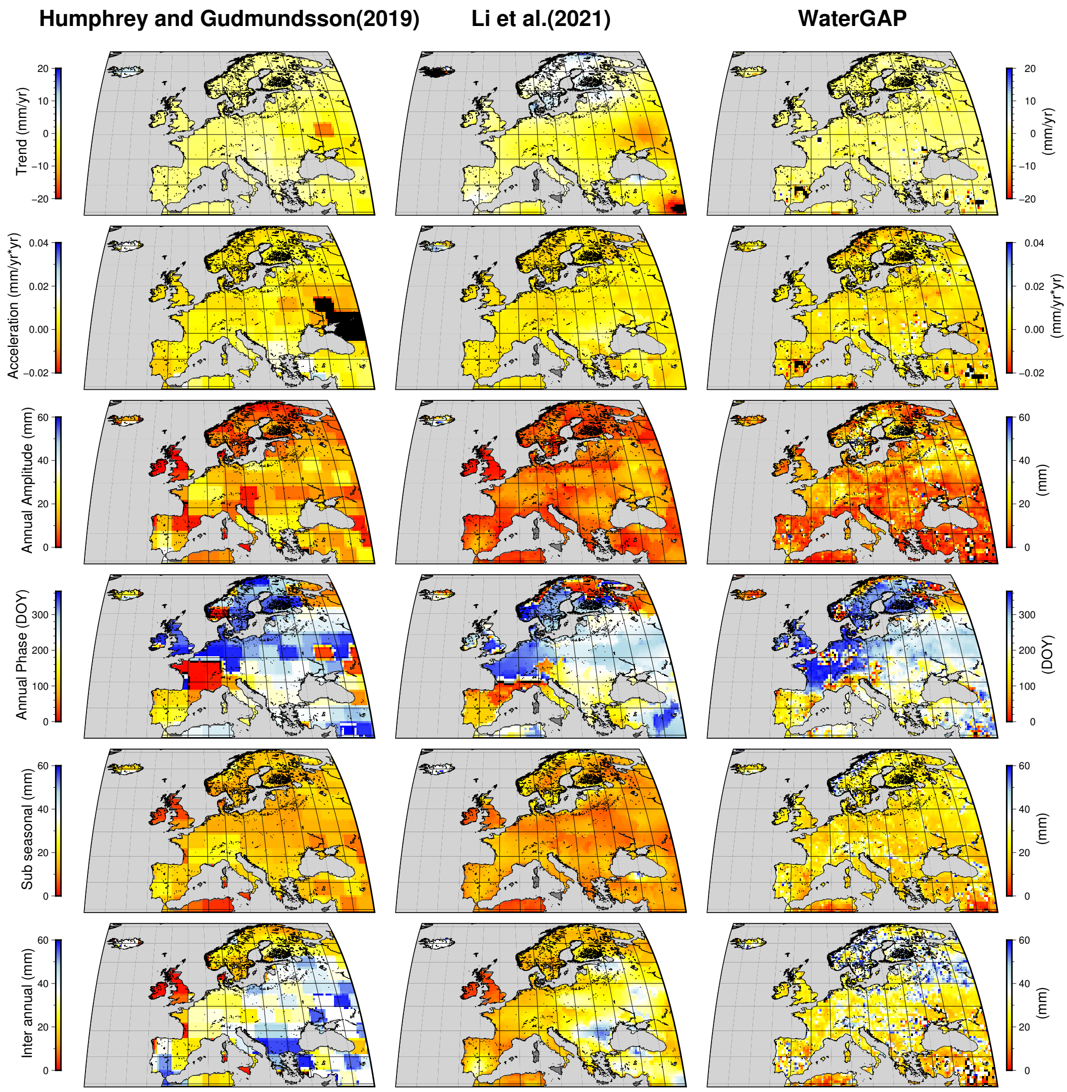

Figure Appendix E.1. Trend, acceleration, annual amplitude and phase, subseasonal and interannual signal for the years 1979-2016 for the reconstructions and WaterGAP for Europe

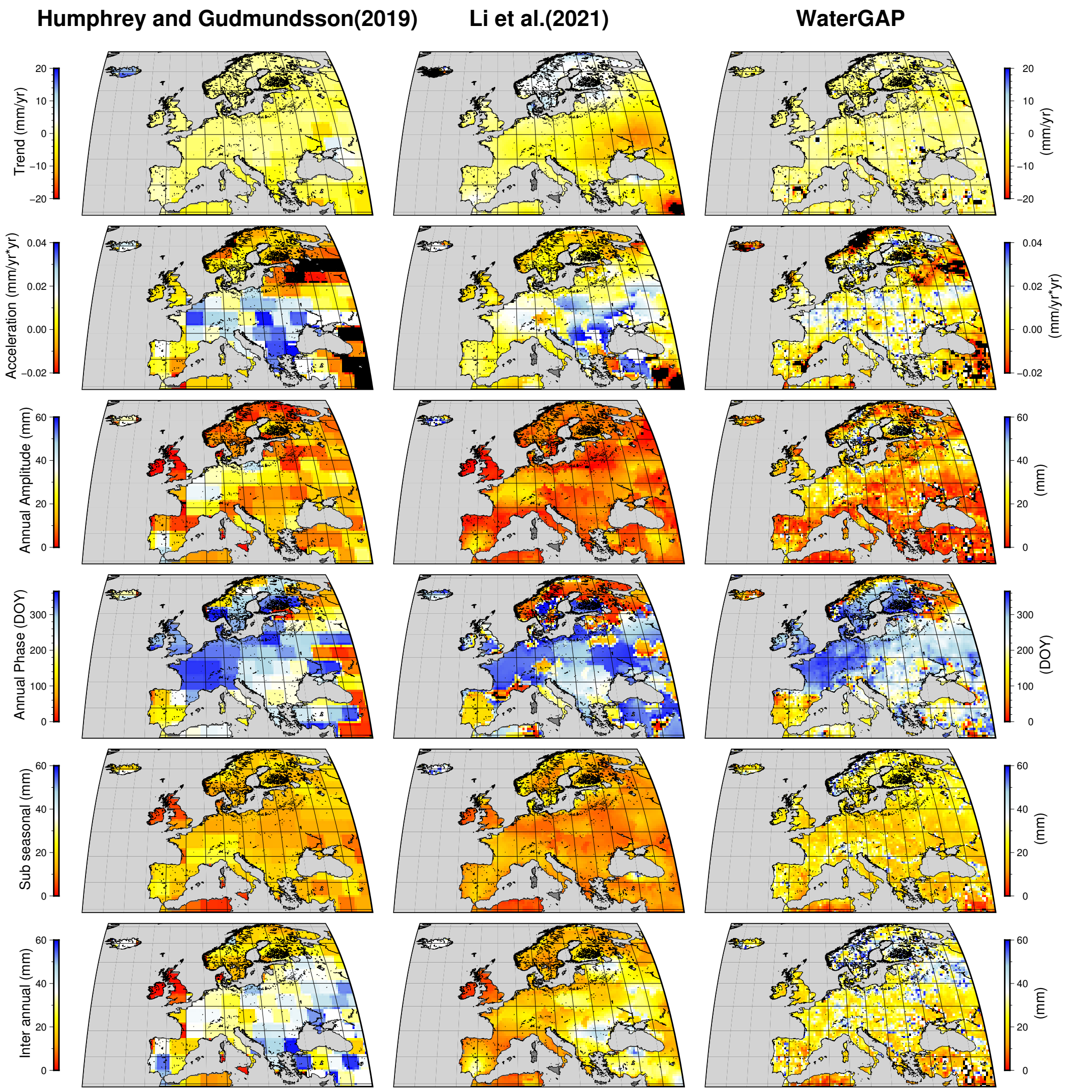

Figure Appendix E.2. Trend, acceleration, annual amplitude and phase, subseasonal and interannual signal for the years 1979-2002 for the reconstructions and WaterGAP for Europe

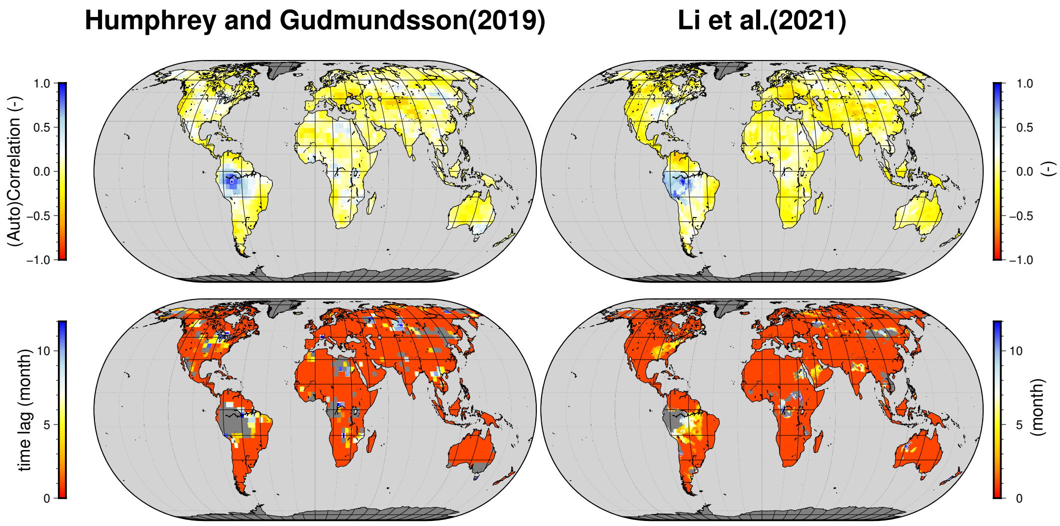

Figure Appendix E.3. *top:* Autocorrelation of the detrended and deseasonalized reconstructions. *bottom:* Time Lag. The point is located in the Amazon basin.

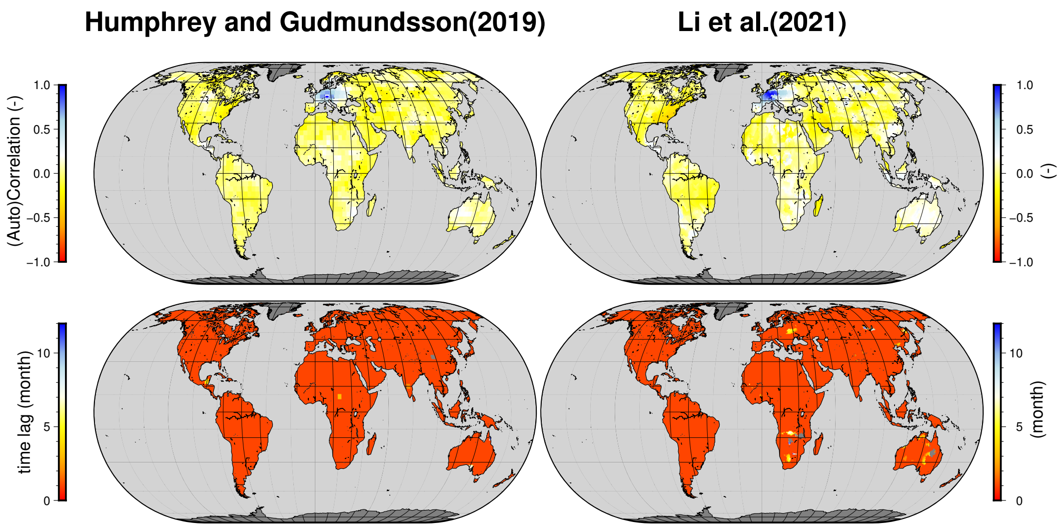

Figure Appendix E.4. *top:* Autocorrelation of the detrended and deseasonalized reconstructions. *bottom:* Time Lag. The point is located in Germany.

How realistic are multi-decadal reconstructions of GRACE-like total water storage anomalies?

C. Hacker[1], J.Kusche[1]

[1]Institute for Geodesy and Geoinformation, University of Bonn, Bonn, Germany.

Key Points:

- Analysis of signal components of pre-GRACE (1979-2002) and long-term (1979-2016) terrestrial water storage based on global GRACE like TWSA reconstructions
- The low frequency part of the reconstructions are evaluated by low degree gravity fields from satellite laser ranging
- The reconstructions reveal similar regions affected by water storage changes over the last four decades.

Corresponding author: Charlotte Hacker, `hacker@geod.uni-bonn.de`

Abstract

[The Gravity Recovery and Climate Experiment (GRACE) mission has monitored total water storage anomalies (TWSA) globally with unprecedented resolution and accuracy since 2002. However, many applications require a data-based, multi-decadal extended record of TWSA prior to the GRACE period and for bridging the eleven-months gap between GRACE and its successor GRACE Follow-On (GRACE-FO), that does not depend on hydrological modelling. Statistical and machine-learning 'reconstruction' approaches have been developed to this end, mostly via identifying relations of GRACE-derived TWSA to climate variables, and some regional or global land data sets are now publicly available.

In this contribution, we compare the two global reconstructions by Humphrey and Gudmundsson (2019) and F. Li et al. (2021) mutually and against output from the water Global Analysis and Prognosis (WaterGAP) hydrological model from 1979 onwards, against large-scale mass-change derived from geodetic satellite laser ranging (SLR) from 1992 onwards, and finally against differing GRACE/-FO solutions from 2002 onwards.

We find that the reconstructions agree surprisingly well in many regions at seasonal and sub-seasonal timescales, even in the pre-GRACE era. We find larger differences at inter-annual timescales which we speculate are in part due to the way reconstructions are trained, and in part on which specific GRACE solution they are trained as well as the climatological characteristics of the region. Our comparison against independent SLR data reveals that reconstructions (only) partially succeed in representing anomalous TWSA for regions that are influenced by large climate modes such as El Niño-Southern Oscillation (ENSO).]

Plain Language Summary

Water is a life sustaining resource, crucial for human survival, agricultural and economical proposes. Since 2002, the Gravity Recovery and Climate Experiment (GRACE) mission monitors total water storage anomalies (TWSA) on a global scale, allowing the analysis of temporal changes in the water cycle. However, the time series is limited to 20 years. Many data analysis applications require a data-based, multi-decadal extended record of TWSA prior to the GRACE period. Reconstructions are directly "build" based on the GRACE observation, finding a relationship between GRACE-derived TWSA and climate variables.

In this contribution, we compare the two global reconstructions by Humphrey and Gudmundsson (2019) and F. Li et al. (2021) mutually and against output from the water Global Analysis and Prognosis (WaterGAP) hydrological model from 1979 onwards, against large-scale mass-change derived from geodetic satellite laser ranging (SLR) from 1992 onwards, and finally against differing GRACE/-FO solutions from 2002 onwards.

We find the reconstructions reveal similar regions affected by water storage changes over the last four decades, especially for basins with strong TWSA signals like the Amazon. Our comparison against independent SLR data reveals that reconstructions (only) partially succeed in representing anomalous TWSA for regions that are influenced by large climate modes such as El Niño-Southern Oscillation (ENSO).

1 Introduction

The GRACE (Gravity Recovery And Climate Experiment) and GRACE-FO (GRACE Follow-On) satellite missions provide time variable gravity field models and estimates of total water storage anomalies (TWSA) with monthly resolution since 2002 (Tapley et al., 2019). GRACE and GRACE-FO (hereafter GRACE-/FO) data have enabled studying the natural variability of the terrestrial water cycle and its response to radiative forcing, land use, and water withdrawal and redirection (Rodell et al., 2018). TWSA maps have been used to quantify groundwater stress (Richey et al., 2015), large-scale droughts

(Zhao et al., 2017; Gerdener et al., 2020) and floods (Reager et al., 2014; Han et al., 2021), vegetation response (Geruo et al., 2017), and soil processes (Swenson & Lawrence, 2014). In general, the long-term temporal evolution of water storage observed by satellites can be linked to modifications of the land boundary conditions and the resulting climate forcing, the direct and indirect impacts of anthropogenic activities such as groundwater abstraction and land use change, and the hydrological response of the system (Eicker et al., 2016), and regions where trends and accelerations share the same sign may be seen as moving away from the long-term equilibrium of the water cycle. In addition, several studies (Zaitchik et al., 2008; Eicker et al., 2014; Tangdamrongsub et al., 2015; Girotto et al., 2016; Khaki et al., 2017; Schumacher et al., 2018; B. Li et al., 2019; Springer et al., 2019; Tangdamrongsub et al., 2020; Gerdener et al., 2020) assimilated GRACE/-FO TWSA data into hydrological and land surface models, to add realism to model simulations. Apart from these climate and hydrology applications, observation-based TWSA maps are increasingly employed in geodesy e.g. for loading computations for satellite altimetry (Ray et al., 2013) and GNSS (Chanard et al., 2018; Mémin et al., 2020), deriving geocenter estimates (Wu & Heflin, 2015), or simulating the hydrological angular momentum contribution to Earth rotation variations (Seoane et al., 2009; Jin et al., 2010).

However, the use of GRACE/-FO data for e.g. deriving changes in the frequency of water storage extremes (Kusche et al., 2016) or for climate model evaluation (Jensen et al., 2019), is severely hampered by the short duration of the time series. Similarly, it is well-known that trend and acceleration estimates derived within the present GRACE/-FO time period are affected by inter-annual variability (Eicker et al., 2016). In addition, there is an eleven-month gap between GRACE and GRACE-FO, and the GRACE record has several shorter gaps due to battery problems on the spacecraft. All this limits not only the use of GRACE/-FO TWSA data for confronting model simulations, but also the assimilation of TWSA data into models, as assimilation frameworks are usually ill-prepared for data sets with gaps. Unfortunately, no geodetic observing technique provides either time-variable gravity fields or any other variable that could be related to global land TWSA in the pre-GRACE era with a spatial resolution comparable to GRACE.

Data-based 'reconstructions' of gridded total water storage seek to provide substitute data and in this way to overcome many of the issues mentioned above, typically by deriving and training a relationship between the monthly GRACE TWSA maps and predictors for which multidecadal data records are available. The mathematical framework to derive a relationship is either based on regression techniques or on machine learning methods. Predictors are mostly chosen as climate variables that play a dominating role in the terrestrial water budget, such as precipitation or land surface temperature, but they can also include other hydrological or space-geodetic observations. Predictors can be sets of single time series (e.g. ENSO index, modes derived via empirical orthogonal function analysis) or spatial fields.

TWSA reconstruction methods can serve, at the same time, for predicting near real-time or even future total water storage variability (Forootan et al., 2014). This may be useful in diagnosing problems e.g. in quick-look data analysis, in identifying real anomalies, or in forecasting e.g. drought or flood conditions at the seasonal timescale (Reager et al., 2014).

Table Appendix B.1 provides an overview of different reconstructions of total water storage. Becker et al. (2011) were upon the first in reconstructing water storage, for the 1980−2008 time frame. In the Amazon basin, gridded TWSA maps are derived based on the most energetic spatial modes seen by GRACE, as identified via from singular value decomposition, and scaled with in-situ water level data from river gauges. This approach is essentially identical to how the sea level community reconstructs past sea surface height maps from a few decades of radar altimetry and long records from tide gauges (Church et al., 2004). Forootan et al. (2020) and Richter et al. (2021) applied similar methods later at the global scale in order to close the eleven-months gap between GRACE and GRACE-FO with data from the Swarm satellites. Both studies reported a notable decrease in the noise level and an improvement in the spatial resolution by combining Swarm and GRACE.

Forootan (2014) developed an approach for predicting water storage anomalies over West Africa based on low-degree autoregressive TWSA modelling driven by external variables, which were selected from independent or principal components of relevant climate fields. Autoregressive modelling was also used to close the gap between the two GRACE missions by Lenczuk et al. (2022). Their process model does not employ any external predictor data sets and thus uncertainty accumulates quickly; to mitigate this Lenczuk et al. (2022) suggest forward- and backward propagation.

