Zhang, L., Dobslaw, H., Thomas, M. (2016): Globally gridded terrestrial water storage variations from GRACE satellite gravimetry for hydrometeorological applications. - *Geophysical Journal International*, *206*, 1, pp. 368–378.

DOI: http://doi.org/10.1093/gji/ggw153

Geophysical Journal International

Geophys. J. Int. (2016) **206,** 368–378
Advance Access publication 2016 April 15
GJI Gravity, geodesy and tides

doi: 10.1093/gji/ggw153

Globally gridded terrestrial water storage variations from GRACE satellite gravimetry for hydrometeorological applications

Liangjing Zhang,[1] Henryk Dobslaw[1] and Maik Thomas[1,2]

[1]*GFZ German Research Centre for Geosciences, Department* 1*: Geodesy and Remote Sensing, Telegrafenberg,* 14473 *Potsdam, Germany, E-mail: jzhang@gfz-potsdam.de*
[2]*Freie Universität Berlin, Institute of Meteorology, Carl-Heinrich-Becker-Weg* 6-10, 12165 *Berlin, Germany*

Accepted 2016 April 13. Received 2016 April 12; in original form 2015 August 21

SUMMARY

Globally gridded estimates of monthly-mean anomalies of terrestrial water storage (TWS) are estimated from the most recent GRACE release 05a of GFZ Potsdam in order to provide non-geodetic users a convenient access to state-of-the-art GRACE monitoring data. We use an ensemble of five global land model simulations with different physics and different atmospheric forcing to obtain reliable gridded scaling factors required to correct for spatial leakage introduced during data processing. To allow for the application of this data-set for large-scale monitoring tasks, model validation efforts, and subsequently also data assimilation experiments, globally gridded estimates of TWS uncertainties that include (i) measurement, (ii) leakage and (iii) re-scaling errors are provided as well. The results are generally consistent with the gridded data provided by Tellus, but deviate in some basins which are largely affected by the uncertainties of the model information required for re-scaling, where the approach based on the median of a small ensemble of global land models introduced in this paper leads to more robust results.

Key words: Satellite geodesy; Time variable gravity; Loading of the Earth; Hydrology.

1 INTRODUCTION

Satellite observations of temporal variations of the Earth's gravity field from the Gravity Recovery and Climate Experiment (GRACE) satellite mission (Tapley *et al.* 2004) offer a new possibility to monitor a wide range of Earth system dynamics that are related to large-scale mass re-distribution. Over the continents, the mission is in particular sensitive to several aspects of the terrestrial branch of the global water cycle: it allows measurements of the mass balance of continental ice-sheets and glaciers (Jacob *et al.* 2012), deep soil moisture variability and its consequences for drought and flood potentials (Houborg *et al.* 2012), as well as groundwater depletion arising from growing water demands for irrigation and human consumption in agricultural regions world-wide (Voss *et al.* 2013).

The anomaly of terrestrial water storage (TWS) with respect to a long-term average is the most direct hydrological quantity obtainable from the GRACE science products. TWS is understood here as the sum of all storage compartments of water at and underneath the land surface. This includes soil moisture; the water content in snow-pack and land ice; ground water in shallow and deep aquifers; canopy water; and also the content of surface water bodies as rivers, lakes, and occasionally flooded wetlands. As a new observable that is available globally from space-based instruments, GRACE-based TWS serves an important role in assessing the closure of the terrestrial water balance at global and regional scales and allows for a new way to assess and even improve the quality of hydrological model simulations (Syed *et al.* 2008; Frappart *et al.* 2013; Eicker *et al.* 2014).

Due to its observing principle, GRACE data is highly accurate at hemispheric spatial scales but provides no information on TWS variability at spatial scales smaller than a few hundred km. GRACE monthly mean gravity fields and their associated uncertainties are typically provided in terms of global spherical harmonic coefficients – a mathematical representation that substantially hampers the application of GRACE mission results in non-geodetic branches of the physical geosciences. In order to overcome this limitation, we describe in this paper a globally gridded data-set of TWS anomalies at monthly resolution that is corrected for known systematic errors and other deficits by applying state-of-the-art post-processing methods. To allow for a subsequent application of the data-set in model validation efforts or even data assimilation experiments, globally gridded observational uncertainties are derived as well. Similar gridded TWS products are also available at the Tellus website (http://grace.jpl.nasa.gov/data/get-data/monthly-mass-grids-land/) which are calculated using a different processing scheme as described by Landerer & Swenson (2012). For simplicity, we use the term Tellus to refer to this data-set throughout the manuscript.

The structure of our paper is as follows: The post-processing procedure to derive globally gridded TWS anomalies out of the GRACE monthly mean gravity fields will be introduced in detail

in Section 2. Subsequently, we introduce our method to correct for the leakage and bias using median scaling factors estimated from a small ensemble of five hydrological model simulations in Section 3. In Section 4, we present our gridded error estimates, whose consistency with uncertainties directly obtained from the Stokes coefficients are demonstrated for the largest 50 basins. We then compare the basin-averaged TWS from our gridded TWS variations with those from Tellus for these basins in Section 5 to demonstrate the consistency and robustness of our method, and end with some concluding remarks on the general applicability of GRACE results for hydrometerological applications in the final Section 6.

2 GRACE DATA-SETS AND POST-PROCESSING

GRACE is a twin satellite mission of NASA and the German space agency DLR that was launched in March 2002 into a polar orbit at an initial altitude of only 450 km. By means of highly accurate range-rate measurements between the two spacecrafts that follow each other with a typical separation distance of 250 km, the mission is able to map the Earth's gravity field at approximately monthly intervals on spatial scales of a few hundred km and larger. After removing short-term variability due to tides in solid Earth (Petit & Luzum 2010), oceans (Savcenko & Bosch 2012) and atmosphere (Biancale & Bode 2006), as well as non-tidal variability in atmosphere and oceans (Dobslaw *et al.* 2013) from the observations, the remaining gravity changes on monthly to inter-annual time scales mainly represent variations in TWS. Since the mission inherently does not provide vertical resolution but is instead equally sensitive to mass variability at or beneath the surface, an unambiguous separation of individual storage compartments by means of GRACE data only is impossible.

We use the monthly GRACE release 05a Level-2 products expanded in spherical harmonics up to degree and order 90—corresponding to a spatial wavelength of 220 km—from GFZ Potsdam (Dahle *et al.* 2012), which show an overall improvement by a factor of two in terms of noise reduction compared to previous releases (Chambers & Bonin 2012). The data can be visually explored and downloaded from the website of the International Centre for Global Earth Models (ICGEM) available at 'icgem.gfz-potsdam.de/ICGEM'. In addition to this routinely updated standard GRACE solution, we further utilize in this study a recent GRACE release from the University of Graz (ITSG-Grace2014; Mayer-Gürr *et al.* 2014), which additionally provides full variance-covariance matrices consistent with its unconstrained monthly gravity field solutions.

For the derivation of globally gridded TWS variations from GRACE gravity fields we essentially follow the strategy chosen by Bergmann & Dobslaw (2012), which is briefly recalled here. First, we add global degree-1 coefficients as derived by Bergmann-Wolf *et al.* (2014) with the methodology of Swenson *et al.* (2008), replace the C_{20}-coefficients with estimates from Satellite Laser Ranging (Cheng *et al.* 2011), and remove a multi-year average over the period of January 2004 to December 2009 which is consistent with Tellus to arrive at anomalies. Next, we apply the approximate de-correlation and non-isotropic smoothing method introduced by Kusche (2007) to remove correlated errors in north-south directions that are related to the anisotropic sensitivity of the track-aligned range-rate measurement system between the two satellites. Smoothing in space is enabled through the tuning parameter $a = 10^{13}$ of the signal covariance matrix, which is known also as DDK2 filter and approximately corresponds to an isotropic Gaussian filter with 680 km full width half maximum (Kusche *et al.* 2009). The filtered spherical harmonic coefficients are finally synthesized on a 1° latitude-longitude grid following the conventions of Wahr *et al.* (1998). Since assuming that mass redistribution occurs at the surface of a sphere may introduce errors when spherical harmonic coefficients are expanded to higher degrees, we use a reference ellipsoid as defined by the IERS conventions which is a more suitable approximation of the Earth's shape. Note that using instead a spherical surface might only impose differences of up to 0.5 cm equivalent water height (e.w.h.) in particular at the higher latitudes.