Many recent studies turned to machine learning algorithms (see table Appendix B.1). Long et al. (2014) studied floods and droughts in a karst plateau in China using temporally extended GRACE data. They used an artificial neural network (ANN) to learn the relationship between monthly precipitation, mean temperature and soil moisture derived from the Global Land Data Assimilation System (GLDAS) (Rodell et al., 2004). A similar study was carried out by Zhang et al. (2016) for the Yangtze basin in China, where they used an ANN with precipitation and soil moisture inputs to extend the GRACE time series and characterize drought impacts on the Yangtze. A. Y. Sun et al. (2019) applied a deep convolutional neural network (CNN) to learn the relation between GLDAS and GRACE-derived TWSA for India. The focus of their study was laid on the success of the CNN training: CNNs are mostly used for image classification and trained on big data sets, whereas the GRACE data set with around 170 monthly maps (Tapley et al., 2019) is comparably small. Yu et al. (2021) also used deep learning to reconstruct GRACE-like TWSA for the Canadian landmass, while training with modelled TWSA. Apart from deep learning, Z. Sun et al. (2020) used multiple linear regression (MLR) and seasonal autoregressive integrated moving average with exogenous variable (SARIMAX) to predict water storage changes in around 60 global basins. They found good agreement across the three methods for the reconstructed signals for humid and low-intensity irrigated areas. For drier regions their results showed significant variations, indicating that the performance of an algorithm may depend on hydroclimatic characteristics of the basin. In a second study A. Y. Sun et al. (2021) focused on the GRACE– GRACE-FO data gap. They used six different machine learning algorithms and multiple groups of meteorological and climatic variables to fill the time series over conterminous U.S., and suggest to combine different ML models to provide a final robust estimate.

From a user perspective, there is no clear picture yet emerging as to what approach might show advantages and disadvantages at which temporal and spatial scales, and in which climate regimes. Some of the above mentioned studies were designed to fill the relatively short gap between the GRACE and GRACE-FO missions, while others aimed indeed at multidecadal time series. Except for the studies utilizing Swarm data, all the works mentioned above provide regionally restricted reconstructions.

In this contribution, we therefore focus on (in the following abbreviated as HG19) Humphrey and Gudmundsson (2019) and (Li21 in the following) F. Li et al. (2021), which are, to our knowledge, the only two published reconstructions of total water storage anomalies for the entire global land excluding ice caps and glaciated regions. Both HG19 and Li21 cover more than four decades and here we will look at 1979-2020, later on called the full time frame. Both HG19 and Li21 reconstruct TWSA variations from long-term climate variable records via relations trained in the GRACE time period, but they rely on very different mathematical approaches.

HG19 formulated a model with deterministic and stochastic components to describe the inter-annual variability of water storage change over time. While the deterministic part relates storage to precipitation and temperature via a simple 1D (i.e. grid-cell based) first-order decay model, Humphrey and Gudmundsson (2019) fitted a spatial autoregressive noise model (Cressie, 1993) to spatial and temporal correlation structure in the GRACE TWSA maps, in order to quantify the underlying stochastic process. Markov Chain Monte Carlo is used to achieve a representative error distribution, and the reconstruction is derived as an ensemble consisting of 100 realisations.

In a very different setting, the Li21 framework combines machine learning with time series analysis and statistical decomposition techniques. In a first step the dominant (both statistically independent and orthogonal) modes of GRACE and climate data are identified. Selected relevant modes are then decomposed into linear trend, seasonal, interannual part, and subseasonal using different approaches, and each temporal signal component is reconstructed by either simple neural network, autoregressive modelling with exogenous variables or multilinear regression. The global land is divided into basins, and for each basin the optimal combination of the methods is learned (F. Li et al., 2021).

For evaluating these two global multidecadal reconstructions in the pre-GRACE era, we use output from the hydrological model WaterGAP (Müller Schmied et al., 2020a) and, for the first time, low degree time-variable gravity fields from the geodetic satellite laser raging (SLR) technique (Löcher & Kusche, 2020).

WaterGAP simulates water flows and water storage in canopy, snow, soil, groundwater, lakes, man-made reservoirs, wetlands and rivers (Müller Schmied et al., 2020a), with total water storage being defined over the sum of these compartments. The model differs from many global hydrological or land surface models in its representation of anthropogenic processes; its global water use models determine the water use for irrigation, livestock, domestic, manufacturing and thermal power, then required net water abstraction is partitioned into net abstraction from surface water and groundwater. In the context of comparing to TWSA reconstructions trained on GRACE/-FO data, it is worth mentioning that WaterGAP is however limited in its representation of water abstraction, that it uses rather simple algorithms for simulating reservoir operations, and that it does not simulate at all glacier mass variability. The model has, however, quite often been compared to GRACE/-FO data with favorable results in particular at seasonal and sub-seasonal timescales (Müller Schmied et al., 2020a).

SLR contributes to the International Terrestrial Reference Frame (ITRF) as well as low degree gravity field models (Pearlman et al., 2019). The sensitivity of the SLR satellites to the time variable gravity field ranges from the Earth's center-of-mass to about degree ten (M. Cheng et al., 2011). To increase the spectral resolution while sacrificing some independence of the procedure from GRACE, Löcher and Kusche (2020) fitted, in addition to certain low-degree spherical harmonics, a few empirical orthogonal functions of the recent GRACE solutions to the SLR ranging data. This hybrid approach resulted, after back-transformation of the empirical orthogonal functions (EOF), in monthly sets of spherical harmonic coefficients complete up to degree and order 60.

The paper is organised as follows: In section 2 we present the data sets amd methods that will be employed in this study. Section 3 describes our analysis of the two global land water storage reconstructions L21 and HG19 with respect to each other, and when compared to WaterGAP. For each data set the trend, acceleration, annual amplitude and phase, the interannual signal and the subseasonal signal are discussed for the pre-GRACE era (1979-2002) and for the full reconstruction period (1979-2016). In section 4 the two reconstructions are compared to different GRACE solutions, i.e. not only those they were trained on, and after spectral truncation, to the SLR data by Löcher and Kusche (2020) for the years 1992−2002. An examination of the long term evolution of the water storage for nine major river basins closes this section. The findings of this study are briefly summarized in section 5.

2 Data and Methods

2.1 Data

2.1.1 GRACE/FO data

Three different GRACE/-FO level-2 (L2) solutions, the ITSG2018 series (Mayer-Gürr et al., 2018), the release 06 (RL06) from the Center for Space Research of the University of Texas (CSR) from the website CSR (2018) and the RL06 GFZ Potsdam data (Dahle et al., 2018, 2019), and the mass concentration (mascon) solution provided by

CSR (Save et al., 2016) were used to evaluate the reconstructions. All GRACE/-FO data sets cover the time from April 2002 to December 2020. As no interpolation was applied the GRACE/FO data contain gaps, including the nearly one year period between the two missions.

The spherical harmonic solutions were expanded up to degree and order 96. The degree-one and C_{20} coefficients were replaced as recommended in M. Cheng et al. (2011). The coefficients were smoothed with a DDK3 filter (Kusche, 2007). Time-variable signals due to glacial isostatic adjustment (GIA), i.e. the ongoing viscoelastic uplift of the Earth's crust in response to the melting of the former ice sheets, was corrected for using the model by Peltier et al. (2018).

2.1.2 Global reconstructions of GRACE-type gridded total water storage

The F. Li et al. (2021) data set represents a reconstruction of monthly, total water storage anomalies from 1979−2020 on a 0.5° x 0.5° grid. Their framework combine machine learning techniques with time series and statistical decomposition techniques. F. Li et al. (2021) employed the RL06 monthly CSR GRACE solution as predictand and tested several meteorological variables (see table Appendix B.1) as predictors. In a first step, the dominant modes of the input data were identified via either the Principal Component Analysis (PCA) or Independent Component Analysis (ICA) technique. Selected modes were then partitioned into linear trend, seasonal, and inter-annual , and the residual variability signals using either Least-Squares (LS) or seasonal-trend decomposition based on loess (STL) (Cleveland et al., 1990) methods. Each signal component is reconstructed by either artificial neural network (ANN), autoregressive exogenous model (ARX) or multi-linear regression (MLR) approaches. The global land area is divided into hydrological basins, and for each basin the optimal combination of the framework described above was determined via evaluation against observed GRACE/-FO data.

The Humphrey and Gudmundsson (2019) global reconstruction covers the time period from 1979 – 2019 on a 0.5° x 0.5° grid. The reconstruction is build as consisting of a deterministic and a stochastic part. The first one was formulated as a first order decay model, linking the effect of temperature and precipitation on water storage in a simplified way. To identify and quantify a stochastic process, underlying the spatial and temporal correlation structure seen in the original GRACE solutions, Humphrey and Gudmundsson (2019) employed a spatial autoregressive (SAR) model as in Cressie (1993). A Markov chain Monte Carlo procedure was then used to generate representative sample distributions. Two GRACE solutions, three precipitation and two temperature datasets were used to generate six different reconstructions. Each solution consists of an ensemble of 100 realisations. Here, we used the reconstruction based on the JPL masconcs solution and the ERA5 forcing data (Hersbach et al., 2020). We chose this combination since, according to Humphrey and Gudmundsson (2019) it is the closest to the "true" GRACE solution.

2.1.3 Low degree gravity fields from SLR

SLR is commonly used to derive low degree spherical harmonic coefficients, the geocenter position, station coordinates, and thus significantly contributes to the International Terrestrial Reference Frame (Altamimi et al., 2016; M. K. Cheng et al., 2013). Within the post-processing of GRACE and GRACE-FO data sets the C_{20} oblateness coefficient from the level-1 analysis is regularly replaced by estimates based on SLR, and this has been recommended for C_{30} as well (M. Cheng & Ries, 2023)). SLR results have been also frequently used to validate hydrological angular momentum (HAM) estimates derived from hydrological modelling (Śliwińska et al., 2021; W. Chen et al., 2017). The sensitivity of the SLR technique to the time variable gravity field is however limited to spherical harmonic degrees of about $n = 5$ or 6, with certain coefficients of higher degree and order being observable due to the SLR orbital geometry while others cannot be separated (Sośnica et al., 2015).

To increase the spectral resolution of the SLR technique Löcher and Kusche (2020) and recently M. Cheng and Ries (2023) employed a parameterization based on the EOF approach in the SLR data reduction instead of spherical harmonics. In this method, the leading spatial patterns of mass variability, i.e. the EOF functions, were derived from the unfiltered spherical harmonics coefficients from GRACE (based on the ITSG-Grace2018 solution in Löcher and Kusche (2020)) in a preprocessing step.

Within a dynamic orbit improvement procedure, in the hybrid approach some low-degree spherical harmonic coefficients and a set of scaling factors fitting the GRACE EOFs were derived, fitting the original laser range observations. Spherical harmonic coefficients complete to higher degrees are finally derived via re-mapping the scaled EOFs. In this way, Löcher and Kusche (2020) derived monthly gravity fields from $1992-2020$ complete up to spherical harmonic degree $n = 60$. These solutions must be viewed as having inherited spatial constraints from GRACE in the sense that EOFs beyond some threshold in signal power, and thus combinations of spherical harmonics, were truncated and effectively set to zero. However, they include information on mass change in the pre-GRACE era of unprecedented spatial resolution (see validations in Löcher and Kusche (2020)), are publicly available, and we thus use them here to assess the reconstructions described earlier.

2.1.4 The Water Global Analysis and Prognosis (WaterGAP) model

The WaterGAP model (Müller Schmied et al., 2020a) consists of three major components, i.e. the global water use model, the linking model Groundwater-Surface Water Use (GWSWUSE) and the WaterGAP Global Hydrology Model (WGHM). The global water use model determines water use for irrigation, livestock, domestic and manufacturing use, and thermal power. GWSWUSE divides the net water abstraction determined by the global water use model into net abstraction for surface water and groundwater. Net abstraction together with climate forcing then form the input for the WGHM model. WGHM represents water flows and storage in ten compartments; i.e. canopy, snow, soil, groundwater, lakes, man-made reservoirs, wetlands and rivers.

From this we derived monthly total water storage for the years $1901-2016$ by accumulating over all storages. It is worth mentioning here that WaterGAP does not simulate glaciers.

2.2 Methods

For all comparisons in this study, total water storage trend, acceleration, annual amplitude and annual phase are estimated within a least squares approach.

The sub-seasonal signal is defined here as the signal with a period shorter than a year. After removing trend and annual signal from the time series, a high pass filter was used to suppress frequencies below 1 cycle per year. The inter-annual signal is defined as the signal with periods longer than one year. It is obtained here via low pass filtering, after the time series were reduced by a trend, annual and semiannual signal. For the inter-annual and sub-seasonal signals the power is represented via the root mean square (RMS) of the variability.

For the comparison with results from satellite laser raging in section 4, all data sets are expanded into spherical harmonics, truncated at the same degree and order as the SLR data, and expressed as mass change fields. This step is required, as the SLR fields used in this study have a coarse resolution, and we use them only until a spherical harmonic degree of $n = 12$.

We excluded Antarctica and Greenland signals from the reconstructions and WaterGAP, as we feel we do not have reliable modelling data here that could complement our approach. We also exclude those regions from the SLR data to facilitate the comparison.

3 Continental total water storage anomalies in the pre-GRACE era

The primary objective of global, data based reconstructions of total water storage is to better understand the collective response of the land system to long-variability in the rainfall and temperature while remaining independent of hydrological modelling, with applications ranging from space geodesy and sea level budgeting to providing constraints to climate modelling studies. It is important however to understand that reconstructions are by construction principle developed to reproduce or disregard certain temporal scales, and are bound to reproduce events and signatures that are represented in the used predictor variables.

We suggest therefore to distinguish internal and external consistency. In what follows, we will first present a systematic analysis of the HG19 and L21 global land water storage reconstructions with respect to each other, and when compared to the WaterGAP model, which also relies on meteorological forcing fields but in a very different way. For each data set the linear trend, acceleration, annual amplitude and phase, interannual signal, and finally the subseasonal signal are discussed separately for the pre-GRACE era (1979-2002) and for the full reconstruction period (1979-2016). In section 4 following thereafter, the focus will then be on an independent 'external' evaluation with GRACE/-FO and SLR data.

3.1 Continental total water storage anomalies in the years 1979-2016

3.1.1 Humphrey and Gudmundsson (2019) reconstruction based on GRACE data

The first column of figure 1 shows the linear trend, acceleration, average annual amplitude and phase, sub seasonal and inter-annual signal variations for HG19 for the years 1979-2016. The metrics reveal a "chessboard"pattern, that is most likely due to the model formulation of HG19. This reconstruction does not include a trend – Humphrey and Gudmundsson (2019) argue that trends seen by GRACE are mainly driven by anthropogenic effects, which cannot be explained by their statistical model. However, changes in temperature and precipitation include a trend, which can be estimated based on a least square adjustment from the reconstructions. Negative trends are visible for the Congo basin, the region around the Lake Victoria, the Mississippi basin, the Tocantins basin, the Parnaiba basin and the Sao Francisco basin. Positive trends are found for the Orinoco basin, the Zambezi basin and around the Thar desert.

Negative accelerations, i.e. increasing rates of water storage decline or 'accelerated drying' are found for the Mississippi, Colorado and Rio Grande basin in North America, along with the the large basins along South America's Atlantic coast, the Amazon basin, for the Congo basin in Africa, the Volga and Don basin, the Amur basin, large swathes in the eastern part of China, and the Murray-Darling basin in Australia. Positive accelerations, i.e. increased rates of water storage due to rainfall increases, are shown for the Zambezi basin in Africa. Increasing mass rates are also shown for Iceland, consistent with Wu and Heflin (2015), but it is unclear to what extent the HG19 method may be trusted over glacierated regions.

Large long-term average annual cycle of total water storage are found for the Amazon basin, with the dryness peak (i.e. minimum water storage, phase expressed as day-of-year DOY) around February to March. Large annual amplitudes are also found for some coastal basins in South America with dryness peaking at around DOY 250 for the northern part to DOY 120-150 for the southern part. Furthermore, large average annual amplitude occurs in the Mississippi basin, with a dryness peak in May/June, and around the Gulf of Alaska. For the latter, the phase shows values around December to January.

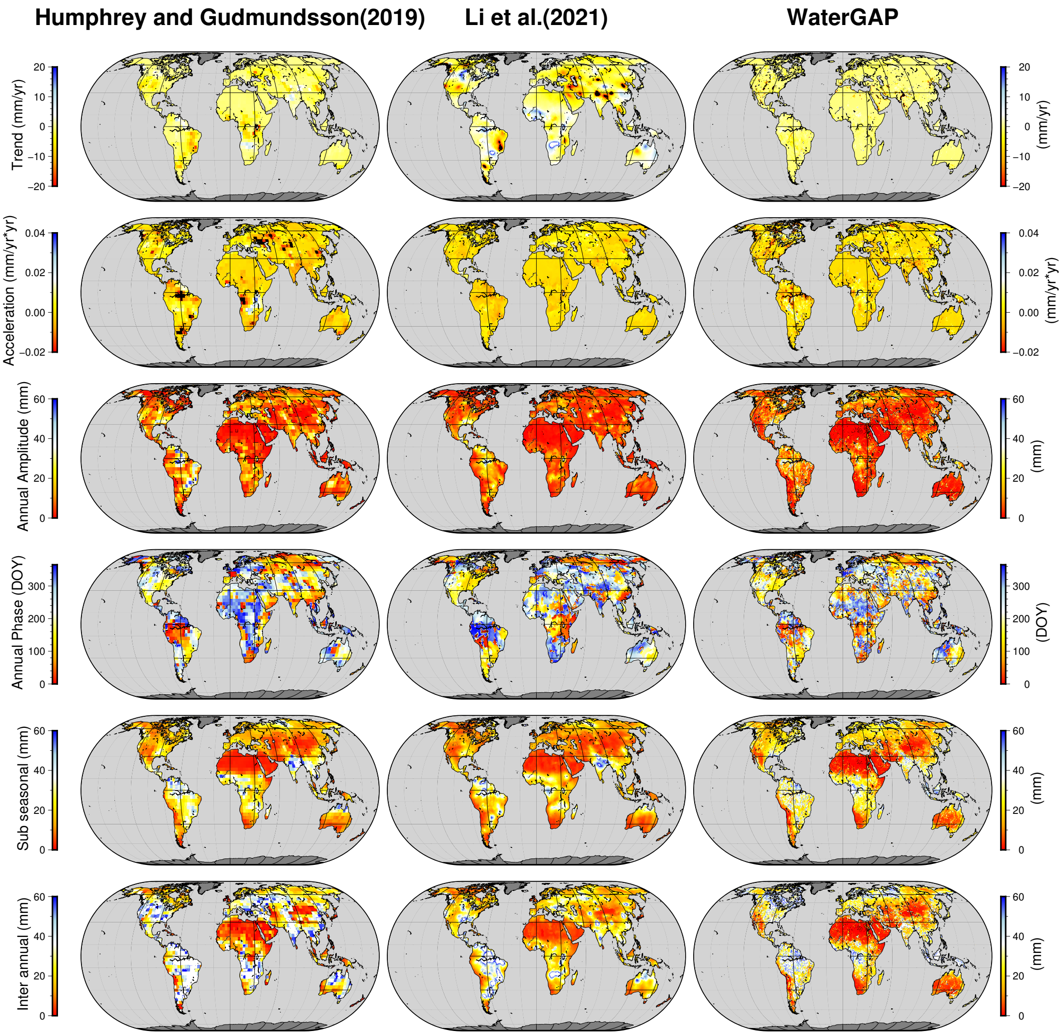

Figure 1. Trend, acceleration, annual amplitude and phase, subseasonal and interannual signal for the years 1979-2016 for the reconstructions and WaterGAP on a global scale

Regions with lower amplitudes (around 2.5 cm) are located in western and southern Europe, around the Caspian Sea, southern Africa, India, southern and eastern China and in the north-west of Australia. For Europe, in terms of peak water deficit, a gradient in magnitude between the north-eastern and south-western part is identified, ranging from DOY 50-160 for the south-western and DOY 250-360 for the north-eastern. For the region around the Caspian Sea, India, southern and eastern China and in north west Australia, minimum peak values for February to March are found, for southern Africa the values are around DOY 260-350. As expected, the annual signals disappear for deserts (Sahara, Arabian Peninsula, Mongolia, Australia) and arid regions (major parts of Canada).

Subseasonal signal power (Fig. 1, first column, fifth row) largely follows the Köppen-Geiger classification (Köppen, 1923; Kottek et al., 2006), with low variations in arid climate zones and large vaiability in tropical areas, e.g. the Amazon and La Plata basins, Niger, Congo and Lake Chad basins, India, southern China, Indonesia and tropical northern Australia. Also regions with polar/snow climate, e.g. northern Canada and Siberia and some associated with mild to warm climate with regular precipitation patterns, like the eastern part of the Mississippi basin and Japan, appear with stronger subseasonal signal variability in the long-term reconstruction.

At the inter-annual scale (bottom right panel) only northern Africa, Arabian Peninsula, Mongolia and south-eastern Australia, i.e. regions with arid climate, exhibit little variability; similar also for the northern hemisphere. Interestingly, large signals of an amplitude comparable to the annual cycle appear to be present (over nearly four decades) for the entire U.S., South America except the Andes, sub-Sahel and southern Africa, the north-eastern coast of Australia, Europe and most parts of Asia.

3.1.2 F. Li et al. (2021) reconstruction based on GRACE data

Even though the Li21 reconstruction does include a trend, this trend has not been reconstructed in the same way as the other signal components; rather it has been separated from the GRACE time series and added back to the finished reconstruction (F. Li et al., 2020, 2021). Therefore, the trends in figures 1 (first row, second column) and 2 (first row, second column) represent mean trends in the used GRACE data. Minor differences in the trend magnitudes can be attributed to the least squares approach used to generate the maps.

Trends in the observed total water storage are in fact due to a mixture of natural variability and anthropogenic effects (Humphrey et al., 2016; Rodell et al., 2018), and since not all human effects can be viewed as a response to natural variability (as e.g. increased water withdrawals in drought periods may be explained) it is unclear to what extent statistical reconstructions trained on such GRACE data sets would represent reality or artefacts. As mentioned above, visible negative trends represent extrapolations from the GRACE period aided by climate data. For the Middle East, they are generally attributed to a superposition of droughts and resulting groundwater depletion for agricultural use (Voss et al., 2013). The decrease in water storage in the Caspian Sea is mostly ascribed to reduced inflow of the Volga river (Loomis & Luthcke, 2017) in the GRACE period.

Accelerated water mass loss can be seen for the Amazon basin, around Victoria Lake and the Caspian Sea. Slightly positive accelerations are found for the Ob basin, around Alaska, the Patagonian glaciers and in Iceland, although for the latter it is not clear whether the Li21 approach is valid.

The Amazon basin, the Zambezi basin (both with a dryness peak around DOY 350) the southern part of South America, the Mississippi basin (both showing minimal phase around April/May), the Caspian Sea region with minimal peaks in the end of Spring, the Victoria lake (with minimal signal magnitudes around January/February) and some ice covered regions, i.e. Iceland, the Patagonian glaciers, Svalbard and the region around the Gulf of Alaska exhibit large amplitudes. For South America, Europe, India, north-eastern China and central Africa moderate annual amplitude of around 1.5−2 cm are shown. The corresponding annual phase, e.g. minimal signal magnitude for South America shows a notable shift between South and North, with minimal water storages around DOY 300-265 for the northern part and DOY 120-180 for the southern part. A similar phase difference can be found for Europe, with minimal signal magnitudes in early Spring / end of Winter for the south-western part and around end of Fall/ begin of Winter for the north-eastern part. North-eastern China and middle Africa both show minimal signal magnitudes for January/ February. The phase for India is around DOY 320 to 260.

Strong sub-seasonal variations are present for the equatorial belt, spanning the Orinoco, Amazon and La Plate basin in South America, the Zambezi, Congo and Niger basin in Africa, Madagascar, India, southern parts of China and the northern part of Australia. The Caspian Sea exhibits also high signal variability, as well as the region around the North China Plain and California, as well as the northern part of the Mississippi basin in North America. Also ice covered regions in Siberia, Canada, around the Gulf of Alaska, around the Patagonian glaciers, Iceland, Svalbard and the Franz-Joseph land depict high signal magnitudes. Most European regions show variability of around $1.5-2$ cm here.

At the inter-annual time scale, the equatorial band becomes even more pronounced, with strong inter-annual signals for most parts of South America, middle Africa, India, southern China and northern Australia likely caused by rainfall variability. Also ice covered regions, like Iceland, the Patagonian glaciers, the Gulf of Alaska, and China's High Mountains glaciers exhibit strong variations. Interestingly, regions where anthropogenic water use prevails, e.g. California, the North China Plain and parts of India exhibit large inter-annual signal magnitudes. The region around the Black Sea, including the Danube, Dnieper, Don, Volga and part of the Ob basin also reveal high signal variability.

3.1.3 WaterGAP Global Hydrological Model simulation

The majority of the trends derived from WaterGAP for $1979-2016$ range between $-0.5cm/yr$ and $0.5cm/yr$ (see figure 1, top left panel). Negative trends can be observed for large parts of America, Australia and Europe, whereas Africa and Asia present a mixed picture of negative and positive trends. Stronger negative trends can be found for the US, over the Arabian Peninsula and India, with patches of stronger trends located all over the world, some of which may be model artefacts. Trends for glacierated regions are not dominant in figure 1 (first row, third column) as WaterGAP does not include glaciers in the simulation (Müller Schmied et al., 2020a).

Trend rates generally tend towards negative accelerations (i.e. increased dryness), with some exceptions for the Congo, the Amazon, the Parana and Ob basin, around the southern region of the Hudson Bay and Indonesia. Notable are negative acceleration pattern over India. The same phenomenon can be observed for the Mississippi basin, where regions showing a strong negative trend also exhibit notable negative acceleration. As mentioned before, the Congo and Ob basin both show partially positive acceleration. The trends for these basins are slightly negative. The opposite sign of trend and acceleration may suggest that long-term reversal from drying or wetting is in progress. Striking is also the negative acceleration in river storage along the main Amazon system, in conjunction with a positive trend for the southern branch, i.e. Madeira and Jurua, and a negative for the northern one, Negro and Branco.

The most prominent annual signal in figure 1 is the Amazon basin. For this area the minimum signal, in terms of water deficit, has been around January to February in the four decades considered here. other regions of higher amplitude in South America can be found for the Orinoco, with minimal phase around DOY 250-300, and Parana basin (lowest storage in April/May). Large annual amplitudes are also found for the Mississippi basin, dryness peaking around DOY 120-160, around Hudson Bay and the Gulf of Alaska. Visible are also large amplitudes for the Congo and the Volga basins, both with minimal signal in February, in the Rhone basin, Norway and parts of Sweden, all showing minimal values around November/December, the Ob basin (March-May), and Indonesia (around DOY 300-350).Moderate annual amplitudes can be found for India, southern China and Central Europe (minmum around DOY 300-360).

In the model simulation, all equatorial regions show high subseasonal signal variations, this includes the majority of South America, Central Africa, India, southern China, Indonesia and the north-eastern coast of Australia. Also regions of polar climate exhibit bigger signal amplitudes. Among the regions with bigger variations is also the Mississippi basin. Europe shows a mixed picture with signal magnitudes around $1.5-2.5$ cm. River storage in WaterGAP appears pronounced in the subseasonal signals.

At inter-annual time scales, regions with an arid warm or cold climate exhibit small signal variations, like the Sahara and the Orange, Okavango and Limpopo basin, the Arabian Peninsula most parts of Australia, Central Asia, the Andes and parts of Mexico. Strong inter-annual signals can be found for the Amazon, Orinoco and La Plata basin and the Patagonian glaciers, the major part of North America, the Congon basin, the northern part of Europe and Siberia, along the Pacific coast of Asia, India, southern China and Indonesia.

3.2 Continental total water storage anomalies for the years 1979-2002

3.2.1 Humphrey and Gudmundsson (2019) reconstruction based on GRACE data

The first panel in the first row of figure 2 shows the trend for HG19 for the years $1979-2002$. HG19 exhibits positive trends for the coastal region of Alaska, the Nelson basin, parts of the Congo basin, the Limpopo basin, over India and the Amazon basin. The latter was not visible for the years 1979-2016. Negative trends are found for the Tocantins basin, the Parnaiba basin and the Sao Francisco basin, the region around the Victoria Lake and the Bandama basin.

Compared to the mean rate changes over the entire time frame (Fig. 1), the storage acceleration in two decades prior to GRACE was larger in magnitude, while also the spatial patterns were apparently different. In HG19, the increasing water storage loss over the US extend to the complete Mississippi basin during this period. Storage anomalies over Ellesmere Island are apparently characterized by positive rate changes, and similar behavior can be found for the Orinoco basin, the East of South America, Africa, Asia and even Europe. Quite notable changes in patterns and magnitude are found for Australia, with increasing wetness trends in the decades prior to GRACE while over the entire time frame (Fig. 1) (and of course over the GRACE period, (van Dijk et al., 2013)) this region experiences long-term drying trends.

The long-term average annual cycle of total water storage over two decades prior to GRACE revealed large amplitudes for the Amazon basin, with a minimum peaking around end of February and March. Large amplitudes were also present along the Pacific coast of North America with minimum storage in around DOY 250, in the Mississippi basin with a minimal peak in May-July and over northern Asia peaking towards the end of February to March. In total, compared to the long term average annual cycle for the entire time period, amplitudes seem bigger for the two decades prior to GRACE. In contrast, the sub-seasonal signal content seems not to depend overly on the time period, and we find the same for the inter-annual signal.

3.2.2 F. Li et al. (2021) reconstruction based on GRACE data

Mean rate changes for the twenty years prior to GRACE are shown in the second column and row of figure 2. Compared to the entire time frame (figure 1, second column and row), we identify a notable change in pattern for Australia – along the southern coasts negative acceleration and in tropical northern regions a positive one. Negative accelerations are also found for the Congo basin, the Parana basin, the Mississippi basin, the Amur basin. The Limpopo basin, the Zambezi basin, the north-western part of Australia,

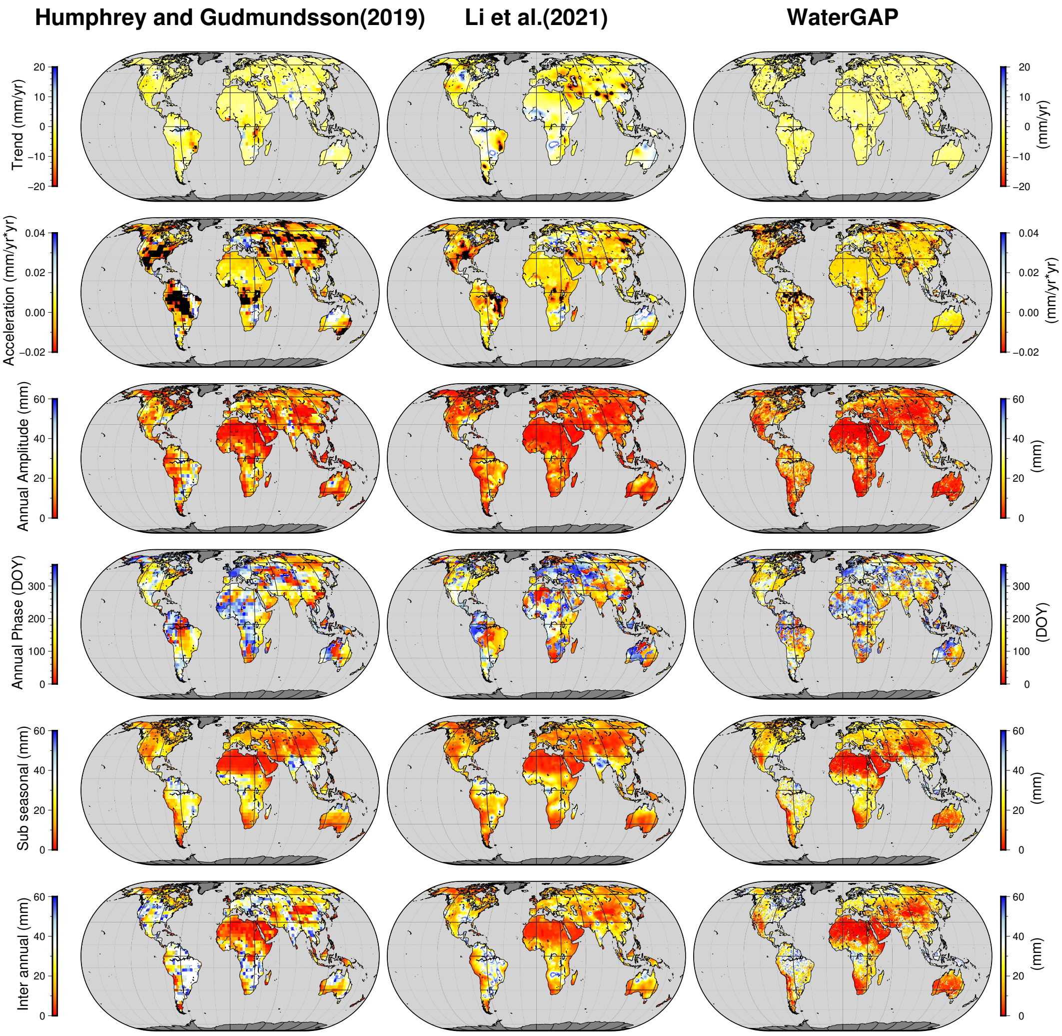

Figure 2. Trend, acceleration, annual amplitude and phase, subseasonal and interannual signal for the years 1979-2002 for the reconstructions and WaterGAP on a global scale

the Danube and Volga basin show positive range rate changes. All mentioned signals are not present in the accelerations derived for the full time frame.

The mean annual amplitude derived for the two decades prior to GRACE shows large annual amplitudes for all coastal basins in South America, with minimal values around January/February and October/November for the upper part of South America and minimal phases in Spring for the lower part. Furthermore large mean annual amplitudes are shown for the southern basins in Africa and the western basins in Siberia. Similar to South America the minimal peak in Africa is in November for the lower part begin of the year in the upper part. For the Northern Asian basins the minimal peak is in early spring, e.g around April/May. In comparison to the mean annual amplitude for the entire time frame the derived amplitude for the twenty years before GRACE has increased in magnitude.

The signal magnitude of the interannual and subseasonal signal is slightly smaller for the two decades prior to GRACE compared to the full time frame. The signal pattern, e.g. regions revealed by the data to experience signal variations, is the same for both time frames.

3.2.3 WaterGAP Global Hydrological Model simulation

Like the trend the mean range rate changes for the two decades prior to GRACE exhibit larger magnitudes compared to the full time frame. Regions with positive mean range rate changes are Indonesia, parts of the Ob basin, Central Europe, parts of the Congo, Zambezi basin in Africa, the Atlantic coast side of South America. The main river branch of the Amazon basin, Alaska, the Mackenzie basin, the Mississippi basin, the Neva basin and northern India reveal negative accelerations. For northern India a negative trend is found, indicating a water depletion in this region.

The average annual signal exhibits large amplitudes in the Amazon basin with minimum water storage ranging from October to March. Other regions with large average annual amplitudes are Indonesia, the region around the Golf of Alaska and Scandinavia. All mentioned regions show minimum storage around September-November. Notable are also the large annual amplitudes for the basins located around the Golf from Mexico (October-March).

3.3 Intercomparisons

In this section, we first briefly discuss general differences between the three data sets. Then, we focus on the Murray-Darling and Amazon basins, regions that are strongly influenced by the El-Niño-Southern Oscillation phenomenon (Nicholls et al., 1997; Trenberth, 1990; Forootan et al., 2016; García-García et al., 2011; Towner et al., 2021; J. L. Chen et al., 2010; Marengo & Espinoza, 2016). We expect the reconstructions to exhibit large storage signals, as both use precipitation in the model formulation. Then we will turn to Europe, where water storage signals (e.g. recent droughts) have large economic impact but are dwarfed by other regions. We suspect Europe to pose more of a challenge to the reconstructions.

Table 2 shows trend (without and with co-estimation of an acceleration term). We find that the trend from Li21 is by more than two times bigger compared to HG19 and WaterGAP. It seams like, WaterGAP (and HG19) underestimate trends in water storage changes compared to Li21. However, as discussed in section 4 the derived trends depend on the used GRACE solution. The trend derived from mascon based GRACE solution, like the one used by Li21 or in the study of Scanlon et al. (2018) exhibit higher water storage signal magnitudes compared to trends derived from spherical harmonic solution.

We note, that the accelerations found for the reconstruction of HG19 are stronger, then those for Li21 or WaterGAP. The acceleration derived from the data sets exhibits higher signal magnitude for the shorter time period. For the full time frame the derived average range rate changes are close to zero.

The data sets reveal similar regions with high average annual amplitudes, like the Amazon basin, the Mississippi basin or the Caspian Sea. Table Appendix C.1 presents values raging from 2.2 cm for HG19 to 2.4 cm for Li21 for the global land. The annual phase indicates similar values in the data sets, except for the Amazon basin, where a time lag of a month is found between HG19 and WaterGAP on the one side and Li21 on the other side.

The subseasonal signal of the two reconstructions exhibits similar signal magnitude and signal pattern and does not change notably for the pre GRACE era compared to the full time series. In comparison, the subseasonal signal of the hydrological model shows higher signal magnitudes, notable especially for North and South America and Europe.

On interannual scale all data sets show small signal variations for arid regions. The signal pattern of the data sets differ for North America, Europe, Asia and Australia. For Asia, Li21 and WaterGAP are in a good agreement, whereas HG19 portrays more region with high signal amplitudes. For Europe, Li21 exhibits smaller signal magnitudes compared to WaterGAP and HG19. For North America, Li21 and HG19 show moderate signal magnitude for the region around the Hudson Bay, whereas WaterGAP indicates signal variations, that are two times larger. The data sets show high signal magnitudes for the Amazon basin and the Orinoco basin. As for the accelerations, HG19 depicts the highest signal magnitudes compared to Li21 and WaterGAP. The signals for the full time period have slightly higher magnitudes compared to the two decades prior to GRACE. The signal pattern does not change.

3.3.1 Australia/ Murray-Darling basin

Australia's climate and rainfall is influenced by its geographic location, with the southwestern Pacific Ocean in the east and the Indian Ocean in the west (Trenberth, 1990; Nicholls et al., 1997). On interannual time scales the most dominant drivers for precipitation are the El-Niño-Southern Oscillation, ENSO, and the Indian Ocean Dipole, IOD (García-García et al., 2011; Risbey et al., 2009). Negative ENSO phases (El Niño) lead to reduced rainfall in the northern and eastern parts of Australia resulting in droughts especially in the center of Australia, while La Niña phases lead to strong precipitations and often flooding. Positive IOD events are linked to a decrease of precipitation in western Australia, where as negative events lead to an increase in rainfall (Trenberth, 1990; Nicholls et al., 1997) and `http://www.bom.gov.au/climate/`. At the beginning of the new century Australia experienced the severe "Millenium drought" (van Dijk et al., 2013). The strong La Niña event starting 2010 led to increased precipitations and caused floods all over the continent (`http://www.bom.gov.au/climate/enso/lnlist`).

Trends from WaterGAP have opposite sign and smaller magnitude as compared to Li21. The small trends for WaterGAP are also found by Schumacher et al. (2018) and Yang et al. (2020). The latter reported a mismatch between GRACE and WaterGAP supposedly due to a lack of ability of the model to represent ground water trends and soil moisture. Trends computed based on the reconstruction by Li21 are mostly positive indicating a shift from dry to wet. The trends from WaterGAP are mostly negative indicating a slightly increase in water mass loss.