3 COMPENSATION OF FILTER-INDUCED SIGNAL ATTENUATION

The gravitational field of the Earth is conservative in space and might be thus continued both upward to the satellite orbit and downward to the Earth surface without loss of generality. The signal decay with height depends on the spatial extent of a gravity disturbance, so that large-scale anomalies generally cause stronger deviations of a spacecraft trajectory from its reference orbit. Errors of GRACE-based TWS variations therefore grow with increasing degree of the spherical harmonics expansion, and low-pass filtering in space or spectral domain is consequently required to reduce the contribution from highly uncertain smaller spatial scales when gridded estimates are to be calculated.

The process of truncation and filtering, however, typically diminishes amplitudes of highly localized signals, and increases spatial leakage of signals from neighbouring regions (Werth *et al.* 2009). In order to approximately account for such effects in the post-processing, Klees *et al.* (2007) suggested to apply local re-scaling coefficients derived from TWS predictions from a global numerical model. The scaling factors estimated from a numerical model are, however, strongly related to the characteristics of the simulated water distribution in a larger area around the point-of-interest. Thus, uncertainties of the numerical model in terms of the model structure, parameter values and meteorological forcing will affect the estimated scaling factors as well. Recent inter-comparison studies of hydrological models demonstrated that there can be a large spread in model performance (see, e.g. Gudmundsson *et al.* 2012) for various regions and frequencies. Thus, the sensitivity of the re-scaling factors to different models is investigated in this paper more closely by using a small ensemble of five different numerical model experiments, which includes land surface schemes of both low and intermediate complexity; global land assimilation systems; and conceptual hydrological models.

(1) The Land Surface Discharge Model (LSDM; Dill 2008) is based on a simplified land surface scheme and a hydrological discharge model developed at the Max-Planck-Institute for Meteorology (MPI-M) in Hamburg, Germany (Hagemann & Gates 2001, 2003). The simulation used in this study is forced with operational ECMWF analysis data, which is also the basis for predictions of hydrologically induced vertical crustal deformations for geodetic applications (Dill & Dobslaw 2013) publicly available at 'http://www.gfz-potsdam.de/esmdata'.

(2) The Global Land Data Assimilation System (GLDAS) operated at the National Center for Environmental Prediction (Rodell *et al.* 2004) incorporates both ground and space-based observations into the model results. We analyse a single realization of GLDAS

Table 1. Description of different model runs.

Model name	Meteorological forcing variables	Storage compartments included	Others
LSDM	ECMWF	Canopy, soil moisture, snow, surface water	
WGHM	WFDEI_CRU	Canopy, soil moisture, snow, surface water, groundwater	Tuned against runoff
JSBACH	WFDEI_CRU	Canopy, soil moisture, snow, ground water	
MPI-HM	WFDEI_CRU	Canopy, soil moisture, snow, surface water	
GLDAS	GDAS meteorological data, CMAP precipitation	Canopy, soil moisture, snow	Observation data assimilated

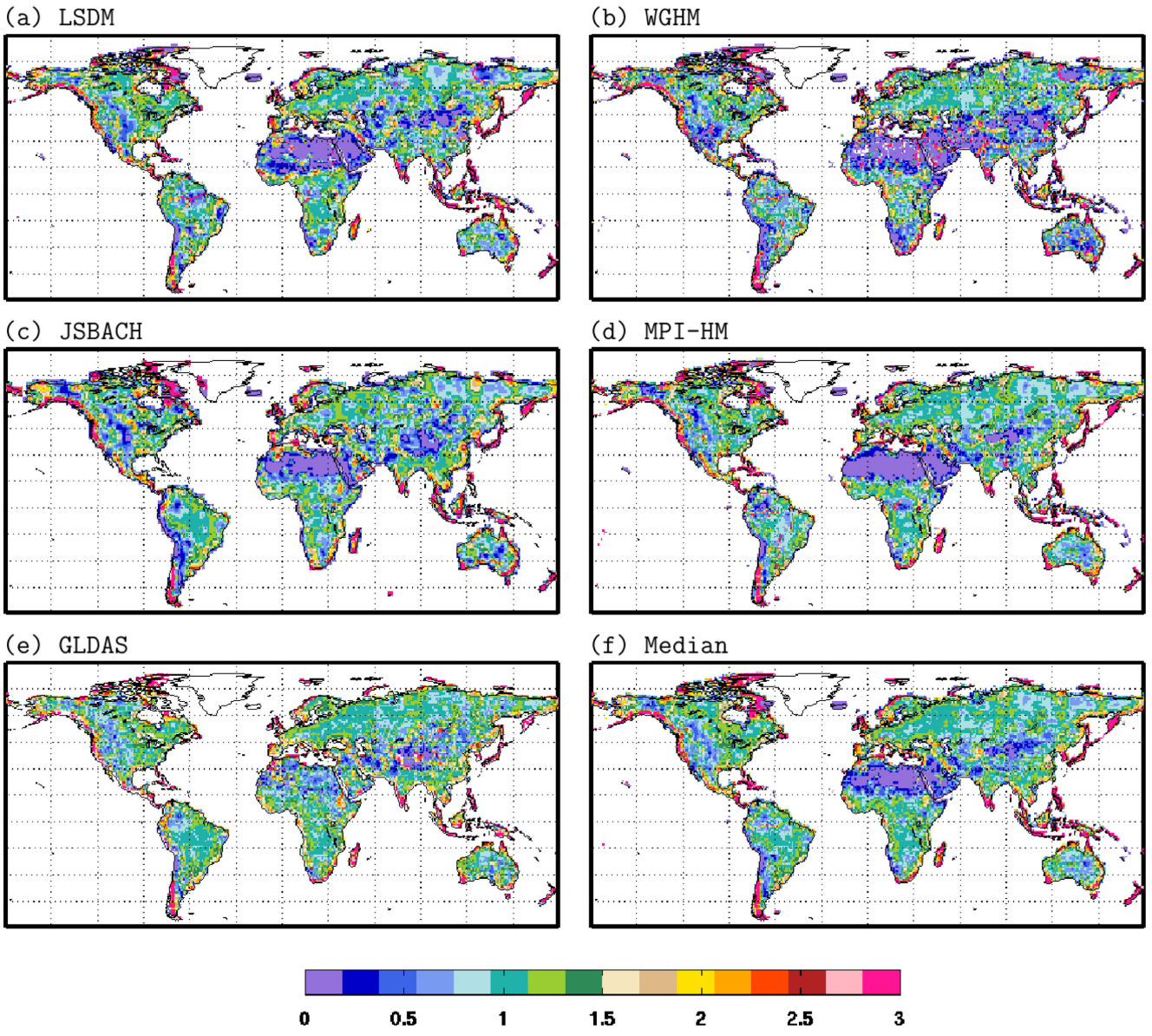

Figure 1. Re-scaling factors obtained as the average ratio of DDK2-filtered and unfiltered TWS time-series from five global land model experiments performed with LSDM (a), WGHM (b), JSBACH (c), MPI-HM (d) and GLDAS (e). In order to reduce the impact of individual model deficiencies, the median value (f) of all five simulations is applied in the GRACE post-processing.