The data sets show higher magnitudes in the average rate changes, when comparing the entire time frame with the twenty years before GRACE. For the entire time frame, the reconstructions by Li21 and WaterGAP both exhibit near-zero small acceleration, whereas HG19 suggests much higher rate range changes. The regions in the Murray-Darling basin exhibiting negative accelerations are associated with agricultural areas (van Dijk et al., 2013). The negative range rate changes might be related to a decline in the groundwater storage from 1993-2009 primary due to pumping for agricultural and domestic purpose (J. L. Chen et al., 2016). HG19 depicts positive accelerations for these regions, which might be due to the model formulation, that does not include any anthropogenic effects.

van Dijk et al. (2013) reported a positive trend in precipitation for the north and west of Australia, both regions show a positive acceleration for Li21, HG19 and WaterGAP in the twenty years prior to GRACE, reflecting the increase in water mass.

For the annual signal the data sets are in a good agreement except for one region, at the northwestern part of Australia. WaterGAP does not show any signal variation in this region. HG19 shows higher annual amplitude magnitudes compared to the other data sets for the entire time frame, whereas Li21 exhibits high magnitudes for the twenty years prior to GRACE. The minimal signal peaks for Australia range from October for the south-western part to February/ March for the southern part.

For the subseasonal signal the reconstruction displays low signal variability in the south and higher ones in the north of Australia. WaterGAP additionally depicts high signal variations for the eastern coast and the region around the Swan coastal area. Both regions are associated with a humid climate (Köppen, 1923) and are mostly affected by ENSO and IOD, both phenomena leading to changes in precipitation patterns and strength (Forootan et al., 2016).

On interannual signal scale WaterGAP exhibits similar signal patterns as for the subseasonal signal, but with higher signal magnitudes. The reconstructions show extended signal patterns covering the north and the eastern and southern coast. The inter-annual signal variation derived from HG19 show higher magnitudes compared to Li21 and WaterGAP, especially for the eastern coast. This region is strongly affected by ENSO events (Forootan et al., 2016; van Dijk et al., 2013). The change in precipitation creates signal variabilities visible in the data sets.

On interannual signal scale WaterGAP exhibits similar signal patterns as for the subseasonal signal, but with higher signal magnitudes. The reconstructions show extended signal patterns covering the north and the eastern and southern coast. The inter-annual signal variation derived from HG19 show higher magnitudes compared to Li21 and WaterGAP, especially for the eastern coast. This region is strongly affected by ENSO events (Forootan et al., 2016; van Dijk et al., 2013). The change in precipitation creates signal variabilities visible in the data sets.

3.3.2 Amazon basin

The Amazon basin is Earth biggest drainage basin, with high seasonal rainfall variability (Costa & Foley, 1998; Marengo, Liebmann, et al., 2012). Frappart et al. (2013) found a high correlation between rainfall and GRACE derived TWS for the years 2003−2010, making it quite reasonable to assume, that rainfall is the main driver for water storage changes in the Amazon river basin. On inter-annual time scale the precipitation patterns are influenced by ENSO, leading to several droughts and floods (Marengo, Tomasella, et al., 2012; Towner et al., 2021; Marengo & Espinoza, 2016; Xavier et al., 2010; Phillips et al., 2012; J. L. Chen et al., 2010).

The trend from Li21 for the Amazon basin is mostly positive, indicating a water mass increase for this region. WaterGAP on the other side depicts a positive trend for the Negro and Madeira sub-basin and a negative one for the Solimoes sub-basin.For their reconstruction Becker et al. (2011) performed an empirical orthogonal function (EOF) analysis of precipitation data over the Amazon for the years 1980−2008, 1990−2008, 2003−2008. The leading EOFs show a strong positive signal (in terms of water accumulation) for the western part and a negative one for the eastern part. The derived dominant signals by Becker et al. (2011) correspond nicely to the trend pictured by WaterGAP, leading to the conclusion, that the derived trend is mainly driven by precipitation. Floods and droughts in the Amazon basin are known to be driven by large-scale climate variability, like ENSO (Towner et al., 2021). According to Rodell et al. (2018) the mostly positive trend derived from Li21 is due to an increase in precipitation and floods mainly driven by La Niña after an decrease in precipitation related to El Niño (J. L. Chen et al., 2010; Marengo & Espinoza, 2016; Towner et al., 2021).

For the entire time period the average rate range changes are mostly negative, except for the Madeira sub-basin. The trend derived for Li21 is mostly positive, the acceleration mostly negative. This suggests, that the Amazon basin may be in a near equilibrium state. We find the same in the WaterGAP simulations, even though the trends magnitude of WaterGAP is smaller compared to Li21. For the twenty years prior to GRACE the range rate change pattern of WaterGAP does not change. A change from positive to negative acceleration for the Negro sub-basin for Li21 is visible. The accelerations for HG19 show an opposite pattern with respect to the accelerations computed from WaterGAP. This mismatch means that at interannual timescales the water storage change in the Amazon cannot be well explained by temperature and precipitation trends; this may be due to processes not well represented in either reconstruction or model, or due to biases in the forcing data.

All three data sets show high annual amplitudes over the main stream. Due to the resolution of the underlying GRACE solution the derived signal patterns for the reconstructions are blurred, where as in WaterGAP the river stream is visible. Comparing the annual amplitude for the entire time period and the twenty years prior to GRACE reveals no change in signal magnitude. However, the signal pattern changes slightly. For Li21 it becomes narrower, mimicking the WaterGAP river routing. For HG19 the annual signal pattern shifts slightly towards the North Atlantic. The annual phase of the datasets displays minimal signal peaks for December to January.

Both reconstructions show high subseasonal signals for the Negro and Amazon subbasins. For the main branch the magnitude of the subseasonal signal of Li21 is smaller compared to HG19. WaterGAP displays high subseasonal signal variations for the whole Amazon basin. The signal magnitude increases from Li21 to HG19 to WaterGAP.

High inter-annual signal variability over most parts of the Amazon basin are visible in all three data sets on both time scales. HG19 and WaterGAP both show stronger inter-annual signals compared to Li21, with WaterGAP showing the biggest inter-annual signal variability, followed by HG19 and Li21.

To better understand the information content and 'effective' spatial resolution of the detrended and deseasonalized reconstructions, we decided to compute the spatial autocorrelation for a given point in the Silimoes sub basin of the Amazon. As expected, highest correlations are found for the Silimoes and Madeira (sub) basin. HG19 also exhibits correlations of around 0.5 for the Congo, Niger, Nil, Zambezi, Amur, Kem, Mississippi, Saint-Laurent, Nelson and Murray basin, as well as Lake Victoria. Simular correlation patterns for the Congo, Amur, Mississippi and Kem basin are also revealed in the Li21 data, however, the correlation itself is smaller compared to HG19. Moderate long-range correlations are indeed visible for both data sets; this can be explained by the irregularities (with respect to the annual cycle) in the precipitation driven by ENSO.

3.3.3 Europe

Europe is facing more and more severe droughts, leading to an decrease in water storage (Gerdener et al., 2020; Boergens et al., 2020; Gudmundsson & Seneviratne, 2015; Spinoni et al., 2015a, 2015b). The largest changes are found for the big European basins: Volga, Danube, Dnieper, Don and Rhine (Rodell et al., 2018; Humphrey et al., 2016).

Li21 shows positive trends for the Mediterranean, Boreal and Atlantic parts of Europe for the full time frame. For Central Europe the reconstructions indicate mass losses increasing from west to eastern Europe, the biggest negative trends allocated around the Black Sea. This findings are in a good agreement with Rodell et al. (2018); Eicker et al. (2016); Tapley et al. (2019). WaterGAP exhibits positive trends for the Boreal regions and western part of the Danube basin and negative ones elsewhere. For the two decades prior to GRACE, WaterGAP displays a shift from negative to positive trends. This is in a good agreement with the findings of Spinoni et al. (2015b) Similar behaviours are found for Li21. The change in trend patterns between the two time series suggest a decrease in water storage after 2002, so strong, that it influences the derived trends and, therefore, dominates the time series. Following WaterGAP, the water storage decrease affects all European countries, where as in Li21 the affected countries are located in the middle of Europe.

For the full time series, WaterGAP and Li21 exhibit similar slightly positive acceleration patterns. HG19 displays high signal magnitudes compared to the other data sets showing mostly negative accelerations, with maximum values around the Black See and parts of France. Positive accelerations are shown for the region around the Adriatic Sea. For the years 1979-2002 the acceleration patterns become more pronounced in all three data sets. HG19 exhibits mostly positive acceleration except for Siberia, parts of Norway and parts of Spain. In contrast to that, Li21 depicts negative acceleration for all countries located between the Adriatic and Black Sea and positive ones over Siberia. These patterns are reflected by WaterGAP. However, WaterGAP nevertheless displays negative accelerations for Sweden, whereas both reconstructions show positive ones.

The annual amplitude of HG19 does not change through the time frames. Regions with higher annual amplitudes are located around the Black Sea, the southern part of Spain and Portugal and the north-eastern part of France and Germany and Iceland. Low annual amplitudes are shown for the Boreas and Atlantic countries. Similar patterns are found for WaterGAP and Li21 for the years 1979-2002. For the full time frame the signal magnitude is smaller compared to the twenty years prior to GRACE, while the signal pattern stays the same. For the full time frame the data sets exhibit values around DOY 80-150 for south-eastern part, values around DOY 200-300 for the eastern part and values above DOY 300 for Central Europe. For the Pre-GRACE time frame, all data sets show, that the minimal day of the annual signal amplitude only changes for the eastern and middle part of Europe, so regions, that are more effected by water mass changes. For the shorter time period the data sets shows a good agreement for central and southern Europe, Scandinavia and the region around the Adriatic Sea. For the region around the Back Sea, Li21 and WaterGAP display values around 150 DOY. The values derived from HG19 are around 230 DOY and closer to the full time frame.

The derived subseasonal signal from the data sets is the same for both time periods. Li21 and HG19 identify similar regions in Central Europe and the coastal border of Norway with high subseasonal signal variations. However, the magnitude of the signal shown by Li21 is smaller compared to HG19. Low subseasonal signal variations are shown for England. WaterGAP exhibit higher signal magnitudes compared to the reconstructions, with high variations for the northern and southern part of Europe, with a small band of lower magnitudes covering the continual part of Europe.

The reconstructions show high interannual signal variations for the eastern part of Europe and southern part of Spain. Especially for eastern Europe the derived magnitude is higher for the full time series compared to the years 1979-2002. For Central Europe Li21 displays moderate interannual signal variations. The derived signal variations from HG19 and WaterGAP are notably higher, with WaterGAP displaying signal variations up to two times bigger compared to Li21.

We also compute the autocorrelation and time lag for the detrended and deseasonalized reconstructions for a point in Central Europe (figure 10). As expected for both reconstructions the correlation is high around the chosen point, and drops to near zero after a few thousand kilometres. High correlation are thus (in this example) shown for the Loire, Rhone, Seine, Garonne Elbe, Rhine and Po basin. We find that for Li21, the correlation pattern is stronger compared to HG19 and extends more to the west. The time lag for both data sets is in the range of a month. As expected, long-range correlations are nearly zero.

4 Evaluation of global TWSA reconstructions with satellite tracking data

4.1 GRACE and GRACE-FO data

We derived a trend, acceleration, annual amplitude and annual phase using a least square approach for four different GRACE solutions and the reconstructions for the years 2002−2020 for the global land. The computed values are shown in table Appendix C.3. The first column of table Appendix C.3 displays the trend, if only trend and annual amplitude and phase are estimated. The second column shows the trend in case an additional parameter for the acceleration is taken into account during the estimation process. As the annual amplitude and annual phase are derived based on orthogonal functions, the annual amplitudes and phase are less sensitive towards additional parameters. The according time series are shown in figure Appendix C.1. Clearly visible are the different slopes depicted by the time series and the slight differences in amplitude between the different data products.

The trends in the first column of table Appendix C.3 all show slightly negative values, indicating a water loss over land. For the spherical harmonic solution the computed trend varies from −0.15 mm/yr to −0.22 mm/yr. The trend from the mass concentration block solution (mascon) from Center of Space Research (CSR) is around ten times bigger compared to the spherical harmonic (SH) solution. Similar results were found by Jing et al. (2019) for the Tibetan Plateau. The reconstruction from Li21 is based on the CSR mascon solution, the magnitude of the trend of this specific solution is reflected by the reconstruction. The negative trends of the mascon based solutions are also clearly visible in figure Appendix C.1. As mentioned before, the HG19 reconstruction is not designed to reproduce a trend, e.g. as contained in the training TWSA data.However, it is nevertheless possible to derive a trend for any given time period and other reseachers have used such trends. The derived trend of −0.16 mm/yr is in the order of the spherical harmonic solutions.

Adding a parameter for the acceleration to the estimation process gives the model more flexibility in terms of model estimation, leading to changes in the estimated magnitude and sign of the trend. For the spherical harmonic solutions the values range from 0.59 mm/yr to 0.69 mm/yr. The negative trend from the mascon solution and Li21 decreases from −1.76 mm/yr to −1.13 mm/yr. Again, the trend for HG19 is within the range of the estimated trends of the spherical harmonic solutions.

The estimated acceleration for all datasets is negative, ranging from −0.04 mm/yr^2 for the spherical harmonic solutions and HG19 to −0.03 mm/yr^2 for the CSR mascon solution. The CSR mascon solution and Li21 are quite close to each other in terms of derived trend. However, the acceleration derived for Li21 is two times bigger compared to the one derived from the CSR mascon.

The annual amplitude varies from 13 mm for the German Research Center for Geoscience (GFZ) spherical harmonic solution to 24 mm for the CSR mascon solution and Li21. The two reconstructions together with the mascon solution depict the highest amplitude with values over 20 mm.

We conclude, that the reconstructions seem to follow the signal properties of the GRACE solution used to train them. This can be seen for the trend of Li21, which is close to the one of the CSR mascon solution, and for the annual amplitude, which is inherited for both data sets from a mascon based GRACE solution. We find, that the annual amplitudes of the reconstructions are closer to the mascon based compared to the spherical harmonic GRACE solution.

4.2 Pre-GRACE era comparison to satellite laser ranging data

For this section the reconstructions have been expanded into spherical harmonics and have been truncated at a spherical harmonic degree of $n = 12$, omitting most of the higher frequency signals. For a spherical harmonic degree of $n = 12$ no filtering is necessary for the SLR data and, therefore, for the reconstructions.

The strongest negative trends in Li21 are related to the mass loss of the Caspian Sea (Loomis & Luthcke, 2017; Rodell et al., 2018) and the groundwater depletion in India, Iraq, Iran and parts of Arabian Peninsula (Joodaki et al., 2014; J. Chen et al., 2014; Rodell et al., 2018). Furthermore, negative trends are found for Alaska, the Mackenzie basin, the Mississippi basin, parts of Mexico, the Sahara, the Congo basin, the Tocantins, the Parnaiba, the Sao Francisco basin, northern part of South America and western Europe. Positive trends are shown for Australia, the Limpopo, Orange, Okavango and Zambezi, the Orinoco basin, parts of the Amazon basin and parts of Canada. Like Li21, HG19 exhibits negative trends for the Mississippi basin and the region around the Caspian sea, the latter with smaller signal magnitude compared to Li21. Notable negative trends are also found for the Parana basin and the Congo basin. Regions with positive trends are Australia, except for the Murray-Darling basin, the Orange basin, Okavango basin, Limpopo basin, Zambezi basin and part of the Amazon basin. For SLR, the strongest negative trends are located over the Ellesmere and the Baffin Islands, the region of the Caspian, the Aral and the Black Sea. The trend derived for Africa is mostly positive, with the highest values depict for the Congo and Zambezi basin. A negative trend is visible for the west-southern part of Africa. Australia exhibits mostly positive trends, with negative trends at the southern western coast. In comparison, HG19 is missing signals around the Hudson Bay, Alaska and Fennoscandia, for which SLR displays the strongest trends. In Li21 these signals are only visible for Ellesmere and Baffin Islands. The data sets display negative trends for the US, the region around the Caspian, the Black and the Aral Sea, India and Siberia.

The accelerations are shown in the second row of figure 3, left to right: HG19, Li21, SLR. Strong acceleration patterns for SLR are located around the Hudson Bay, the Orinoco basin and spreading over the Black Sea, the Caspian Sea, India and China. Positive accelerations are derived for Alaska, the region under the Orinoco basin, the Congo and Zambezi basin, the Ob basin and northern Australia. Li21 exhibits negative range rate changes for the Parana basin, the Tocantins basin, the Parnaiba basin, the Sao Francisco basin, the Sierra Neveda, the Zambezi basin, the Black Sea, the Caspian Sea, northern Siberia, India and northern Australia. Positive accelerations are situated over the Ob basin, the Nil basin, the Japura and Solimoes basin (sub-basins of the Amazon basin), the Yukon basin, the southern land of the Hudson Bay and the southwestern part of Australia. HG19 exhibits positive accelerations for the Amazon basin and the northern part of South America, the Sahara and Arabian Peninsula, the Zambezi, Orange, Okagavo and Limpopo basin, Australia with the exception of the Murray-Darling basin, Alaska and Siberia. For the other regions negative range rate changes are shown. The magnitude as well as location of the accelerations differ widely across the data sets.