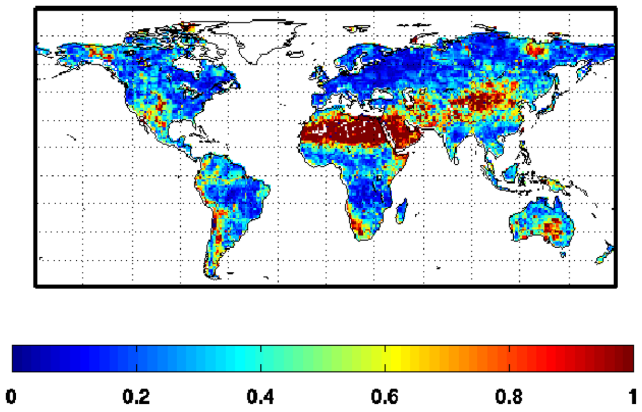

Figure 2. Variation coefficients for the re-scaling factors as obtained as the RMS between individual model-derived re-scaling factors and the median value normalized by the median.

based on the Noah community land surface model downloaded from 'mirador.gsfc.nasa.gov'.

(3) As a compartment of the Water Global Assessment and Prognosis model (WaterGAP; Döll *et al.* 2003), the WaterGAP Global Hydrological Model (WGHM) has been widely applied in several GRACE-related studies (Güntner *et al.* 2007). The model realization available to us is based on version 2.2 as described by Müller Schmied *et al.* (2014) and has been forced with atmospheric data specifically prepared for the Water and Global Change (WATCH) project, which is based on ERA-Interim re-analysis data (Dee *et al.* 2011) and bias corrections for precipitation from the Climate Research Unit (CRU) station-based monthly climatologies (WFDEI_CRU; Weedon *et al.* 2011).

(4) JSBACH (Raddatz *et al.* 2007; Brovkin *et al.* 2009) is a land surface model that forms together with ECHAM6 (Stevens *et al.* 2013) and MPIOM (Jungclaus *et al.* 2013) the current Max-Planck-Institute for Meteorology's Earth System Model (MPI-ESM). We use here an experiment with an un-coupled version of JSBACH that is also driven by daily the WFDEI_CRU atmospheric data.

(5) The final model experiment available is from the Max-Planck-Institute of Meteorology's Hydrology Model (MPI-HM; Stacke & Hagemann 2012), which is a global hydrological model with several more sophisticated parametrizations, as, for example, the Penman Montheith for potential evapotranspiration. Its water flux computations are of similar complexity to land surface models, but it does not account, however, for any energy fluxes. We have access to a model run that is once more integrated with WFDEI_CRU atmospheric forcing data.

Besides the apparent differences in model structure, parametrization, and atmospheric forcing, the model experiments considered here also differ with each other by what particular water storage compartments are actually included in TWS (Table 1). The results from all the model runs are aggregated into monthly averages and – where necessary – conservatively interpolated onto a regular latitude-longitude grid with a horizontal spacing of 1°. Since GRACE estimates over Greenland and Antarctica are dominated by surface mass balance and ice dynamics of the continental ice-sheets, both regions have been masked out of the model results together with the global oceans and will not be considered further in the reminder of this paper.

All model experiments and GRACE data are available to us over the period 2003–2012. Since local trends are usually poorly predicted by the models considered here, and since the influence of GIA signals cannot be completely removed due to a lack of knowledge on both glacial ice load history and mantle viscosity (Steffen *et al.* 2008), we concentrate on TWS variability on seasonal to inter-annual time-scales only. Linear trends are therefore estimated and subtracted from all data-sets considered.

Local re-scaling factors that compensate for filter-induced signal alteration have been obtained from all five model runs by calculating the average ratio between monthly TWS before and after filtering with DDK2 (Fig. 1). Re-scaling factors are almost zero for arid regions dominated by leakage-in from neighbouring regions with stronger signal variability. But factors are also as large as three for (i) isolated high-variability regions close to the coast, which are affected by leakage-out due to the much weaker ocean bottom pressure variability near-by; at (ii) step gradients in orography that are responsible for locally intensified precipitation; or (iii) along spatially concentrated surface water bodies with high storage variability. By comparing re-scaling factors for the different models, we note moreover substantial discrepancies among the model runs, in particular at smaller spatial scales. There is, however, no prior knowledge of which model performs best at different areas of the

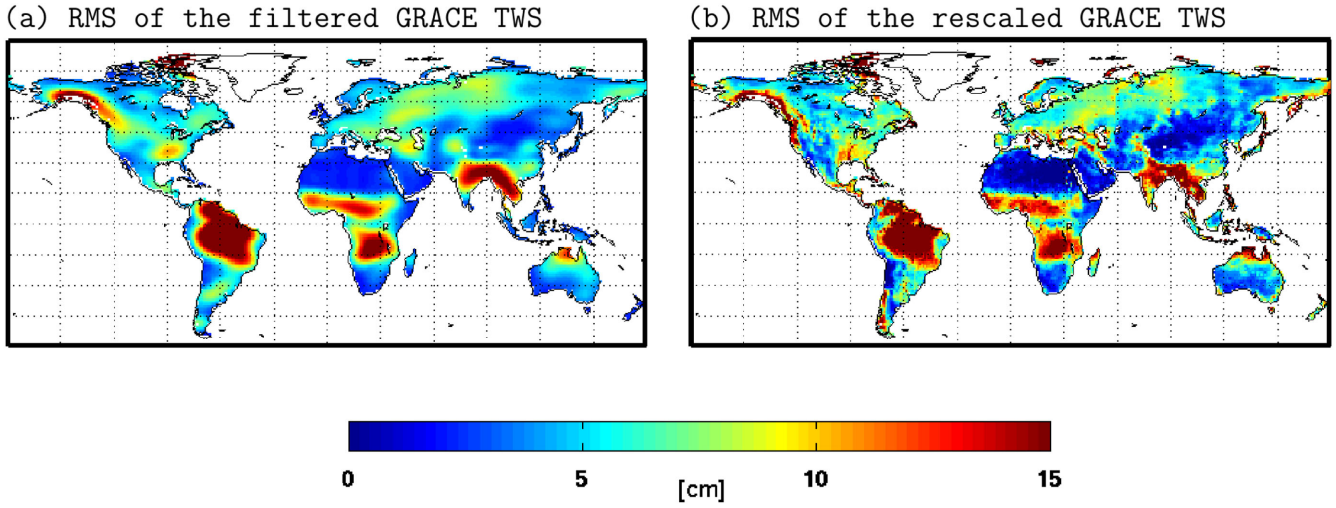

Figure 3. RMS variability of GRACE-based terrestrial water storage after applying the DDK2 filter (a), and after re-scaling of the filtered results (b) with spatially variable median re-scaling factors obtained from five different model data sets.

Table 2. Re-scaling factors estimated from the WGHM, LSDM, GLDAS, JSBACH and MPI-HM models for the 50 largest discharge basins of the world. Additionally provided are the median re-scaling factors subsequently applied in the GRACE processing; the variation (VAR) coefficients of the re-scaling factors quantifying the uncertainty associated with the re-scaling; total error and signal amplitude of the basin-averaged TWS (unit: cm) and the GRACE signal-to-noise ratio that includes also the impact of uncertain re-scaling; the RMS differences between rescaled TWS from Tellus and our calculation with median scaling factors. Basin names are taken from the Simulated Topological Network data set (STN-30p; Vörösmarty *et al.* 2000). GHAAS#Number indicate unidentified basins, where the number is the unique integer identifier adopted in STN-30p.

Discharge Basin	Re-scaling coefficients						VAR-coeff.	Total error	Signal amp.	GRACE-SNR	Tellus-Median scaled
	LSDM	WGHM	JSBACH	MPI-HM	GLDAS	Median					
Amazon	1.03	1.01	1.02	0.96	1.01	1.01	0.02	1.46	14.28	9.76	1.81
Nile	1.08	1.16	0.97	1.05	0.91	1.05	0.08	1.06	3.46	3.26	1.33
Zaire	1.08	0.96	1.04	1.01	0.99	1.01	0.04	1.32	5.05	3.82	0.98
Mississippi	1.00	1.00	0.97	0.95	0.97	0.97	0.02	0.86	5.67	6.60	0.59
Amur	1.01	0.99	0.96	0.96	0.94	0.96	0.03	0.68	2.17	3.18	0.67
Parana	1.14	1.10	1.06	1.08	1.01	1.08	0.04	1.32	5.92	4.50	1.51
Yenisei	0.96	0.92	1.03	0.98	0.97	0.97	0.03	0.68	4.52	6.67	0.64
Ob	0.94	0.98	1.02	0.96	0.99	0.98	0.03	0.68	5.64	8.31	0.66
Lena	0.94	0.88	1.00	0.98	0.96	0.96	0.04	0.68	4.12	6.01	0.60
Niger	1.04	1.01	1.03	1.05	0.99	1.03	0.02	1.29	6.37	4.93	0.54
Zambezi	1.02	0.98	1.02	1.00	0.97	1.00	0.02	1.57	10.66	6.80	0.91
GHAAS #14	1.42	1.74	1.77	0.70	0.99	1.42	0.30	0.77	0.68	0.88	0.47
Chang Jiang	1.37	0.98	0.89	0.98	0.89	0.98	0.19	1.49	4.61	3.09	1.26
Mackenzie	0.69	1.02	1.00	0.97	0.93	0.97	0.13	0.83	5.15	6.20	0.92
Ganges	0.96	1.09	1.00	1.00	0.97	1.00	0.05	1.94	11.62	5.99	0.81
Chari	1.19	1.08	1.01	1.02	0.97	1.02	0.08	1.50	5.14	3.42	0.92
Volga	1.00	0.99	0.98	0.98	0.98	0.98	0.01	0.84	7.07	8.43	0.68
St. Lawrence	1.02	0.98	0.85	0.75	0.79	0.85	0.13	1.14	5.63	4.94	1.14
Indus	1.17	0.87	1.00	0.95	0.89	0.95	0.12	1.54	3.72	2.42	1.32
Syr-Darya	1.20	0.92	1.01	0.99	1.00	1.00	0.10	1.12	4.08	3.65	0.84
Nelson	0.93	1.01	0.94	0.87	0.94	0.94	0.05	1.12	4.28	3.82	1.26
Orinoco	0.99	1.09	1.07	1.08	0.93	1.07	0.07	3.14	14.88	4.74	1.97
Murray	1.31	1.04	1.07	1.13	0.91	1.07	0.12	1.88	5.13	2.73	1.38
Great Artesian	1.76	1.17	0.97	0.97	1.29	1.17	0.25	1.33	3.54	2.67	1.10
Shatt el Arab	1.19	0.90	1.15	1.01	0.88	1.01	0.12	1.49	5.67	3.81	1.40
Orange	1.00	1.02	1.20	1.08	0.91	1.02	0.09	1.65	2.75	1.67	1.29
Huang He	1.28	0.52	0.99	1.05	1.02	1.02	0.25	1.28	3.00	2.35	0.91
Yukon	1.08	0.83	0.90	0.88	1.12	0.90	0.15	1.19	9.13	7.68	1.24
GHAAS #34	1.14	1.06	1.40	0.84	0.90	1.06	0.19	1.04	1.13	1.09	0.67
Colorado (Ari)	1.39	0.82	1.04	1.11	0.97	1.04	0.18	1.41	3.93	2.78	0.99
Danube	1.01	1.04	0.96	1.00	0.95	1.00	0.03	1.50	7.45	4.96	0.86
Mekong	1.11	1.07	1.02	1.08	0.90	1.07	0.08	3.86	14.42	3.73	2.32
Tocantins	1.06	1.09	1.04	1.21	1.00	1.06	0.07	2.81	16.72	5.95	2.69
Columbia	0.98	0.99	0.98	1.07	0.92	0.98	0.05	1.85	9.87	5.32	1.76
GHAAS #49	0.80	0.97	0.94	0.84	0.92	0.92	0.08	1.86	3.06	1.65	0.97
Kolyma	0.99	0.78	1.03	0.92	1.04	0.99	0.11	0.98	4.55	4.65	1.27
Sao Francisco	1.06	1.04	0.99	1.09	1.04	1.04	0.03	2.81	9.65	3.43	1.80
Amu-Darya	1.32	1.27	1.11	1.20	0.95	1.20	0.13	2.47	7.27	2.95	2.34
Dnepr	0.99	0.98	0.97	0.99	0.98	0.98	0.01	1.32	6.37	4.81	1.09
Don	1.14	1.07	1.12	1.12	1.05	1.12	0.03	1.68	8.67	5.17	1.16
GHAAS #50	1.01	0.85	1.72	0.47	0.81	0.85	0.48	1.20	0.89	0.74	0.70
Zhu jiang	1.07	1.18	1.19	1.37	1.11	1.18	0.10	3.18	7.92	2.49	2.39
Irrawaddy	1.32	1.43	1.40	1.06	1.26	1.32	0.11	4.18	17.44	4.17	10.07
Volta	0.63	0.88	0.98	1.06	0.97	0.97	0.18	3.36	9.94	2.96	1.83
GHAAS #54	0.72	0.74	1.28	0.83	1.10	0.83	0.27	1.66	3.25	1.96	1.03
Khatanga	0.87	0.94	0.87	0.89	0.93	0.89	0.04	1.01	5.48	5.42	1.43
Dvina	1.04	1.03	1.10	1.01	1.00	1.03	0.04	1.25	7.58	6.05	1.02
Urugay	1.21	1.12	1.30	1.21	1.16	1.21	0.06	2.70	7.28	2.70	3.00
Qarqan	0.49	1.07	0.23	1.04	1.33	1.04	0.54	1.14	1.34	1.18	1.63
GHAAS #75	0.26	1.02	0.77	0.51	0.83	0.77	0.44	0.94	0.75	0.80	0.48

world, so that we calculate a median value of the re-scaling coefficients from the five model runs at each grid point, which is finally applied to the GRACE data.

To further quantify the uncertainties in the re-scaling factors, we calculate root mean square (RMS) estimates of the difference between each re-scaling factor and the median value normalized by the median itself (Fig. 2). Those variation coefficients show similarities with the results from Long *et al.* (2015), which is quite opposite with the pattern of the signal variability of GRACE-based TWS (Fig. 3). Largest variability of the re-scaling factors occurs in