The annual amplitudes and corresponding phases are shown in the third and fourth row of figure 3. SLR depicts high values for the Amazon basin, the region around the Hudson Bay and the Gulf of Alaska, the Congo and Zambezi basin, the Niger and Lake Chad basin, the Ob basin, western India and southern China and north eastern Australia. For the Amazon basin, the Congo basin and north-eastern Australia the derived minimal signal magnitude is around January. The Mississippi basin and the region around the Hudson Bay display minimal phase values around DOY 120-150. For the Zambezi basin and ice covered regions the minimum is found for January. The Niger and Lake Chad basin, western India and southern China show minimal phase values around DOY 200-300. For the Ob basin it is around DOY 50-100. The average annual amplitudes in the Amazon river, around the Gulf of Alaska and in the Ob basin are also reflected by the reconstructions, with HG19 showing higher values compared to Li21. Li21 also shows a signal over the La Plata basin, the Mississippi basin and for the region between the Black and the Caspian Sea. The strong annual signal in the Congo and Zambezi basin detected by SLR is missing. HG19 pictures high annual signal amplitudes over Alaska, California, in the Mississippi basin, in the Amazon and La Plata basin, in the Zambezi

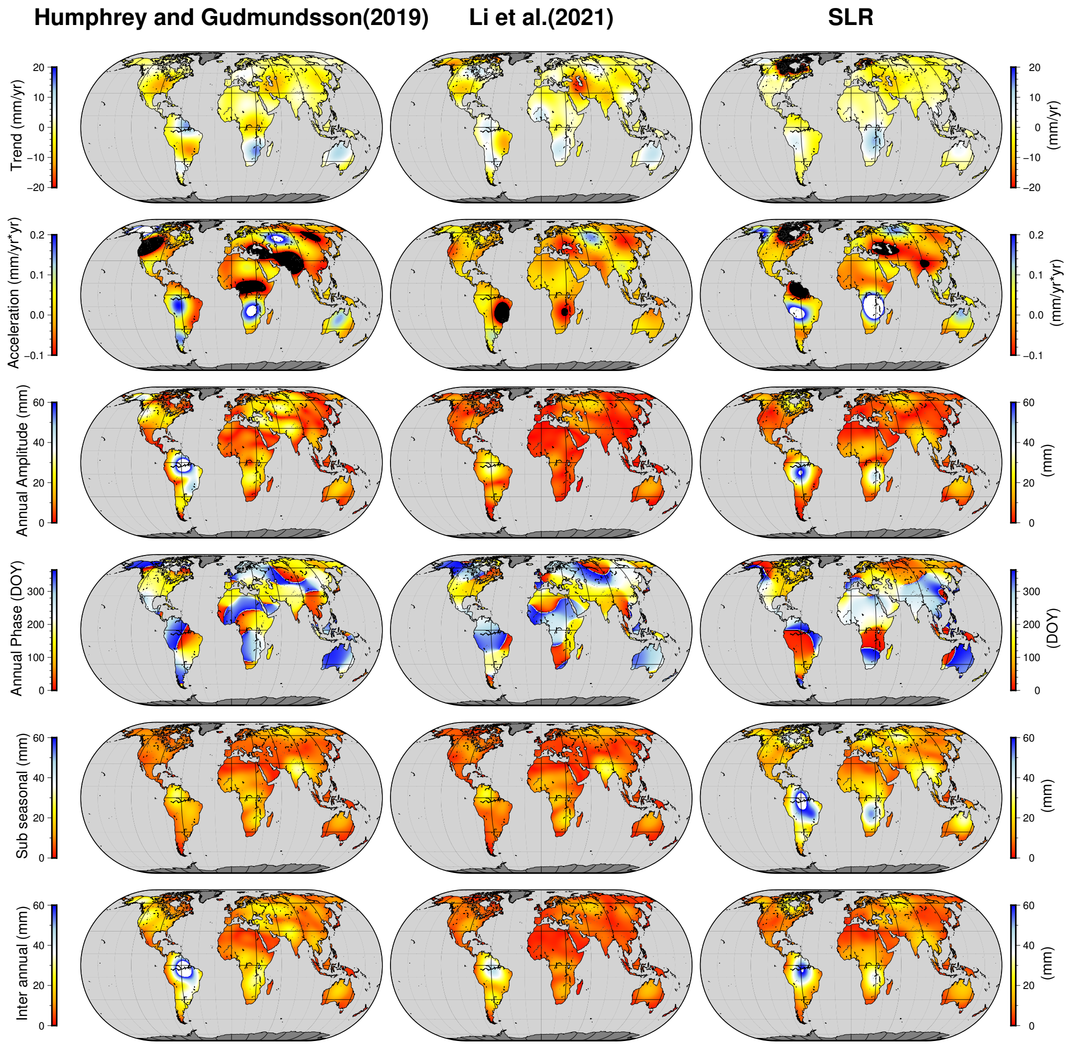

Figure 3. Trend, acceleration, annual amplitude and phase, subseasonal and interannual signal for the years 1992-2002 for the reconstructions and SLR on a global scale

basin, over the Victoria Sea, around the Black and the Caspian Sea, the Ob basin, north-western India and north-eastern Australia. The signal in the Mississippi basin, Australia and the signal spreading from the Black over the Caspian Sea to India can also be found for Li21 and partially for the SLR data, even though the signal magnitude is smaller. The annual amplitudes related to the La Plata basin and around the Victoria Sea are not reflected by Li21 or SLR. The signal in the Hudson Bay detected by SLR is not found in the reconstructions. For the Amazon basin and Alaska Li21 and HG19 depict minimal signal magnitudes around DOY 360. SLR exhibits values around zero, suggesting a small frequency shift between SLR and the reconstructions. A similar behaviour is found for Siberia, China and India.

The subseasonal signal is displayed in the fifth row of figure 3. Similar to the annual amplitude SLR exhibits high signal variations for the area around the Hudson Bay, Alaska, the Amazon basin and the Congo and Zambezi basin. Regions with a moderate subseasonal signal magnitudes are India and southern China, northern Australia, the region above the Black Sea and Alaska. Notable are also signals for the Mackenzie and Mississippi basin and the Caspian Sea. Both reconstructions show high subseasonal signal variation for the Orinoco basin, the Zambezi basin and over north-western India. Regions with lower signal variations are the northern part of Siberia, northern Australia, the Negro and Japura basin (both part of the Amazon basin), the Magdalena basin and in the Niger basin. HG19 exhibits higher signal magnitudes compared to Li21. The regions identified by the reconstructions are only partly found in the SLR data. The signal over South America in the SLR data is more pronounced and extends over the majority of the continent. The same holds for the signal over Africa and Australia. The high subseasonal signal variations over the Hudson Bay, and north-eastern Europe are not visible in the reconstructions.

The inter-annual signal variation is shown in the last row of figure 3. Similar regions compared to the annual amplitudes are found to have high inter-annual signal changes. In comparison to SLR and Li21, HG19 exhibits higher magnitudes for Siberia, Alaska and the region around the Caspian Sea. SLR shows bigger magnitudes for the Amazon basin, the Zambezi basin and the region around the Hudson Bay.

Except for region affected by GIA, the data sets exhibit similar signal patterns in all metrics. However, the magnitude of the signal differs. Strong water storage changes are, especially but not exclusively, found in the Amazon basin, the Congo, Zambezi, Limpopo, Orange and Okavango basin. As expected, the SLR data reveals larger trends and accelerations over regions affected by GIA, which are not present in the reconstructions.

4.3 Long-term evolution of water storage at river basin scale

We derived water storage changes on an interannual scale for the years 1979−2020 for the reconstructions for nine major river basins. The selected basin are the Amazon basin in South America, the Mississippi and Mackenzie basin in North America, the Nile, Niger and Congo basin in Africa and the Danube and Volga basin in Europe. The basins are visualised in figure Appendix D.2.

Several flood and drought events in the Amazon basin fall into the observation period (Marengo & Espinoza, 2016). The droughts in the years 1983 and 1997−1998 were linked to strong El Niño events (Marengo & Espinoza, 2016) and are clearly visible in the time series. Additionally, the droughts in the years 1980, 1995, 2005, the strong drought in 2015 and the floods from 1989, 1999 and 2005 are reflected by the data sets. Except for the strong drought in 2015, for which both data sets show similar signal magnitude, HG19 reveals stronger changes in the water storage compared to Li21.

The Congo basin is situated in about the same geographical latitude as the Amazon basin (Nicholson, 2022; Amarasekera et al., 1997) and is comparable to the Amazon basin in terms of climate and ecology (Nicholson, 2022). However, precipitation is lower in the Congo basin compared to the Amazon basin and, therefore, seasonal and inter-annual variability is smaller. Over the whole time series, HG19 shows a decrease in water storage for the basin, a negative trend is clearly visible. The time series from Li21 oscillates around zero. In comparison to HG19, Li21 seams to overestimate the water storage changes for the GRACE era and underestimates them for the pre GRACE era. Interestingly, the reconstructions show a good agreement for the years 1998−2002.

For the Danube basin both reconstructions reveal a loss of water mass from the year 1983 until 1990. After the year 1992 both reconstructions exhibit an increase of water storage followed by several peaks in the years 2006, 2011 and 2015.

For the Ganges basin the inter-annual time series for Li21 do not show any significant change in the water storage. HG19 on the other side, shows fluctuations of around 60 Gt over the whole time series. The precipitation over the Ganges and India depends highly on the Indian summer monsoon (MIRZA et al., 1998; Kumar et al., 2010). The moonson cycles are linked to ENSO and the IOD modes. The strong flood in 1998 partly due to the strong La Niña event is clearly visible in the reconstruction of HG19, as well as the decrease in water mass during the drought in 2002. Both events are only slightly to not visible in Li21.

The inter-annual time series for the Mackenzie river basin are in a partly agreement, for the years 1986−1989 or 2004−2015. For the years 1979−1984 and 2011−2015 the decrease in water storage shown by HG19 is more pronounced compared to Li21. The opposite is the case for the years 2002−2005 and 2016 onward. The time series for HG19 seam to be dominated by oscillating signal for the years 1992−199 with a wavelength of 3 years. This signal pattern is not reflected by the Li21 reconstruction.

The time series of the interannual signal of the Mississippi basin show a strong decrease in water mass for the years 1981, 1989 and 2013 and an increase in water storage for the years 1983 − 1987, 1994 and 2010. In comparison, HG19 exhibits stronger signal changes compared to Li21.

For the Nile basin a similar effect as for the Congo basin can be observed. HG19 shows a negative slope over the whole period, where as Li21 depicts no visible change in water storage.

During the GRACE era the reconstruction shows a strong decrease in water mass for the Niger basin for 2011. Before 2003 the reconstructions vary greatly, exhibiting partially inverse storage changes, like for the beginning of the time series. Compared to Li21, HG19 displays stronger signal oscillations for the years 1979 − 2003 and smaller ones for the GRACE era.

The reconstructions display a slight decrease in water storage for the Volga basin for the years 2010 to 2015 with a shift towards an increase for the years after 2015. The magnitude of the water mass loss for HG19 is higher compared to Li21. For the years 2000−2008 and the first years of the reconstructions, the data sets exhibit similar water mass changes. Notable, especially for HG19, but also for Li21 is an increase in water content for 1991.

We conclude, that the time series of the reconstructions are surprisingly close beyond the annual cycle for basins with signal amplitudes, dominated by seasonal precipitation, like the Amazon, Danube and Mississippi basin. For basins with low inter-annual signal variability, like the Congo and Nile basin we find differing long-term trends.

5 Conclusions

In this study we derived trend, acceleration, annual amplitude, annual phase, interannual signal parts and subseasonal TWSA signals from the two global reconstructions from Humphrey and Gudmundsson (2019) and F. Li et al. (2021) and from the hydrological model WaterGAP for the years 1979 − 2016 (full time frame) and 1979 − 2002 (pre GRACE time frame). Furthermore, we compared the reconstructions to the low degree gravity fields derived from SLR (Löcher & Kusche, 2020) for the pre GRACE time frame 1992 − 2002 for a spherical harmonic degree of expansion of $n = 12$. We also derived trend, acceleration, annual amplitude and phase from four different GRACE solutions and the reconstructions for the years 2002 − 2020.

We find similar sign and magnitude of water storage changes over the full time period and pre GRACE time period for South America and the large basins in southern Africa, like the Congo and Zambezi basin. These regions also stick out with highest TWSA in the SLR-derived maps of mass variability. For Europe, Siberia, India, southern China, Australia and North America the magnitude of TWSA over the four decades differs between the reconstructions and the hydrological model. The SLR observations exhibit higher magnitudes of water storage changes for India, northern China, northern Australia, Alaska and the north-eastern part of Siberia. However we caution that the SLR data also include gravity changes due to glacier mass variability, and a direct comparison is most meaningful in areas free of glaciers. Except for the acceleration (and trend), the signal pattern and magnitude of TWSA is similar for the full and pre GRACE time frame. The acceleration for the pre GRACE time period exhibit notably higher signal magnitudes and patterns compared to the full time frame. We recap here, that HG19 is not tuned to reproduce a trend and the one by Li21 is not reconstructed, but recovered from the GRACE period.

Given the good agreement of all data sets, we suggest, that the Congo basin has indeed suffered a prolonged loss of water storage over the last 40 years. The US show a similar picture, with negative accelerations and trends, also suggesting a water mass loss over the last fourty years. For the Amazon basin high TWSA variations are found. According to the trend and range rate changes the Amazon basin appears, over the last four decades in a-near equilibrium. For Europe and Australia no significant trend or acceleration pattern is found. The huge drying patterns of the Caspian Sea and the Aral Sea, clearly visible in GRACE observations, are not reflected by trend or accelerations over the fourty years. For this region moderate TWSA magnitudes are found on subseasonal and interannual scale. SLR on the other side, exhibits strong negative acceleration, with a slightly negative trend in this region.

Open Research Section

Data availability statement

CSR mascon (Save et al., 2016; Save, 2020): `https://www2.csr.utexas.edu/grace/RL06_mascons.html`; GRACE L2 data: GFZ (Dahle et al., 2018), ITSG2018 (Mayer-Gürr et al., 2018), CSR CSR (2018), all GRACE data sets were downloaded from `http://icgem.gfz-potsdam.de/series`; Reconstruction by F. Li (2021): `https://datadryad.org/stash/dataset/doi:10.5061/dryad.z612jm6bt`; Reconstruction by Humphrey (2019): `https://doi.org/10.6084/m9.figshare.7670849`; SLR hybrid solution by Löcher and Kusche (2020): `http://icgem.gfz-potsdam.de/series/04_SLR/IGG_SLR_HYBRID`; WaterGAP data (Müller Schmied et al., 2020b): `https://doi.pangaea.de/10.1594/PANGAEA.918447?format=html#download`

Acknowledgments

The authors acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—SFB 1502/1-2022—Project No. 450058266.

References

Altamimi, Z., Rebischung, P., Métivier, L., & Collilieux, X. (2016). ITRF2014: A new release of the International Terrestrial Reference Frame modeling nonlinear station motions. *Journal of Geophysical Research: Solid Earth*, *121*(8), 6109–6131. Retrieved 2023-03-13, from `https://onlinelibrary.wiley.com/doi/abs/10.1002/2016JB013098` doi: 10.1002/2016JB013098

Amarasekera, K. N., Lee, R. F., Williams, E. R., & Eltahir, E. A. B. (1997). ENSO and the natural variability in the flow of tropical rivers. *Jour-*

nal of Hydrology, *200*(1), 24–39. Retrieved 2022-12-02, from https://www.sciencedirect.com/science/article/pii/S0022169496033409 doi: 10.1016/S0022-1694(96)03340-9

Becker, M., Meyssignac, B., Xavier, L., Cazenave, A., Alkama, R., & Decharme, B. (2011). Past terrestrial water storage (1980–2008) in the Amazon Basin reconstructed from GRACE and in situ river gauging data. *Hydrology and Earth System Sciences*, *15*(2), 533–546. Retrieved 2022-07-08, from https://hess.copernicus.org/articles/15/533/2011/ doi: 10.5194/hess-15-533-2011

Boergens, E., Güntner, A., Dobslaw, H., & Dahle, C. (2020). Quantifying the Central European Droughts in 2018 and 2019 With GRACE Follow-On. *Geophysical Research Letters*, *47*(14). Retrieved 2023-01-09, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2020GL087285 doi: 10.1029/2020GL087285

Chanard, K., Fleitout, L., Calais, E., Rebischung, P., & Avouac, J.-P. (2018). Toward a Global Horizontal and Vertical Elastic Load Deformation Model Derived from GRACE and GNSS Station Position Time Series. *Journal of Geophysical Research: Solid Earth*, *123*(4), 3225–3237. Retrieved 2023-01-09, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2017JB015245 doi: 10.1002/2017JB015245

Chen, J., Li, J., Zhang, Z., & Ni, S. (2014). Long-term groundwater variations in Northwest India from satellite gravity measurements. *Global and Planetary Change*, *116*, 130–138. Retrieved 2022-11-25, from https://www.sciencedirect.com/science/article/pii/S0921818114000526 doi: 10.1016/j.gloplacha.2014.02.007

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*(12). Retrieved 2022-11-29, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2010WR009383 doi: 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 observations. *Global and Planetary Change*, *139*, 56–65. Retrieved 2022-11-25, from https://www.sciencedirect.com/science/article/pii/S0921818116300108 doi: 10.1016/j.gloplacha.2016.01.002

Chen, W., Li, J., Ray, J., & Cheng, M. (2017). Improved geophysical excitations constrained by polar motion observations and GRACE/SLR time-dependent gravity. *Geodesy and Geodynamics*, *8*(6), 377–388. Retrieved 2023-03-14, from https://www.sciencedirect.com/science/article/pii/S1674984717300800 doi: 10.1016/j.geog.2017.04.006

Cheng, M., & Ries, J. (2023). C20 and C30 Variations From SLR for GRACE/GRACE-FO Science Applications. *Journal of Geophysical Research: Solid Earth*, *128*(2). Retrieved 2023-03-13, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2022JB025459 doi: 10.1029/2022JB025459

Cheng, M., Ries, J. C., & Tapley, B. D. (2011). Variations of the Earth's figure axis from satellite laser ranging and GRACE. *Journal of Geophysical Research: Solid Earth*, *116*(B1). Retrieved 2023-03-13, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2010JB000850 doi: 10.1029/2010JB000850

Cheng, M. K., Ries, J. C., & Tapley, B. D. (2013). Geocenter Variations from Analysis of SLR Data. In Z. Altamimi & X. Collilieux (Eds.), *Reference Frames for Applications in Geosciences* (pp. 19–25). Berlin, Heidelberg: Springer. doi: 10.1007/978-3-642-32998-2_4

Church, J. A., White, N. J., Coleman, R., Lambeck, K., & Mitrovica, J. X. (2004). Estimates of the Regional Distribution of Sea Level Rise over the

1950–2000 Period. *Journal of Climate*, *17*(13), 2609–2625. Retrieved 2023-03-13, from https://journals.ametsoc.org/view/journals/clim/17/13/1520-0442_2004_017_2609_eotrdo_2.0.co_2.xml doi: 10.1175/1520-0442(2004)017⟨2609:EOTRDO⟩2.0.CO;2

Cleveland, R., Cleveland, W., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on Loess. *Journal of official statistics*, *6*, pp. 3–73.

Costa, M. H., & Foley, J. A. (1998). A comparison of precipitation datasets for the Amazon Basin. *Geophysical Research Letters*, *25*(2), 155–158. Retrieved 2022-11-29, from https://onlinelibrary.wiley.com/doi/abs/10.1029/97GL03502 doi: 10.1029/97GL03502

Cressie, N. (1993). *Statistics for Spatial Data.* Wiley-Interscience.