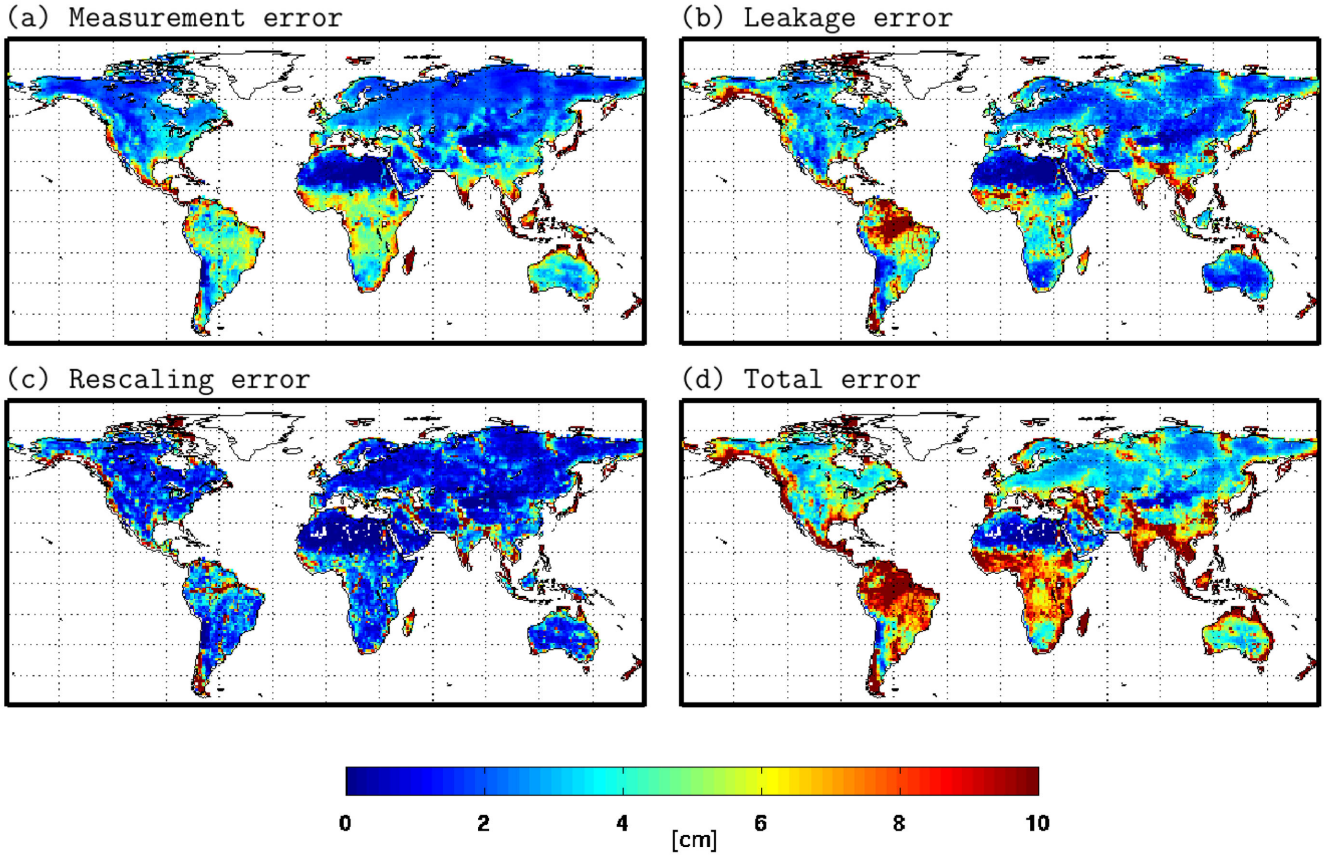

Figure 4. Estimates of GRACE-based TWS errors calculated at 1° grid-scale level: measurement errors (a), leakage errors (b), re-scaling errors (c) and total errors (d).

rather dry areas, like North Africa, South Australia, Middle East, and Northwest China, which is partly explained by the small values of the median scaling factors in these regions. Areas dominated by surface water variability and affected by groundwater abstraction also show large spreads of the re-scaling factors.

We also look into the re-scaling coefficients estimated from different models at basin-scale level (Table 2). Despite of the differences indicated above, variation coefficients of the re-scaling factors are mostly below 0.15 when averaged over areas of several 10^6 km^2, thereby indicating the high consistency of the TWS simulated by all models in these large areas. Uncertainties grow when smaller basins with a lower GRACE signal-to-noise ratio (SNR) are considered. The SNR is calculated as the ratio of the RMS of the GRACE TWS time-series and the total error estimated from the basin-scale method which is introduced in Section 4. Both the low GRACE SNR and the large spread of the scaling factors from hydrological models confirm the poor ability of GRACE and hydrological models to capture the TWS signal in overly dry areas (Gunkel & Lange 2011). Care should also be taken in some humid areas where a large spread of the re-scaling factors is found, as, for example, in the Yukon basin, where a substantial contribution of mountain glaciers on observed TWS variability can be expected which is rather poorly represented in all the models. Considering that we only focus on the seasonal and interannual signals, it is therefore not suitable to apply such scaling factors in areas which have large contribution from glaciers. Further, observed TWS in the Chang Jiang basin is affected by surface water variability, which is not represented properly in both GLDAS and JSBACH. In the Indus catchment, intensive irrigation takes place which uses both surface water and groundwater resources. However, only WGHM and JSBACH accommodate groundwater storage changes in their physical models. Nevertheless, the application of the median value makes the re-scaling factors less affected by such deficiencies in a single model and therefore contributes to the robustness of the GRACE post-processing methodology applied here.

4 TWS ERRORS FROM GRACE

In order to provide a quantitative estimate for the uncertainties associated with GRACE-based TWS errors, we individually assess the contributions of measurement errors, leakage errors, and re-scaling errors as suggested by Landerer & Swenson (2012). We estimate the measurement error by error propagation from the 'calibrated errors' provided together with the GRACE monthly-mean Stokes coefficients. The calibrated errors are further re-scaled to fit the non-seasonal GRACE residuals after subtracting a constant as well as annual and semiannual sinusoids following Wahr *et al.* (2006). Then gridded re-scaling factors are multiplied with the measurement errors to get the final measurement error distribution in the spatial domain. The leakage error compartment is calculated for all five model realizations according to

$$E_{\text{leak}} = \text{RMS}(\Delta S_o - k\Delta S_F)\frac{\text{RMS}_{\text{GRACE}}}{\text{RMS}_{\text{Model}}}, \quad (1)$$

where ΔS_o and ΔS_F are the original and filtered signals from the models, respectively, k is the re-scaling factor, and $\text{RMS}_{\text{GRACE}}$ and $\text{RMS}_{\text{Model}}$ are the RMS of the TWS from filtered GRACE and from one of the unfiltered model data-sets. For the re-scaling error, we

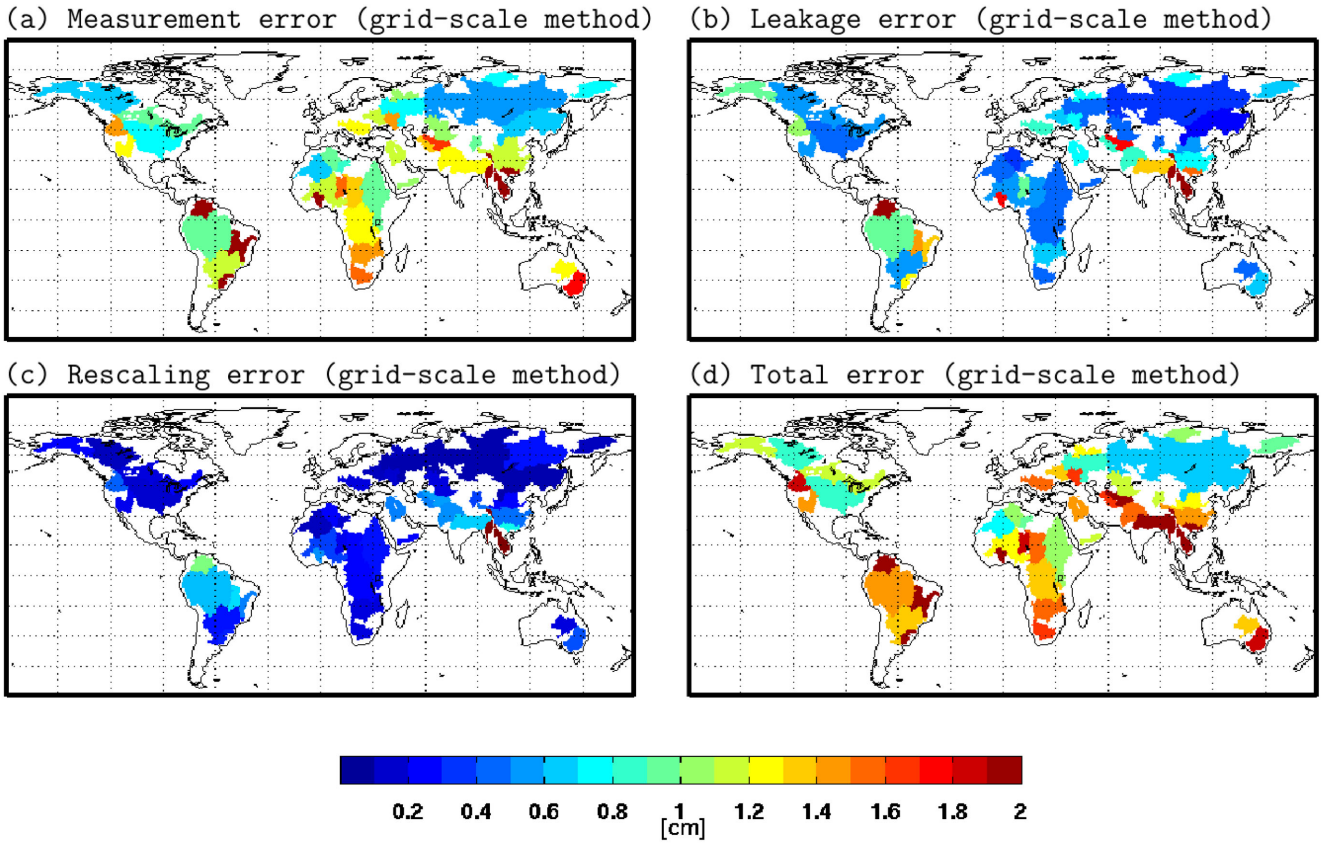

Figure 5. Estimates of GRACE-based TWS errors for the 50 largest discharge basins derived from 1° grid point estimates: measurement errors (a), leakage errors (b), re-scaling errors (c) and total errors (d).

make use of the scaling ratio for the individual model run shown in Fig. 1 by multiplying the RMS of the GRACE signals with the difference between each realization-based re-scaling factor with the median value. The total error at each grid point is subsequently taken as the sum of the measurement error, leakage error, and re-scaling error in quadrature.

While calculating these error compartments individually for each grid point, we obtain total errors of up to 10 cm in e.w.h. (Fig. 4). The water storage variation estimated from GRACE is not a point measurement, but rather a regional spatial average. It therefore does not make sense to compare GRACE at 1° grid-scale level directly to any other data-set. We use, instead, the gridded error estimates as a starting point for deriving error estimates of arbitrarily shaped regional averages. Here, the shapes of the 50 largest river basins from Table 2 are chosen, but the methodology can be similarly applied to other areas as well, as for example, climate or altitude zones associated with particular precipitation regimes, the spatial extent of a specific aquifer system, or an area particularly affected by land use and land cover changes. Since the error contributors of GRACE-based TWS at the grid scale are spatially correlated, the basin-averaged water storage errors cannot be obtained by simply averaging the gridded errors for an arbitrarily shaped region. We use the squared exponential covariance function to estimate the statistical covariance between two grids as proposed by Landerer & Swenson (2012) and estimate the error variance of the regional mean TWS estimate by the following equation:

$$\mathrm{var} = \sum_{i=1}^{m}\sum_{j=1}^{m} w_i w_j \theta_i \theta_j \exp\left(\frac{-d_{ij}^2}{2d_0^2}\right), \qquad (2)$$

where θ_i is the error for grid point i, w_i represents the area weight at the grid i, d_{ij} is the distance between the two points and d_0 is the parameter in the Gaussian window representing the de-correlation length scale. We choose 300, 100 and 10 km as d_0 for measurement error, leakage error and re-scaling error (Fig. 5) separately by fitting the error budget from the gridded data set to the ones obtained from estimating errors directly at basin-scale level (Chen *et al.* 2007; Klees *et al.* 2007) as shown in Fig. 6. Both methods provide generally consistent results down to a level of about 20 per cent , indicating that the errors at 1° spatial resolution might indeed serve as a starting point for the derivation of realistic errors for regions of arbitrary shape.

Since only diagonal elements of the covariance matrix have been provided for GRACE release 05a of GFZ Potsdam, the error correlations between the individual Stokes coefficients are typically ignored. To assess the impact of neglecting the error correlations, error estimates from both the diagonal and full variance-covariance matrices of the ITSG-Grace2014 release (Mayer-Gürr *et al.* 2014) are calculated (Fig. 7). The ITSG-Grace2014 measurement errors propagated from only the diagonal part show consistent results with GFZ RL05a. The differences between the measurement errors from the diagonal and full covariance matrix reach 0.56 cm at some basins (Fig. 7). Generally, when the error correlations are neglected, the measurement errors at the lower latitudes are overestimated, while at higher latitudes they are underestimated. GFZ is planning to provide the error variance-covariance matrix as well (Ch. Dahle, personal communication, 2015), which we believe is necessary to further improve the reliability of the TWS error estimates.

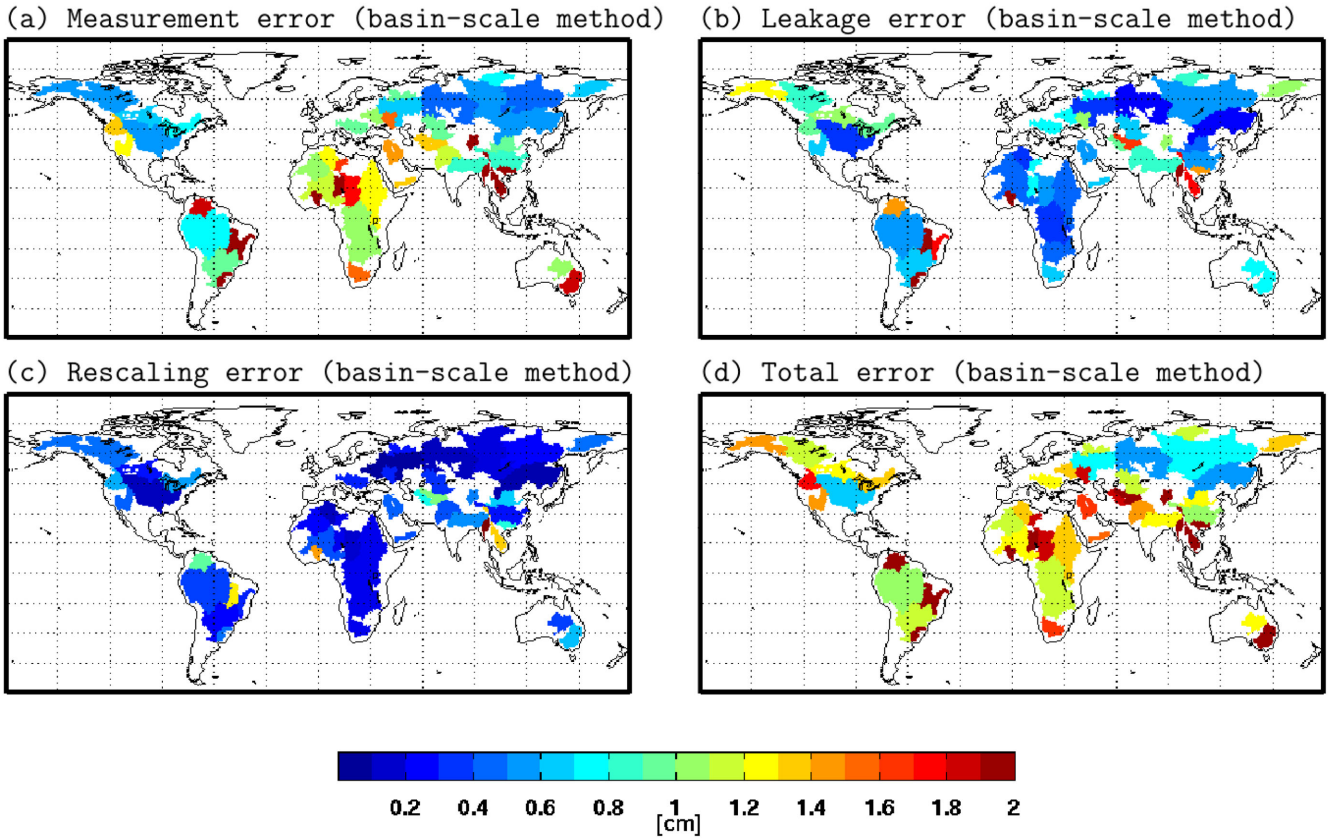

Figure 6. Estimates of GRACE-based TWS errors for the 50 largest discharge basins directly calculated at the basin-scale level out of spherical harmonic coefficients: measurement errors (a), leakage errors (b), re-scaling errors (c) and total errors (d).