CSR. (2018). *GRACE FIELD GEOPOTENTIAL COEFFICIENTS CSR RELEASE 6.0| Physical Oceanography Distributed Active Archive Center (PO.DAAC).* Retrieved 2023-03-13, from https://podaac.jpl.nasa.gov/dataset/GRACE_GSM_L2_GRAV_CSR_RL06(lastaccess:23November2022)

Dahle, C., Flechtner, F., Murböck, M., Michalak, G., Neumayer, H., Abrykosov, O., … König, R. (2018). GRACE Geopotential GSM Coefficients GFZ RL06. V. 6.0. GFZ Data Services.
doi: https://doi.org/10.5880/GFZ.GRACE_06_GSM

Dahle, C., Flechtner, F., Murböck, M., Michalak, G., Neumayer, K.-H., Abrykosov, O., … König, R. (2019). GRACE-FO Geopotential GSM Coefficients GFZ RL06. V. 6.0. GFZ Data Services.
doi: https://doi.org/10.5880/GFZ.GRACEFO_06_GSM

Eicker, A., Forootan, E., Springer, A., Longuevergne, L., & Kusche, J. (2016). Does GRACE see the terrestrial water cycle "intensifying"? *Journal of Geophysical Research: Atmospheres*, *121*(2), 733–745. Retrieved 2022-09-23, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2015JD023808 doi: 10.1002/2015JD023808

Eicker, A., Schumacher, M., Kusche, J., Döll, P., & Schmied, H. M. (2014). Calibration/Data Assimilation Approach for Integrating GRACE Data into the WaterGAP Global Hydrology Model (WGHM) Using an Ensemble Kalman Filter: First Results. *Surveys in Geophysics*, *35*(6), 1285–1309. Retrieved 2022-09-08, from https://doi.org/10.1007/s10712-014-9309-8 doi: 10.1007/s10712-014-9309-8

Ferreira, V. G., Andam-Akorful, S. A., Dannouf, R., & Adu-Afari, E. (2019). A Multi-Sourced Data Retrodiction of Remotely Sensed Terrestrial Water Storage Changes for West Africa. *Water*, *11*(2), 401. Retrieved 2022-07-25, from https://www.mdpi.com/2073-4441/11/2/401 doi: 10.3390/w11020401

Forootan, E. (2014). *Statistical Signal Decomposition Techniques for Analyzing Time-Variable Satellite Gravimetry Data* (Thesis, Universitäts- und Landesbibliothek Bonn). Retrieved 2022-08-02, from https://bonndoc.ulb.uni-bonn.de/xmlui/handle/20.500.11811/5858

Forootan, E., Khandu, Awange, J. L., Schumacher, M., Anyah, R. O., van Dijk, A. I. J. M., & Kusche, J. (2016). Quantifying the impacts of ENSO and IOD on rain gauge and remotely sensed precipitation products over Australia. *Remote Sensing of Environment*, *172*, 50–66. Retrieved 2022-11-28, from https://www.sciencedirect.com/science/article/pii/S0034425715301796 doi: 10.1016/j.rse.2015.10.027

Forootan, E., Kusche, J., Loth, I., Schuh, W.-D., Eicker, A., Awange, J., … Shum, C. K. (2014). Multivariate Prediction of Total Water Storage Changes Over West Africa from Multi-Satellite Data. *Surveys in Geophysics*, *35*(4), 913–940. Retrieved from https://doi.org/10.1007/s10712-014-9292-0 doi: 10.1007/s10712-014-9292-0

Forootan, E., Schumacher, M., Mehrnegar, N., Bezděk, A., Talpe, M. J., Farzaneh, S., … Shum, C. K. (2020). An Iterative ICA-Based Reconstruction Method to Produce Consistent Time-Variable Total Water Storage Fields Using GRACE and Swarm Satellite Data. *Remote Sensing*, *12*(10), 1639. Retrieved 2022-07-07, from `https://www.mdpi.com/2072-4292/12/10/1639` doi: 10.3390/rs12101639

Frappart, F., Seoane, L., & Ramillien, G. (2013). Validation of GRACE-derived terrestrial water storage from a regional approach over South America. *Remote Sensing of Environment*, *137*, 69–83. Retrieved 2022-11-29, from `https://www.sciencedirect.com/science/article/pii/S0034425713002010` doi: 10.1016/j.rse.2013.06.008

García-García, D., Ummenhofer, C. C., & Zlotnicki, V. (2011). Australian water mass variations from GRACE data linked to Indo-Pacific climate variability. *Remote Sensing of Environment*, *115*(9), 2175–2183. Retrieved 2022-11-25, from `https://www.sciencedirect.com/science/article/pii/S0034425711001179` doi: 10.1016/j.rse.2011.04.007

Gerdener, H., Engels, O., & Kusche, J. (2020). A framework for deriving drought indicators from the Gravity Recovery and Climate Experiment (GRACE). *Hydrology and Earth System Sciences*, *24*(1), 227–248. Retrieved 2023-01-09, from `https://hess.copernicus.org/articles/24/227/2020/` doi: 10.5194/hess-24-227-2020

Geruo, A., Velicogna, I., Kimball, J. S., Du, J., Kim, Y., Colliander, A., & Njoku, E. (2017). Satellite-observed changes in vegetation sensitivities to surface soil moisture and total water storage variations since the 2011 Texas drought. *Environmental Research Letters*, *12*(5). Retrieved 2023-01-09, from `https://dx.doi.org/10.1088/1748-9326/aa6965` doi: 10.1088/1748-9326/aa6965

Girotto, M., De Lannoy, G. J. M., Reichle, R. H., & Rodell, M. (2016). Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model. *Water Resources Research*, *52*(5), 4164–4183. Retrieved 2023-03-13, from `https://onlinelibrary.wiley.com/doi/abs/10.1002/2015WR018417` doi: 10.1002/2015WR018417

Gudmundsson, L., & Seneviratne, S. I. (2015). European drought trends. *Proceedings of the International Association of Hydrological Sciences*, *369*, 75–79. Retrieved 2023-01-09, from `https://piahs.copernicus.org/articles/369/75/2015/` doi: 10.5194/piahs-369-75-2015

Han, S.-C., Ghobadi-Far, K., Yeo, I.-Y., McCullough, C. M., Lee, E., & Sauber, J. (2021). GRACE Follow-On revealed Bangladesh was flooded early in the 2020 monsoon season due to premature soil saturation. *Proceedings of the National Academy of Sciences*, *118*(47). Retrieved from `https://www.pnas.org/doi/abs/10.1073/pnas.2109086118` doi: 10.1073/pnas.2109086118

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., … Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. Retrieved 2022-05-18, from `https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803` doi: 10.1002/qj.3803

Humphrey, V. (2019). GRACE-REC: A reconstruction of climate-driven water storage changes over the last century [dataset]. doi: 10.6084/m9.figshare.7670849.v3

Humphrey, V., & Gudmundsson, L. (2019). GRACE-REC: a reconstruction of climate-driven water storage changes over the last century. *Earth System Science Data*, *11*(3), 1153–1170. Retrieved from `https://essd.copernicus.org/articles/11/1153/2019/` doi: 10.5194/essd-11-1153-2019

Humphrey, V., Gudmundsson, L., & Seneviratne, S. I. (2016). Assessing Global Water Storage Variability from GRACE: Trends, Seasonal Cycle, Subseasonal Anomalies and Extremes. *Surveys in Geophysics*, *37*(2), 357–395. Retrieved

2022-05-24, from http://link.springer.com/10.1007/s10712-016-9367-1 doi: 10.1007/s10712-016-9367-1

Jensen, L., Eicker, A., Dobslaw, H., Stacke, T., & Humphrey, V. (2019). Long-Term Wetting and Drying Trends in Land Water Storage Derived From GRACE and CMIP5 Models. *Journal of Geophysical Research: Atmospheres*, *124*(17-18), 9808–9823. Retrieved 2023-03-13, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2018JD029989 doi: 10.1029/2018JD029989

Jin, S., Chambers, D. P., & Tapley, B. D. (2010). Hydrological and oceanic effects on polar motion from GRACE and models. *Journal of Geophysical Research: Solid Earth*, *115*(B2). Retrieved 2023-01-09, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2009JB006635 doi: 10.1029/2009JB006635

Jing, W., Zhang, P., & Zhao, X. (2019). A comparison of different GRACE solutions in terrestrial water storage trend estimation over Tibetan Plateau. *Scientific Reports*, *9*(1). Retrieved 2022-12-14, from https://www.nature.com/articles/s41598-018-38337-1 doi: 10.1038/s41598-018-38337-1

Joodaki, G., Wahr, J., & Swenson, S. (2014). Estimating the human contribution to groundwater depletion in the Middle East, from GRACE data, land surface models, and well observations. *Water Resources Research*, *50*(3), 2679–2692. Retrieved 2022-09-23, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2013WR014633 doi: 10.1002/2013WR014633

Khaki, M., Hoteit, I., Kuhn, M., Awange, J., Forootan, E., van Dijk, A. I. J. M., ... Pattiaratchi, C. (2017). Assessing sequential data assimilation techniques for integrating GRACE data into a hydrological model. *Advances in Water Resources*, *107*, 301–316. Retrieved 2023-03-13, from https://www.sciencedirect.com/science/article/pii/S0309170816307564 doi: 10.1016/j.advwatres.2017.07.001

Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World Map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, *15*(3), 259–263. Retrieved 2022-07-12, from http://www.schweizerbart.de/papers/metz/detail/15/55034/World_Map_of_the_Koppen_Geiger_climate_classificat?af=crossref doi: 10.1127/0941-2948/2006/0130

Kumar, V., Jain, S. K., & Singh, Y. (2010). Analysis of long-term rainfall trends in India. *Hydrological Sciences Journal*, *55*(4), 484–496. Retrieved 2022-12-14, from https://doi.org/10.1080/02626667.2010.481373 doi: 10.1080/02626667.2010.481373

Kusche, J. (2007). Approximate decorrelation and non-isotropic smoothing of time-variable GRACE-type gravity field models. *Journal of Geodesy*, *81*(11), 733–749. Retrieved 2022-11-15, from https://doi.org/10.1007/s00190-007-0143-3 doi: 10.1007/s00190-007-0143-3

Kusche, J., Eicker, A., Forootan, E., Springer, A., & Longuevergne, L. (2016). Mapping probabilities of extreme continental water storage changes from space gravimetry. *Geophysical Research Letters*, *43*(15), 8026–8034. Retrieved 2023-01-09, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2016GL069538 doi: 10.1002/2016GL069538

Köppen, W. (1923). *Die Klimate der Erde: Grundriss der Klimakunde.* De Gruyter. Retrieved 2022-07-14, from https://www.degruyter.com/document/doi/10.1515/9783111491530/html?lang=de doi: 10.1515/9783111491530

Lenczuk, A., Weigelt, M., Kosek, W., & Mikocki, J. (2022). Autoregressive Reconstruction of Total Water Storage within GRACE and GRACE Follow-On Gap Period. *Energies*, *15*(13). Retrieved 2022-07-07, from https://www.mdpi.com/1996-1073/15/13/4827 doi: 10.3390/en15134827

Li, B., Rodell, M., Kumar, S., Beaudoing, H. K., Getirana, A., Zaitchik, B. F., ... Bettadpur, S. (2019). Global GRACE Data Assimilation for Groundwater and Drought Monitoring: Advances and Challenges. *Wa-*

ter Resources Research, *55*(9), 7564–7586. Retrieved 2023-01-09, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2018WR024618 doi: 10.1029/2018WR024618

Li, F. (2021). Long-term (1979-present) total water storage anomalies over the global land derived by reconstructing GRACE data [dataset]. Retrieved from http://datadryad.org/stash/dataset/doi:10.5061/dryad.z612jm6bt doi: 10.5061/DRYAD.Z612JM6BT

Li, F., Kusche, J., Chao, N., Wang, Z., & Löcher, A. (2021). Long-Term (1979-Present) Total Water Storage Anomalies Over the Global Land Derived by Reconstructing GRACE Data. *Geophysical Research Letters*, *48*(8). Retrieved 2022-07-18, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2021GL093492 doi: 10.1029/2021GL093492

Li, F., Kusche, J., Rietbroek, R., Wang, Z., Forootan, E., Schulze, K., & Lück, C. (2020). Comparison of data-driven techniques to reconstruct (1992-2002) and predict (2017-2018) GRACE-like gridded total water storage changes using climate inputs. *Water Resources Research, 56, e2019WR026551*. doi: 10.1029/2019wr026551

Long, D., Shen, Y., Sun, A., Hong, Y., Longuevergne, L., Yang, Y., … Chen, L. (2014). Drought and flood monitoring for a large karst plateau in Southwest China using extended GRACE data. *Remote Sensing of Environment*, *155*, 145–160. Retrieved from https://www.sciencedirect.com/science/article/pii/S0034425714003034 doi: https://doi.org/10.1016/j.rse.2014.08.006

Loomis, B. D., & Luthcke, S. B. (2017). Mass evolution of Mediterranean, Black, Red, and Caspian Seas from GRACE and altimetry: accuracy assessment and solution calibration. *Journal of Geodesy*, *91*, 195–206. Retrieved from https://doi.org/10.1007/s00190-016-0952-3 doi: 10.1007/s00190-016-0952-3

Löcher, A., & Kusche, J. (2020). A hybrid approach for recovering high-resolution temporal gravity fields from satellite laser ranging. *Journal of Geodesy*, *95*(1), 6. Retrieved from https://doi.org/10.1007/s00190-020-01460-x doi: 10.1007/s00190-020-01460-x

Marengo, J. A., & Espinoza, J. C. (2016). Extreme seasonal droughts and floods in Amazonia: causes, trends and impacts. *International Journal of Climatology*, *36*(3), 1033–1050. Retrieved 2022-11-30, from https://onlinelibrary.wiley.com/doi/abs/10.1002/joc.4420 doi: 10.1002/joc.4420

Marengo, J. A., Liebmann, B., Grimm, A. M., Misra, V., Silva Dias, P. L., Cavalcanti, I. F. A., … Alves, L. M. (2012). Recent developments on the South American monsoon system. *International Journal of Climatology*, *32*(1), 1–21. Retrieved 2022-11-30, from https://onlinelibrary.wiley.com/doi/abs/10.1002/joc.2254 doi: 10.1002/joc.2254

Marengo, J. A., Tomasella, J., Soares, W. R., Alves, L. M., & Nobre, C. A. (2012). Extreme climatic events in the Amazon basin. *Theoretical and Applied Climatology*, *107*(1), 73–85. Retrieved 2022-11-30, from https://doi.org/10.1007/s00704-011-0465-1 doi: 10.1007/s00704-011-0465-1

Mayer-Gürr, T., Behzadpur, S., Ellmer, M., Kvas, A., Klinger, B., Strasser, S., & Zehentner, N. (2018). *ITSG-Grace2018 - Monthly, Daily and Static Gravity Field Solutions from GRACE.* Retrieved 2022-12-14, from https://dataservices.gfz-potsdam.de/icgem/showshort.php?id=escidoc:3600910 doi: 10.5880/ICGEM.2018.003

MIRZA, M. Q., WARRICK, R. A., ERICKSEN, N. J., & KENNY, G. J. (1998). Trends and persistence in precipitation in the Ganges, Brahmaputra and Meghna river basins. *Hydrological Sciences Journal*, *43*(6), 845–858. Retrieved 2022-12-14, from https://doi.org/10.1080/02626669809492182 doi: 10.1080/02626669809492182

Mémin, A., Boy, J.-P., & Santamaría-Gómez, A. (2020). Correcting GPS measurements for non-tidal loading. *GPS Solutions*, *24*(2), 45. Retrieved 2023-01-09, from https://doi.org/10.1007/s10291-020-0959-3 doi: 10.1007/s10291-020-0959-3

Müller Schmied, H., Cáceres, D., Eisner, S., Flörke, M., Herbert, C., Niemann, C., ... Döll, P. (2020a). The global water resources and use model WaterGAP v2.2d - Standard model output. Retrieved 2022-07-27, from https://doi.pangaea.de/10.1594/PANGAEA.918447 doi: 10.1594/PANGAEA.918447

Müller Schmied, H., Cáceres, D., Eisner, S., Flörke, M., Herbert, C., Niemann, C., ... Döll, P. (2020b). The global water resources and use model WaterGAP v2.2d - Standard model output [dataset]. Retrieved from https://doi.pangaea.de/10.1594/PANGAEA.918447 doi: 10.1594/PANGAEA.918447

Nicholls, N., Drosdowsky, W., & Lavery, B. (1997). Australian rainfall variability and change. *Weather*, *52*(3), 66–72. Retrieved 2023-05-04, from https://onlinelibrary.wiley.com/doi/abs/10.1002/j.1477-8696.1997.tb06274.x doi: 10.1002/j.1477-8696.1997.tb06274.x

Nicholson, S. E. (2022). The Rainfall and Convective Regime over Equatorial Africa, with Emphasis on the Congo Basin. In *Congo Basin Hydrology, Climate, and Biogeochemistry* (pp. 25–48). Retrieved 2022-12-02, from https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119657002.ch3 doi: 10.1002/9781119657002.ch3

Pearlman, M., Arnold, D., Davis, M., Barlier, F., Biancale, R., Vasiliev, V., ... Bloßfeld, M. (2019). Laser geodetic satellites: a high-accuracy scientific tool. *Journal of Geodesy*, *93*(11), 2181–2194. Retrieved 2023-01-09, from https://doi.org/10.1007/s00190-019-01228-y doi: 10.1007/s00190-019-01228-y

Peltier, R. W., Argus, D. F., & Drummond, R. (2018). Comment on "An Assessment of the ICE-6G_c (VM5a) Glacial Isostatic Adjustment Model" by Purcell et al. *Journal of Geophysical Research: Solid Earth*, *123*(2), 2019–2028. Retrieved 2023-01-25, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2016JB013844 doi: 10.1002/2016JB013844

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*(16). Retrieved 2022-12-02, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2012GL052495 doi: 10.1029/2012GL052495

Ray, R. D., Luthcke, S. B., & Dam, T. v. (2013). Monthly Crustal Loading Corrections for Satellite Altimetry. *Journal of Atmospheric and Oceanic Technology*, *30*(5), 999–1005. Retrieved 2023-01-09, from https://journals.ametsoc.org/view/journals/atot/30/5/jtech-d-12-00152_1.xml doi: 10.1175/JTECH-D-12-00152.1

Reager, J. T., Thomas, B. F., & Famiglietti, J. S. (2014). River basin flood potential inferred using GRACE gravity observations at several months lead time. *Nature Geoscience*, *7*(8), 588–592. Retrieved 2023-01-09, from https://www.nature.com/articles/ngeo2203 doi: 10.1038/ngeo2203

Richey, A. S., Thomas, B. F., Lo, M.-H., Reager, J. T., Famiglietti, J. S., Voss, K., ... Rodell, M. (2015). Quantifying renewable groundwater stress with GRACE. *Water Resources Research*, *51*(7), 5217–5238. Retrieved 2022-09-23, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2015WR017349 doi: 10.1002/2015WR017349

Richter, H. M. P., Lück, C., Klos, A., Sideris, M. G., Rangelova, E., & Kusche, J. (2021). Reconstructing GRACE-type time-variable gravity from the Swarm satellites. *Scientific Reports*, *11*(1), 1117. Retrieved 2022-07-08, from https://www.nature.com/articles/s41598-020-80752-w doi: 10.1038/s41598-020-80752-w

Risbey, J., Pook, M., McIntosh, P., Wheeler, M., & Hendon, H. (2009). On the Remote Drivers of Rainfall Variability in Australia. *Monthly Weather Review - MON WEATHER REV*, *137*, 3233–3253. doi: 10.1175/2009MWR2861.1

Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., & Lo, M.-H. (2018). Emerging trends in global freshwater availability. *Nature*, *557*(7707), 651–659. Retrieved 2022-10-13, from https://www.nature.com/articles/s41586-018-0123-1 doi: 10.1038/s41586-018-0123-1

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., ... Toll, D. (2004). The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society*, *85*(3), 381–394. Retrieved 2023-01-09, from https://journals.ametsoc.org/view/journals/bams/85/3/bams-85-3-381.xml doi: 10.1175/BAMS-85-3-381

Save, H. (2020). *"CSR GRACE and GRACE-FO RL06 Mascon Solutions v02.* doi: doi:10.15781/cgq9-nh24

Save, H., Bettadpur, S., & Tapley, B. D. (2016). High-resolution CSR GRACE RL05 mascons. *Journal of Geophysical Research: Solid Earth*, *121*(10), 7547–7569. Retrieved 2022-12-14, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2016JB013007 doi: 10.1002/2016JB013007

Scanlon, B. R., Zhang, Z., Save, H., Sun, A. Y., Müller Schmied, H., van Beek, L. P. H., ... Bierkens, M. F. P. (2018). Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. *Proceedings of the National Academy of Sciences*, *115*(6). Retrieved 2022-11-29, from https://www.pnas.org/doi/full/10.1073/pnas.1704665115 doi: 10.1073/pnas.1704665115

Schumacher, M., Forootan, E., van Dijk, A. I. J. M., Müller Schmied, H., Crosbie, R. S., Kusche, J., & Döll, P. (2018). Improving drought simulations within the Murray-Darling Basin by combined calibration/assimilation of GRACE data into the WaterGAP Global Hydrology Model. *Remote Sensing of Environment*, *204*, 212–228. Retrieved 2022-11-28, from https://www.sciencedirect.com/science/article/pii/S0034425717304923 doi: 10.1016/j.rse.2017.10.029

Seoane, L., Nastula, J., Bizouard, C., & Gambis, D. (2009). The use of gravimetric data from GRACE mission in the understanding of polar motion variations. *Geophysical Journal International*, *178*(2), 614–622. Retrieved 2023-01-09, from https://doi.org/10.1111/j.1365-246X.2009.04181.x doi: 10.1111/j.1365-246X.2009.04181.x

Sośnica, K., Jäggi, A., Meyer, U., Thaller, D., Beutler, G., Arnold, D., & Dach, R. (2015). Time variable Earth's gravity field from SLR satellites. *Journal of Geodesy*, *89*(10), 945–960. Retrieved 2023-01-09, from https://doi.org/10.1007/s00190-015-0825-1 doi: 10.1007/s00190-015-0825-1

Spinoni, J., Naumann, G., Vogt, J., & Barbosa, P. (2015b). European drought climatologies and trends based on a multi-indicator approach. *Global and Planetary Change*, *127*, 50–57. Retrieved 2023-01-09, from https://www.sciencedirect.com/science/article/pii/S0921818115000284 doi: 10.1016/j.gloplacha.2015.01.012