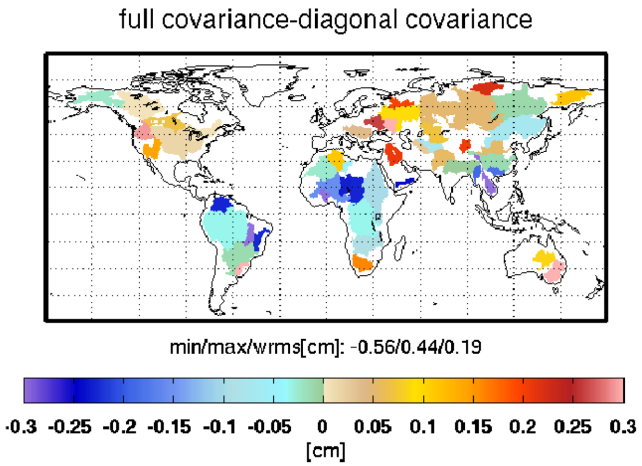

Figure 7. Changes in measurement errors for ITSG-Grace2014 when covariances are taken into account in addition to the usually considered variances.

5 COMPARISON WITH GRIDDED TWS RESULTS FROM TELLUS

We now compare the basin-averaged TWS time-series from our calculation with gridded TWS products downloaded from the Tellus website, which are based on GRACE Stokes coefficients truncated at degree and order 60, destriped and smoothed by using a 300 km Gaussian filter following Swenson & Wahr (2006) and then re-scaled by scaling factors derived from NCAR's CLM4.0 land surface model (Landerer & Swenson 2012). The RMS of the differences between the TWS variations filtered with two different methods generally lie within the bounds of the GRACE error estimates (*cf.* Fig. 5d), which indicates consistency between the different filtering and smoothing methods applied (Fig. 8a).

By comparing the two re-scaled TWS time-series, we find several basins with much larger differences in particular in South America and Southeast Asia and there are ten out of fifty basins where the RMS of the differences are larger than the GRACE error estimates (Fig. 8b; Table 2). The relative differences as the percentage of the RMS of the basin-averaged TWS from Tellus are shown as well (Figs 8c and d). The large relative differences mainly occur at small basins. By taking Irrawaddy as an example, we present time-series of the TWS variations re-scaled by the scaling factors from all five hydrological models and the median value and compare them with the re-scaled TWS from Tellus (Fig. 9). In this catchment, the TWS variations re-scaled by different scaling factors from the models show a comparably large spread, but they are consistent in a way that all the scaling factors are larger than one indicating signal loss caused by filtering. Besides, the TWS time-series re-scaled by median scaling factors generally lie in the middle. The results from Tellus show, however, a damping effect from the scaling factors. The TWS differences caused by applying two different scaling factors reach 10 cm, much larger than the total error (Table 2). This could be related to both the shape of the basin and also the hydrological signal within and around the basin. The Irrawaddy catchment is rather elongated and shares a long border with its neighbouring basins which also have high water storage variability. Both factors make it highly susceptible to spatial leakage effects and thereby

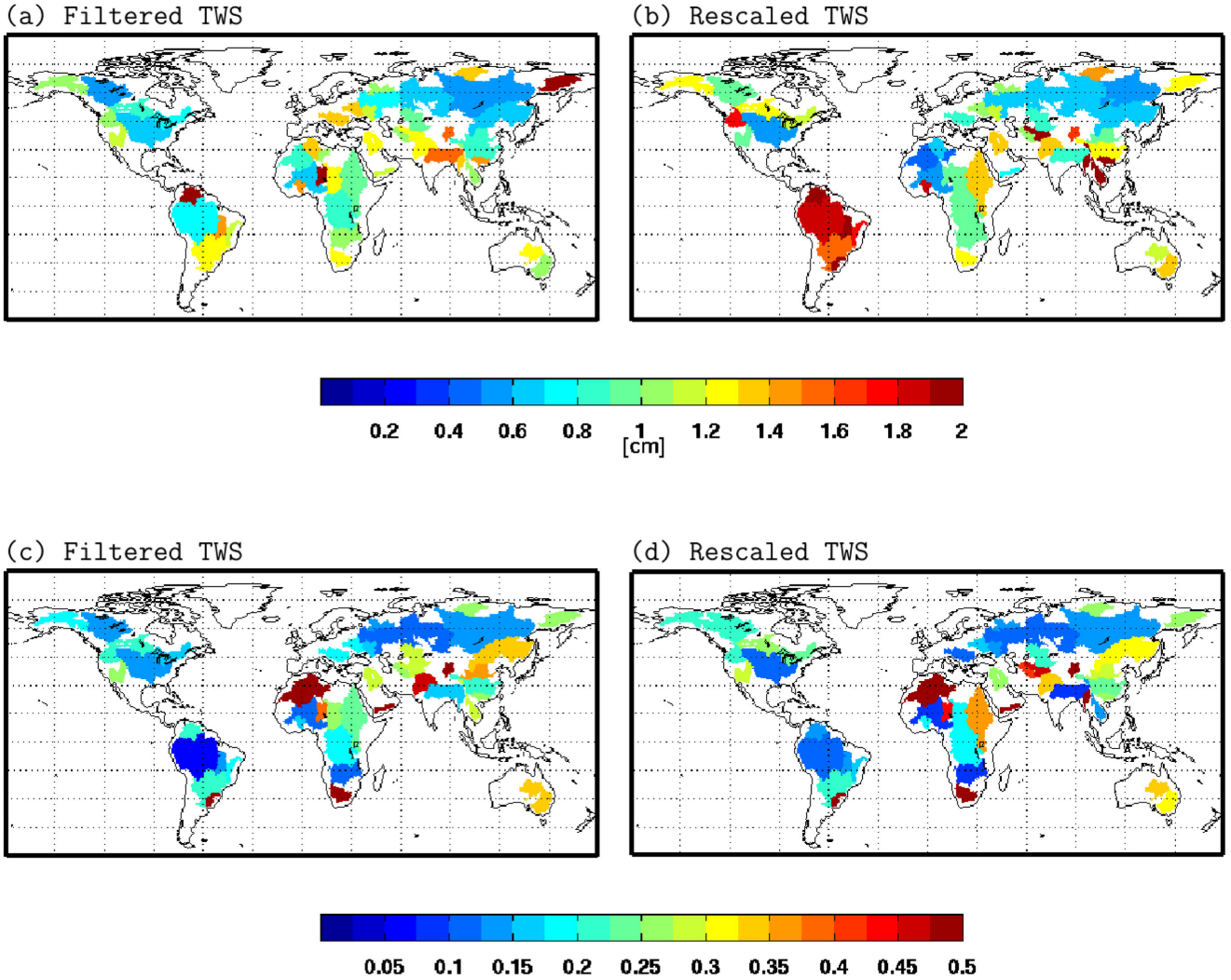

Figure 8. RMS of the differences between our estimated basin-averaged TWS filtered with DDK2 and the ones from Tellus for GFZ RL05a for the 50 largest discharge basins (a); and RMS of the differences between our filtered TWS re-scaled by the median scaling factors and the results from gridded TWS re-scaled by the scaling factors provided separately also on the Tellus website (b). (c) and (d) are the relative differences for (a) and (b) where the RMS of the differences are divided by the RMS of the basin-averaged TWS from Tellus.

vulnerable to uncertainties in simulated water storage in and outside the basin. If only a single model is used for the re-scaling, inherent uncertainties of this processing step remain inaccessible and might lead to additional errors in the GRACE-based TWS series that are not accounted for in the associated error estimates.