Spinoni, J., Naumann, G., Vogt, J. V., & Barbosa, P. (2015a). The biggest drought events in Europe from 1950 to 2012. *Journal of Hydrology: Regional Studies*, *3*, 509–524. Retrieved 2023-01-09, from https://www.sciencedirect.com/science/article/pii/S2214581815000026 doi: 10.1016/j.ejrh.2015.01.001

Springer, A., Karegar, M. A., Kusche, J., Keune, J., Kurtz, W., & Kollet, S. (2019). Evidence of daily hydrological loading in GPS time series over Europe. *Journal of Geodesy*, *93*(10), 2145–2153. Retrieved 2023-01-09, from https://doi.org/10.1007/s00190-019-01295-1 doi: 10.1007/s00190-019-01295-1

Sun, A. Y., Scanlon, B. R., Save, H., & Rateb, A. (2021). Reconstruction of GRACE Total Water Storage Through Automated Machine Learning. *Water Resources Research*, *57*(2). Retrieved 2022-07-07, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2020WR028666 doi: 10.1029/2020WR028666

Sun, A. Y., Scanlon, B. R., Zhang, Z., Walling, D., Bhanja, S. N., Mukherjee, A., & Zhong, Z. (2019). Combining Physically Based Modeling and Deep Learning for Fusing GRACE Satellite Data: Can We Learn From Mismatch? *Water Resources Research*, *55*(2), 1179–1195. Retrieved 2022-07-25, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2018WR023333 doi: 10.1029/2018WR023333

Sun, Z., Long, D., Yang, W., Li, X., & Pan, Y. (2020). Reconstruction of GRACE Data on Changes in Total Water Storage Over the Global Land Surface and 60 Basins. *Water Resources Research*, *56*(4). Retrieved 2022-07-08, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2019WR026250 doi: 10.1029/2019WR026250

Swenson, S. C., & Lawrence, D. M. (2014). Assessing a dry surface layer-based soil resistance parameterization for the Community Land Model using GRACE and FLUXNET-MTE data. *Journal of Geophysical Research: Atmospheres*, *119*(17), 10,299–10,312. Retrieved 2023-01-09, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2014JD022314 doi: 10.1002/2014JD022314

Tang, S., Wang, H., Feng, Y., Liu, Q., Wang, T., Liu, W., & Sun, F. (2021). Random Forest-Based Reconstruction and Application of the GRACE Terrestrial Water Storage Estimates for the Lancang-Mekong River Basin. *Remote Sensing*, *13*(23), 4831. Retrieved 2022-07-07, from https://www.mdpi.com/2072-4292/13/23/4831 doi: 10.3390/rs13234831

Tangdamrongsub, N., Han, S.-C., Yeo, I.-Y., Dong, J., Steele-Dunne, S. C., Willgoose, G., & Walker, J. P. (2020). Multivariate data assimilation of GRACE, SMOS, SMAP measurements for improved regional soil moisture and groundwater storage estimates. *Advances in Water Resources*, *135*, 103477. Retrieved 2023-01-09, from https://www.sciencedirect.com/science/article/pii/S0309170819302970 doi: 10.1016/j.advwatres.2019.103477

Tangdamrongsub, N., Steele-Dunne, S. C., Gunter, B. C., Ditmar, P. G., & Weerts, A. H. (2015). Data assimilation of GRACE terrestrial water storage estimates into a regional hydrological model of the Rhine River basin. *Hydrology and Earth System Sciences*, *19*(4), 2079–2100. Retrieved 2023-01-09, from https://hess.copernicus.org/articles/19/2079/2015/ doi: 10.5194/hess-19-2079-2015

Tapley, B. D., Watkins, M. M., Flechtner, F., Reigber, C., Bettadpur, S., Rodell, M., … Velicogna, I. (2019). Contributions of GRACE to understanding climate change. *Nature Climate Change*, *9*(5), 358–369. Retrieved 2022-07-25, from http://www.nature.com/articles/s41558-019-0456-2 doi: 10.1038/s41558-019-0456-2

Towner, J., Ficchí, A., Cloke, H. L., Bazo, J., Coughlan de Perez, E., & Stephens, E. M. (2021). Influence of ENSO and tropical Atlantic climate variability on flood characteristics in the Amazon basin. *Hydrology and Earth System Sciences*, *25*(7), 3875–3895. Retrieved 2022-11-30, from https://hess.copernicus.org/articles/25/3875/2021/ doi: 10.5194/hess-25-3875-2021

Trenberth, K. E. (1990). Recent Observed Interdecadal Climate Changes in the Northern Hemisphere. *Bulletin of the American Meteorological Society*, *71*(7), 988–993. Retrieved 2023-05-04, from https://journals.ametsoc.org/view/journals/bams/71/7/1520-0477_1990_071_0988_roicci_2_0_co_2.xml doi: 10.1175/1520-0477(1990)071⟨0988:ROICCI⟩2.0.CO;2

van Dijk, A. I. J. M., Beck, H. E., Crosbie, R. S., de Jeu, R. A. M., Liu, Y. Y., Podger, G. M., ... Viney, N. R. (2013). The Millennium Drought in southeast Australia (2001–2009): Natural and human causes and implications for water resources, ecosystems, economy, and society. *Water Resources Research*, *49*(2), 1040–1057. Retrieved 2022-11-28, from https://onlinelibrary.wiley.com/doi/abs/10.1002/wrcr.20123 doi: 10.1002/wrcr.20123

Voss, K. A., Famiglietti, J. S., Lo, M., de Linage, C., Rodell, M., & Swenson, S. C. (2013). Groundwater depletion in the Middle East from GRACE with implications for transboundary water management in the Tigris-Euphrates-Western Iran region. *Water Resources Research*, *49*(2), 904–914. Retrieved 2022-11-29, from https://onlinelibrary.wiley.com/doi/abs/10.1002/wrcr.20078 doi: 10.1002/wrcr.20078

Wu, X., & Heflin, M. B. (2015). A global assessment of accelerations in surface mass transport. *Geophysical Research Letters*, *42*(16), 6716–6723. Retrieved 2022-11-22, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2015GL064941 doi: 10.1002/2015GL064941

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*(8), 1629–1637. Retrieved 2022-11-30, from https://www.sciencedirect.com/science/article/pii/S0034425710000647 doi: 10.1016/j.rse.2010.02.005

Yang, X., Tian, S., Feng, W., Ran, J., You, W., Jiang, Z., & Gong, X. (2020). Spatio-Temporal Evaluation of Water Storage Trends from Hydrological Models over Australia Using GRACE Mascon Solutions. *Remote Sensing*, *12*(21). Retrieved 2022-11-28, from https://www.mdpi.com/2072-4292/12/21/3578 doi: 10.3390/rs12213578

Yu, Q., Wang, S., He, H., Yang, K., Ma, L., & Li, J. (2021). Reconstructing GRACE-like TWS anomalies for the Canadian landmass using deep learning and land surface model. *International Journal of Applied Earth Observation and Geoinformation*, *102*. Retrieved 2022-07-07, from https://www.sciencedirect.com/science/article/pii/S0303243421001112 doi: 10.1016/j.jag.2021.102404

Zaitchik, B. F., Rodell, M., & Reichle, R. H. (2008). Assimilation of GRACE Terrestrial Water Storage Data into a Land Surface Model: Results for the Mississippi River Basin. *Journal of Hydrometeorology*, *9*(3), 535–548. Retrieved 2023-01-09, from https://journals.ametsoc.org/view/journals/hydr/9/3/2007jhm951_1.xml doi: 10.1175/2007JHM951.1

Zhang, D., Zhang, Q., Werner, A. D., & Liu, X. (2016). GRACE-Based Hydrological Drought Evaluation of the Yangtze River Basin, China. *Journal of Hydrometeorology*, *17*(3), 811–828. Retrieved 2022-07-25, from https://journals.ametsoc.org/view/journals/hydr/17/3/jhm-d-15-0084_1.xml doi: 10.1175/JHM-D-15-0084.1

Zhao, M., A, G., Velicogna, I., & Kimball, J. S. (2017). A Global Gridded Dataset of GRACE Drought Severity Index for 2002–14: Comparison with PDSI and SPEI and a Case Study of the Australia Millennium Drought. *Journal of Hydrometeorology*, *18*(8), 2117–2129. Retrieved 2023-01-09, from https://journals.ametsoc.org/view/journals/hydr/18/8/jhm-d-16-0182_1.xml doi: 10.1175/JHM-D-16-0182.1

Śliwińska, J., Nastula, J., & Wińska, M. (2021). Evaluation of hydrological and cryospheric angular momentum estimates based on GRACE, GRACE-FO and SLR data for their contributions to polar motion excitation. *Earth, Planets and Space*, *73*(1), 71. Retrieved 2023-03-14, from https://doi.org/10.1186/s40623-021-01393-5 doi: 10.1186/s40623-021-01393-5

Appendix A Data processing

The functional relationship to derive trend, acceleration, annual amplitude and phase reads

$$\vec{f}(\vec{t}) = a_0 + a_1(\vec{t} - t_0) + a_2(\vec{t} - t_0)^2 + c_1 \cos(2\pi(\vec{t} - t_0)) + s_1 \sin(2\pi(\vec{t} - t_0)), \tag{A1}$$

where $\vec{x} = [a_0, a_1, a_2, c_1, s_1]^T$ are the parameters estimated in the least square process and t_0 is a reference epoch. For this study the data has been referenced to the time of the first observation. As HG19 does not include a trend the equation above reduces to

$$\vec{f}(\vec{t}) = a_0 + a_2(\vec{t} - t_0)^2 + c_1 \cos(2\pi(\vec{t} - t_0)) + s_1 \sin(2\pi(\vec{t} - t_0)). \tag{A2}$$

As trend and acceleration estimated by the least squares approach differ, depending on whether parameters are estimated separately or joinly, we first estimated the trend for the other data sets, reduced them by the estimated trend and then derived acceleration, annual amplitude and phase. The annual amplitude and phase are computed based on the estimated coefficients c_1, s_1. For better interpretability the annual phase is expressed as the day of the signal minimum via

$$D_{\mathrm{Min}} = \frac{\arctan(s_1, c_1) + \pi}{2\pi} 365, 25. \tag{A3}$$

The annual amplitude is computed as

$$A = \sqrt{s_1^2 + c_1^2} \tag{A4}$$

We define interannual signal as all signal components with a period longer than a year. The interannual signal was derived based on a combination of filtering and least squares adjustment. In a first step the annual and semi-annual periods are computed and reduced from the data sets based on a least square adjustment fitting annual and semi-annual coefficients to the time series. A low pass filter is applied to reduce all frequencies greater than 1. The filter weights are derived for an ideal lowpass filter, which is defined as a filter that allows all frequencies greater than a given cutting frequency (in our case $>$ 1 year) to pass and set all frequencies smaller then the cutting frequency ($<$ 1 year) to zero. The derived coefficients, $c_{|k|}$, reads

$$c_{|k|} = c_{|k|}\sigma^N_{|k|} = \sum_{k=0}^{N} \frac{\nu_{\mathrm{stop}}}{\nu_N} \mathrm{sinc}\left(k\frac{\nu_{\mathrm{stop}}}{\nu_N}\right) \mathrm{sinc}\left(\frac{k}{N+1}\right) \tag{A5}$$

The filter length, N, was set to 13 months and the cutting frequency ν_{stop} to 1 year. The derived filter coefficients are finite. A Lanczos smoothing hamming$_{d}igital_{1}989, heredenotedas$ $\sigma^N_{|k|}$ is used to smooth the signal around the cutting frequency ν_{stop}.

We define subseasonal to be all signal parts $<$ 1 year. The dominant annual signal is removed using a least square adjustment. All signals $>$ 1 year are reduced using a high pass filter. Assuming, the input signal,$\{u_n\}_{\Delta x}$, can be written as $\{u_n\}_{\Delta x} = \{y_n\}_{\Delta x} + \{z_n\}_{\Delta x}$ with $\{y_n\}_{\Delta x}$ containing the low frequencies, i.e the output of the low pass filter, and $\{z_n\}_{\Delta x}$ the high frequencies, the high pass filter can be derived as $\{z_n\}_{\Delta x} = \{u_n\}_{\Delta x} - \{y_n\}_{\Delta x}$.

It should be mentioned, that the filtering operation, even though effectively reducing all unwanted frequencies also damps signal we are interested in. The frequency operation might also cause leakage, so a smearing of frequencies in others.

Appendix B Different reconstructions of GRACE like TWSA

Table Appendix B.1 gives an overview of different GRACE like TWSA reconstructions, including reconstruction period, region and employed methods.

Authors	Predictors	employed methods	Time period	Area	Data access
Humphrey and Gudmundsson (2019)	P, T	exponential first order decay function	01.1979 – 07.2019	global	`https://doi.org/10.6084/m9.figshare.7670849`
F. Li et al. (2021)	P, T, SST, climate indices	PCA,ICA,STL, LS, ANN; MLR, ARX	07.1979 – 06.2020	global	`https://doi.org/10.5061/dryad.z612jm6bt`
A. Y. Sun et al. (2021)	P, T, SST, NAO, MEI,GLDAS TWS	AutoML	06.2017 – 12.2019	conterminous U.S (CONUS)	on request
Forootan et al. (2020)	Swarm	ICA	2002 – 2020	global	`https://www.mdpi.com/2072-4292/12/10/1639/s1`
Z. Sun et al. (2020)	P, T, GLDAS Noah TWS	DNN, MLR, SARIMAX	04.2002 – 06.2018	60 basins	on request
Tang et al. (2021)	GLDAS Noah TWS, P, T, meteorological data	RF	1980 – 2014	Lancang-Mekong River Basin	on request
Yu et al. (2021)	EALCO TWSA	CNN, cGAN, DCAE, ConvL-STM	1979 – 2002	Canadian landmass	on request
Lenczuk et al. (2022)		forward and backward AR process	07.2017 – 05.2018	global	on request
A. Y. Sun et al. (2019)	GLDAS Noah TWS	CNN	04.2002 – 06.2017	India	on request
Ferreira et al. (2019)	T, P, E, SM, R, climate indices	NARX	1979 – 2013	West Africa	on request
Becker et al. (2011)	in situ river level records	PCA / EOF	1980 – 2008	Amazon basin	on request
Long et al. (2014)	SMS, P, T	MLP ANN	02.1979 – 12.2002	karst Plateau basin	on request
Richter et al. (2021)	SWARM	PCA	07.2017 – 05.2018	global	on request
Zhang et al. (2016)	SM, P	ANN	1979 – 2012	Yangtze river basin	on request

Table Appendix B.1: Overview: Different reconstructions of GRACE like TWSA

Note: P precipation, T land surface temperature, SST sea surface temperature, NAO North Atlantic Oscillation, MEI Multivariate ENSO index, E evaporation, SM soil moisture, R rainfall, NDVI normalized difference vegetation index, M mascon, SH spherical harmonics, PCA principal component analysis, ICA indipentend component analysis, STL seasonal-trend decomposition based on loess, LS least square, ANN artificial neural network, MLR multi linear regression, ARX autoregressive exogenous model, AutoML automated machine learning, DNN deep neural network, SARIMAX seasonal ARIMA (autoregressiv integrated moving average) with exogenous variables, RF random forest, CNN convolutional neural network, cGAN conditional generative adversarial network, DCAE deep convolutional autoencoder, ConvLSTM convolutional long short term memory, AR (process) auto regressive (process), NARX nonlinear autoregressive with exogenous input, MLP multi-layer percepton

Appendix C Trend, acceleration and annual amplitude and phase for the reconstructions, WaterGAP and GRACE

Table Appendix C.1 shows the derived average trend, acceleration, average annual phase and minimal day of the seasonal cycle for the reocntructions and the hydrological model WaterGAP for 1979-2016 for the global land. Table Appendix C.2 reveal the same, but for the Pre GRACE time frame. Average trend, average rate range change, annual amplitude and minimal day of the seasonal cycle for the GRACE time frame for four different GRACE solutions and the reconstructions for the global land are shown in table Appendix C.3. Figure Appendix C.1 present the respective time series.

Appendix D Long term evolution of TWS

Figure Appendix D.1 reveal the water mass change for the detrended and deseasonalized reconstructions for 1979-2016 for nine major river basins. The time series of HG19 is presented in red, Li21 in black. The according location of the basins are shown in figure Appendix D.2. The time series are discussed in section 4.3.

Appendix E Supplementary results

The average trend, acceleration, average annual amplitude and phase, subseasonal and interannual signal variability of water storage for Europe is shown in figure Appendix E.1, for the years 1979-2016, and in figure Appendix E.2 for the pre GRACE era. The results are discussed in section 3.3.3. Figure Appendix E.3 reveals the autocorrelation pattern of the detrended and deseasonalized reconstructions for a point in the Amazon basin. Figure Appendix E.4 shows the same, but for a point in Europe. The figures are discussed in section 3.3.2 and 3.3.2.

	Trend* $[mm/yr]$	Trend$^{\diamond}$ $[mm/yr]$	Acceleration$^{\diamond}$ $[mm/yr^2]$	annual amplitude$^{*\diamond}$ $[mm]$	annual phase$^{*\diamond}$ $[d]$
Li21	-1.59	-1.63	0.001	24.29	83
HG19	-0.57	-0.0135	-0.0137	21.95	266
WaterGAP	-0.69	-0.56	-0.003	23.85	246

Table Appendix C.1. Trend, acceleration and annual amplitude and annual phase derived from the two reconstructions and the hydrological model WaterGAP for the global land for the time frame 1979 – 2016, without (1st column) and with (2nd column) estimating acceleration

	Trend* $[mm/yr]$	Trend$^{\diamond}$ $[mm/yr]$	Acceleration$^{\diamond}$ $[mm/yr^2]$	annual amplitude$^{*\diamond}$ $[mm]$	annual phase$^{*\diamond}$ $[d]$
Li21	-1.62	-2.61	0.04	24.29	83
HG19	-0.04	0.57	-0.03	21.73	267
WaterGAP	-0.65	-1.22	0.03	23.85	246

Table Appendix C.2. Trend, acceleration and annual amplitude and annual phase derived from the two reconstructions and the hydrological model WaterGAP for the global land for the time frame 1979 – 2002, without (1st column) and with (2nd column) estimating acceleration

	Trend* $[mm/yr]$	Trend$^{\diamond}$ $[mm/yr]$	Acceleration$^{\diamond}$ $[mm/yr^2]$	annual amplitude$^{*\diamond}$ $[mm]$	annual phase$^{*\diamond}$ $[d]$
ITSG SH 2018	-0.22	0.63	-0.04	18.58	180
CSR SH RL06	-0.15	0.69	-0.04	18.5	181
GFZ SH RL06	-0.15	0.59	-0.04	13.31	180
CSR M RL06, v.2	-1.76	-1.13	-0.03	24.30	173
Li21	-1.66	-0.62	-0.06	24.30	174
HG19	-0.16	0.61	-0.04	22.30	205

Table Appendix C.3. Trend, acceleration and annual amplitude and annual phase derived from different GRACE/GRACE-FO solutions and the reconstructions for the global land for the time frame 2002 – 2020, without (1st column) and with (2nd column) estimating acceleration

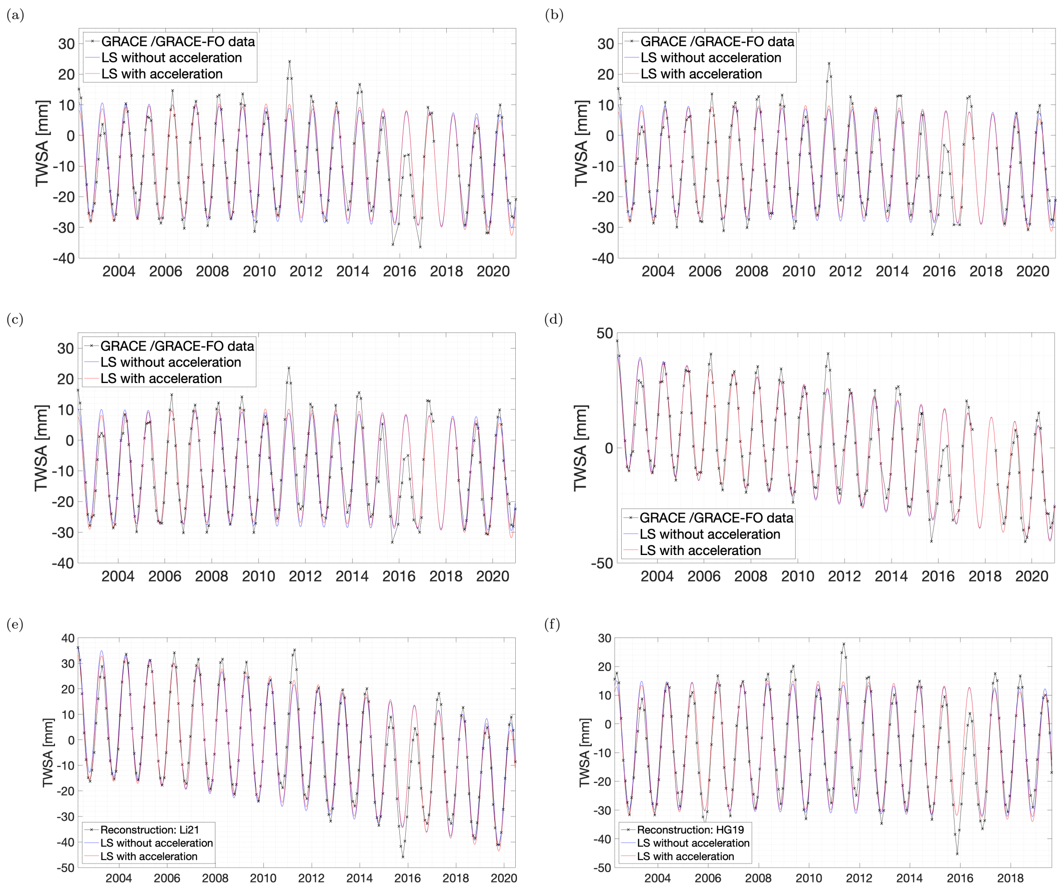

Figure Appendix C.1. Time series for four different GRACE solutions and the two global reconstructions for the global land for the time frame 2002 – 2020

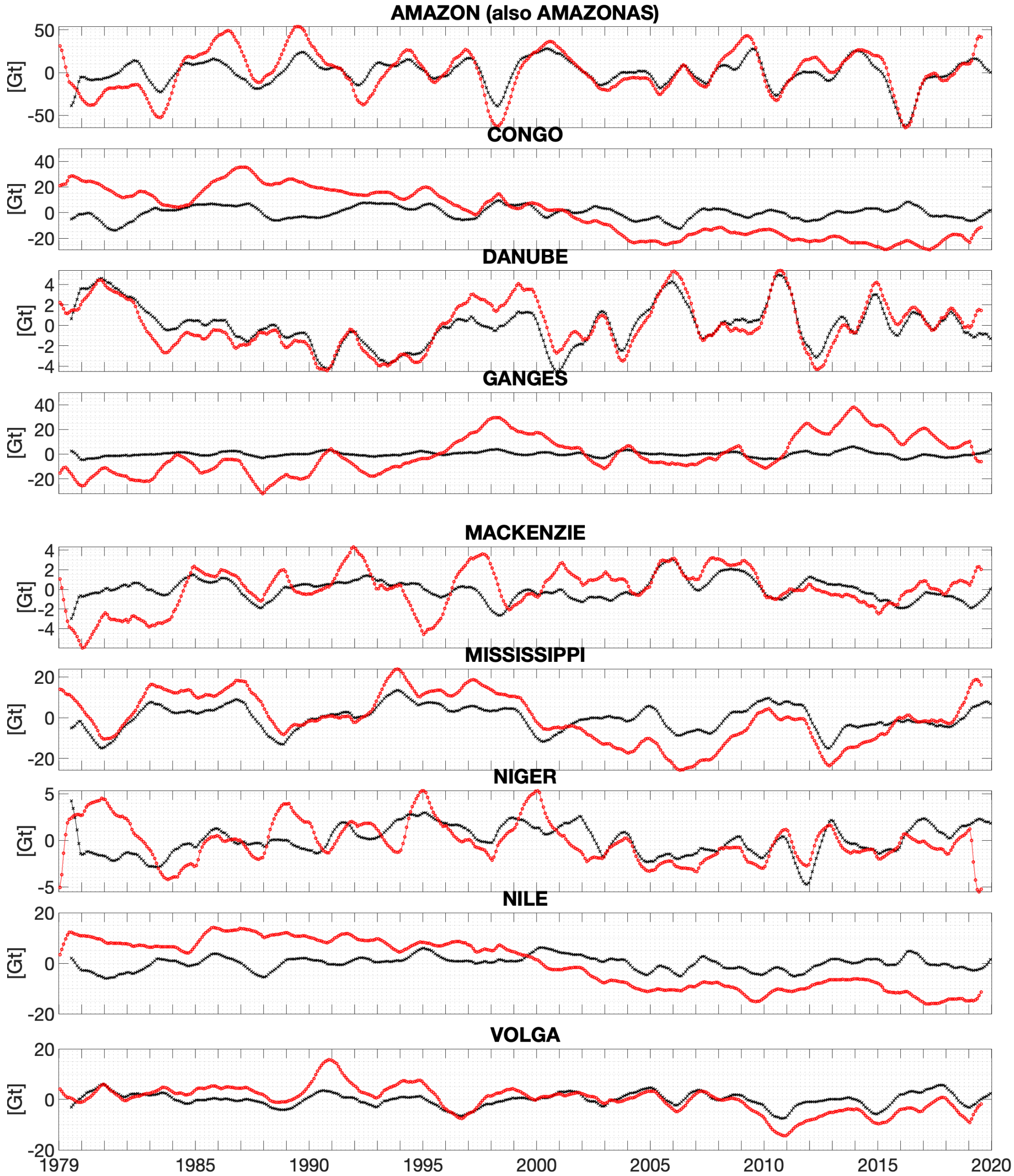

Figure Appendix D.1. Detrended time series for Li21 (black crosses) and HG19 (red circles) for the Amazon, Mississippi, Mackenzie, Volga, Danube, Ganges, Nile, Niger, Congo basin. The seasonal signal is reduced using a 12 month moving average filter.

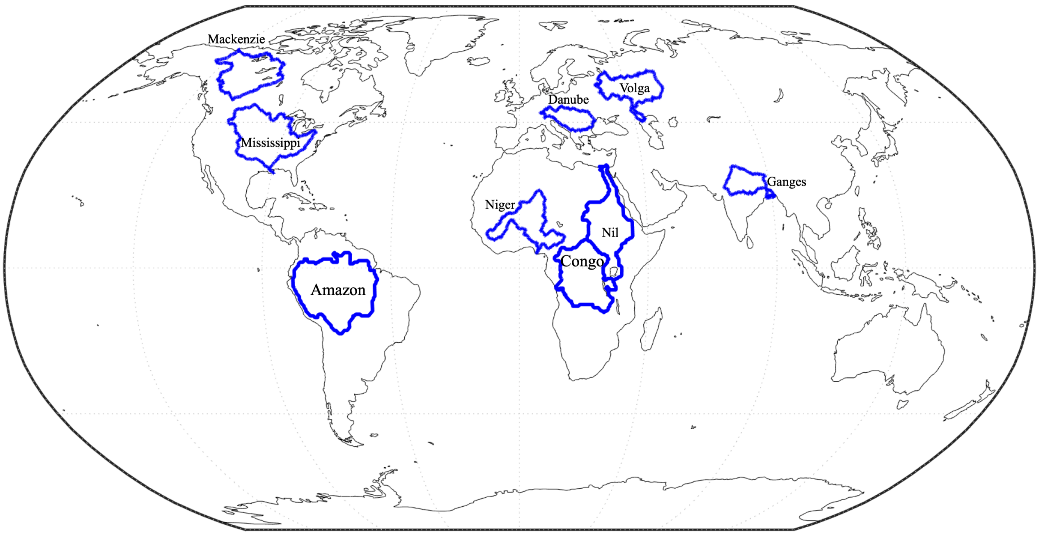

Figure Appendix D.2. Overview: Selected basins

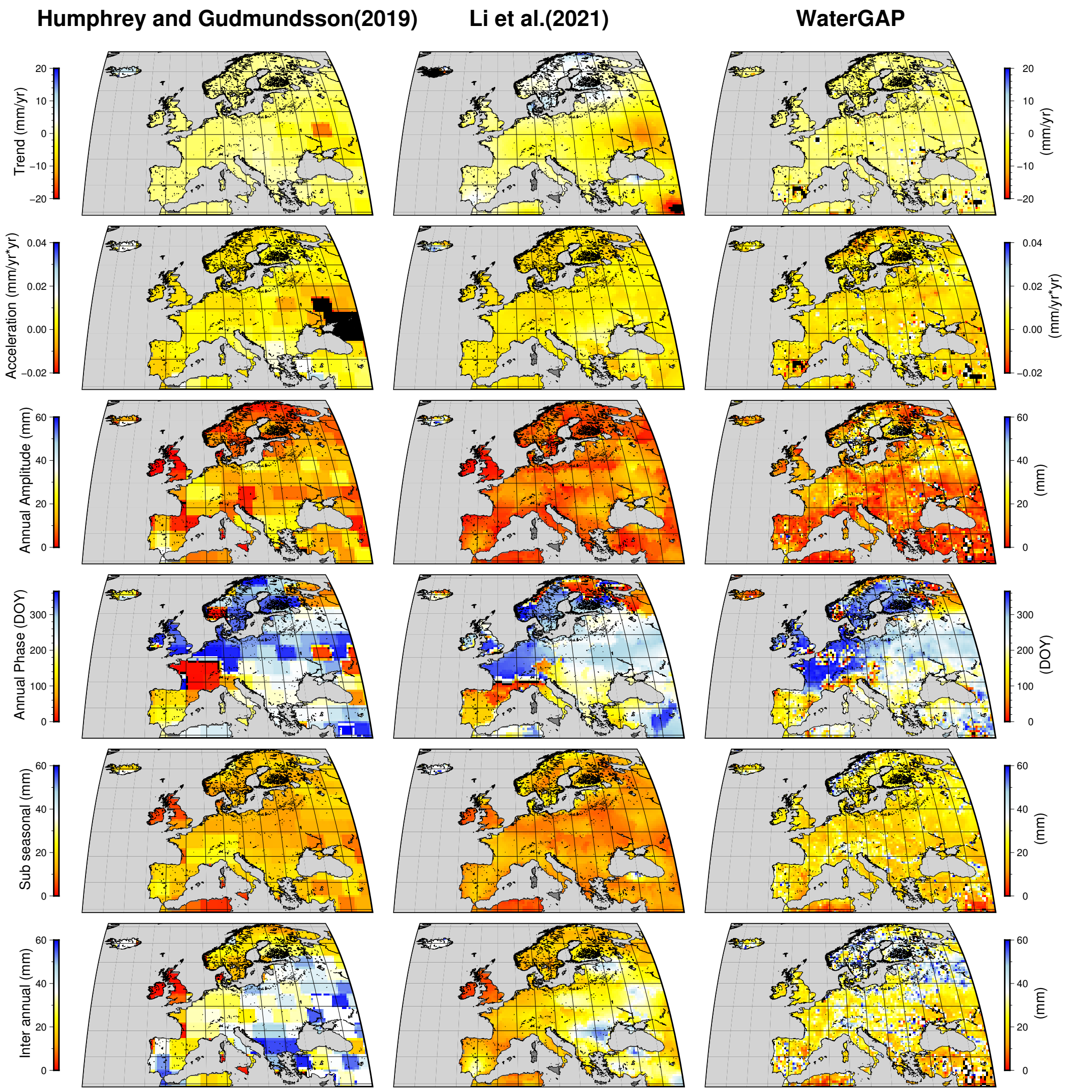

Figure Appendix E.1. Trend, acceleration, annual amplitude and phase, subseasonal and interannual signal for the years 1979-2016 for the reconstructions and WaterGAP for Europe

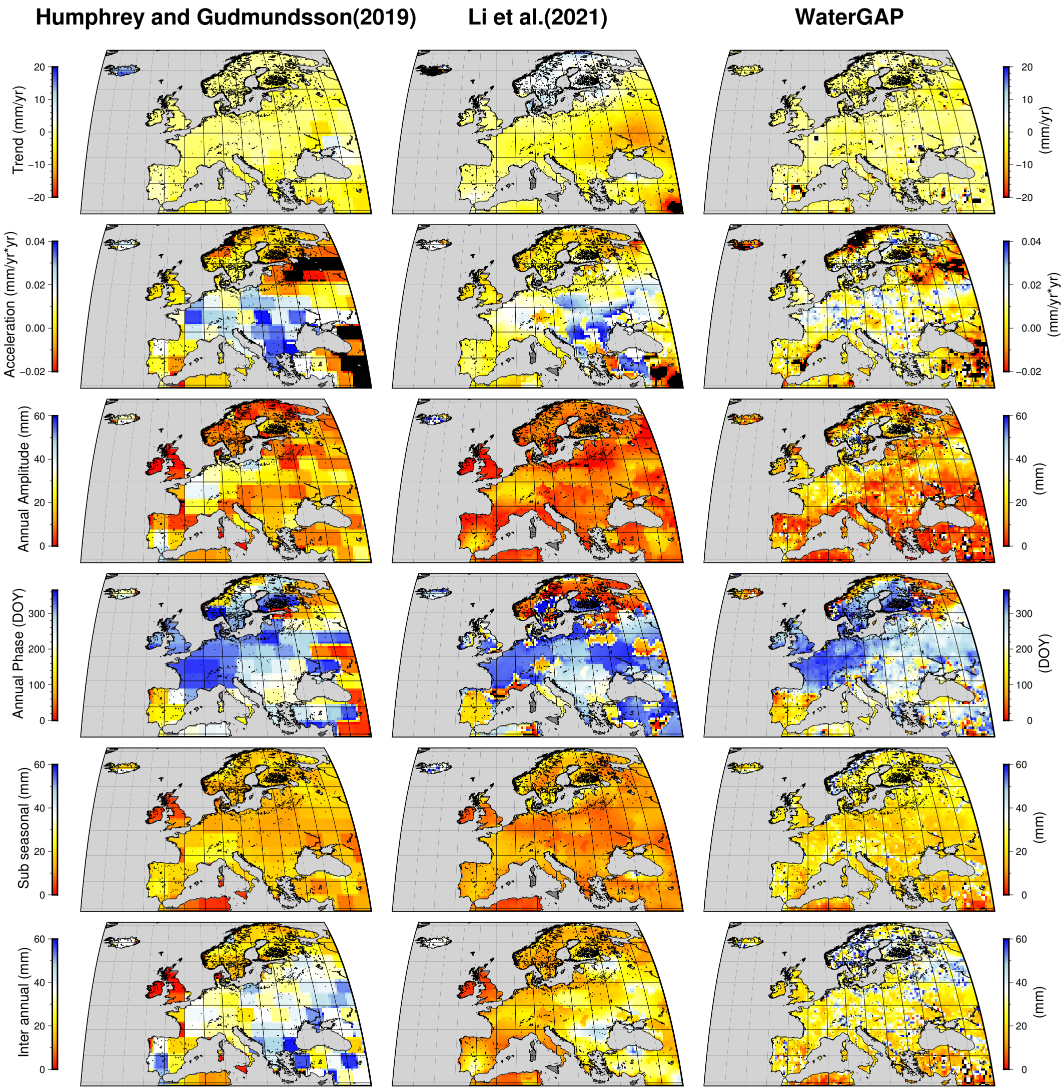

Figure Appendix E.2. Trend, acceleration, annual amplitude and phase, subseasonal and interannual signal for the years 1979-2002 for the reconstructions and WaterGAP for Europe

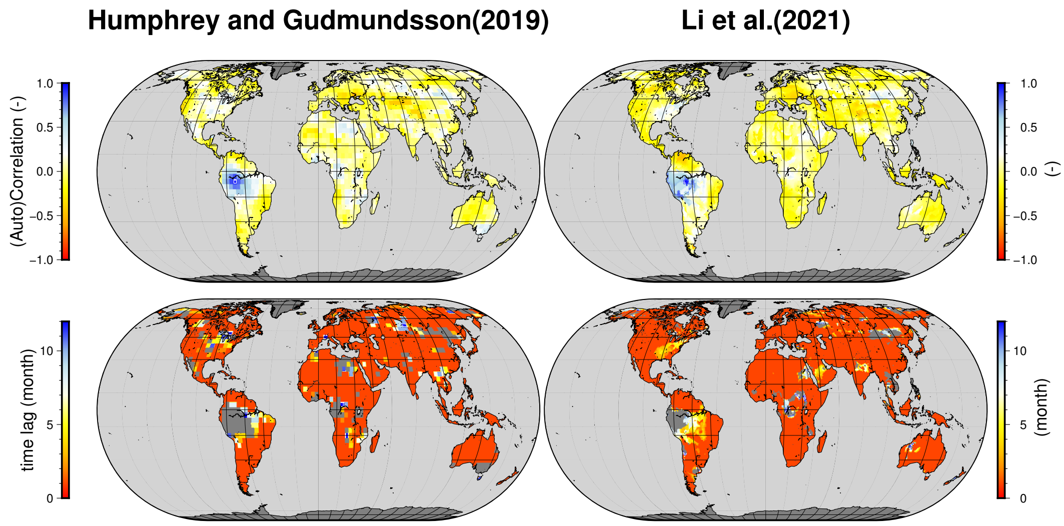

Figure Appendix E.3. *top:* Autocorrelation of the detrended and deseasonalized reconstructions. *bottom:* Time Lag. The point is located in the Amazon basin.

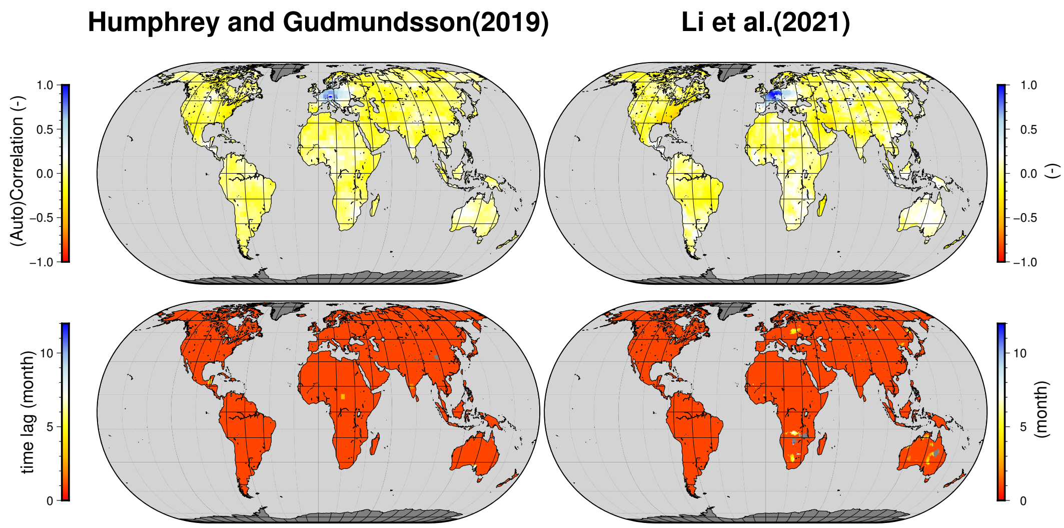

Figure Appendix E.4. *top:* Autocorrelation of the detrended and deseasonalized reconstructions. *bottom:* Time Lag. The point is located in Germany.