6 SUMMARY AND CONCLUSIONS

Globally gridded estimates of TWS anomalies have been processed from the GRACE release 05a monthly-mean gravity fields from GFZ Potsdam (Dahle *et al.* 2012) by applying state-of-the-art post-processing methodologies. The de-correlation filter of DDK2 has been chosen with the goal of minimizing signal loss while maximizing noise reduction. Re-scaling factors required to account for signal loss during filtering were obtained from the median values of a small ensemble of five global models in order to make the re-scaling more robust against particular weaknesses of a single model. We therefore intend to include those globally gridded TWS data-sets as an additional Level-3 product into the ICGEM website accessible at 'icgem.gfz-potsdam.de/ICGEM', so that it will be routinely updated as soon as new Level-2 gravity fields become available, and thereby contributes to a better accessibility of near real-time GRACE information to users that are not willing or not able to process Stokes coefficients themselves.

In addition to the monthly TWS estimates, we prepared realistic globally gridded error estimates by assessing individually the contributions of measurement errors, leakage errors, and re-scaling errors. The error estimates account for spatial correlations and yield largely consistent results for most basins when compared to estimates that are directly derived from the spherical harmonics representation. Thus, errors might guide users in selecting proper averaging regions and remind them that GRACE is in particular sensitive to the largest spatial scales as demonstrated by the fact that the SNR values generally decrease when going from large basins to small basins.

The RMS of the differences between the filtered TWS from our calculation and those from Tellus generally lie below the TWS error level estimated, which underlines the consistency of the two post-processing strategies. Larger differences found at certain basins

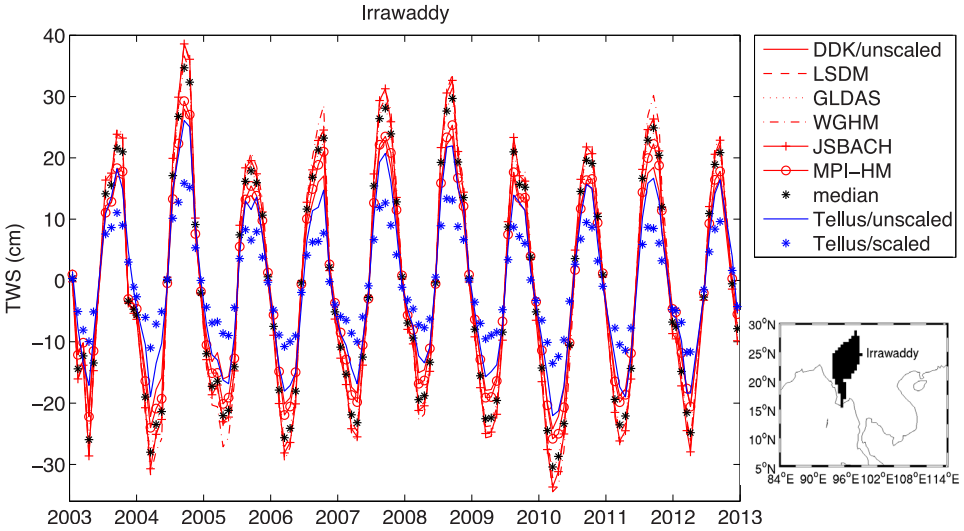

Figure 9. The comparison of the basin-averaged TWS time-series from our calculation and from Tellus for the Irrawaddy catchment. DDK/unscaled indicates the TWS variations that are filtered by DDK2; LSDM indicates the DDK2 filtered TWS re-scaled by scaling factors from LSDM, and the same for WGHM, GLDAS, JSBACH and MPI-HM; median is the TWS variations re-scaled by the median scaling factors from the five models; Tellus/unscaled is the basin average of the gridded TWS from Tellus website and Tellus/scaled is the re-scaled TWS from Tellus.

between the rescaled TWS time-series, however, emphasize the importance of model-based information required to account for spatial leakage. Since global land models perform differently in simulating the TWS variability in different areas of the world, an ensemble of multiple models is helpful to make scaling factors less affected by deficiencies in certain models. In view of the important role of such models for the GRACE processing, a detailed evaluation of the quality of such models and their systematic weaknesses is recommended.

Currently, the GRACE mission has been in orbit for more than 13 years and continues to provide monthly-mean snap-shots of the global gravity field. The GRACE Follow-On mission is already in its implementation phase and scheduled for launch in 2017 (Flechtner *et al.* 2014), thereby improving the prospects of establishing a long-term monitoring of global TWS variability with gravimetric methods. The GRACE mission has already contributed unique observations to five out of six current Grand Challenges of the World Climate Research Programme: (1) Melting Ice and Global Consequences (Sasgen *et al.* 2010), (2) Climate Extremes (Reager *et al.* 2014), (3) Regional Sea-Level Change (Chambers *et al.* 2010), (4) Water Availability (Famiglietti & Rodell 2013), and also (5) Decadal Climate Prediction (Zhang *et al.* 2015). The observing concept is therefore in a good position to be considered as a contribution to the 'Essential Climate Variables' (Hollmann *et al.* 2013) as defined by the World Meteorological Organization. To foster more applications of satellite gravimetry in scientific fields like hydrometeorology and climatology, conveniently pre-processed data-sets as described in this study are an essential prerequisite.

ACKNOWLEDGEMENTS

The authors thank Robert Dill, Andreas Güntner and Tobias Stacke for providing the LSDM, WGHM, JSBACH and MPI-HM model data. We would like also to thank Jennifer Bonin, two anonymous reviewers and the editor Kosuke Heki for their helpful and valuable comments. This study has been supported by the German Federal Ministry of Education and Research within the FONA research program under grants 03F0654A and 01LP1151A.

REFERENCES

Bergmann, I. & Dobslaw, H., 2012. Short-term transport variability of the antarctic circumpolar current from satellite gravity observations, *J. geophys. Res.,* **117**(C5), doi:10.1029/2012JC007872.

Bergmann-Wolf, I., Zhang, L. & Dobslaw, H., 2014. Global eustatic sea-level variations for the approximation of geocenter motion from GRACE, *J. Geod. Sci.,* **4**(1), 37–48.

Biancale, R. & Bode, A., 2006. Mean annual and seasonal atmospheric tide models based on 3-hourly and 6-hourly ECMWF surface pressure data, Scientific Technical Report STR06/01, GFZ, Helmholtz-Zentrum, Potsdam.

Brovkin, V., Raddatz, T., Reick, C.H., Claussen, M. & Gayler, V., 2009. Global biogeophysical interactions between forest and climate, *Geophys. Res. Lett.,* **36**(7), doi:10.1029/2009GL037543.

Chambers, D.P. & Bonin, J.A., 2012. Evaluation of Release-05 GRACE time-variable gravity coefficients over the ocean, *Ocean Sci.,* **8**(5), 859–868.

Chambers, D.P., Wahr, J., Tamisiea, M.E. & Nerem, R.S., 2010. Ocean mass from GRACE and glacial isostatic adjustment, *J. geophys. Res.,* **115**(B11), B11415, doi:10.1029/2010JB007530.

Chen, J., Wilson, C., Famiglietti, J. & Rodell, M., 2007. Attenuation effect on seasonal basin-scale water storage changes from GRACE time-variable gravity, *J. Geod.,* **81**(4), 237–245.

Cheng, M., Ries, J.C. & Tapley, B.D., 2011. Variations of the Earth's figure axis from satellite laser ranging and GRACE, *J. geophys. Res.,* **116,** 1409, doi:10.1029/2010JB000850.

Dahle, C., Flechtner, F., Gruber, C., König, D., König, R., Michalak, G. & Neumayer, K., 2012. GFZ GRACE Level-2 Processing Standards Document for Level-2 Product Release 0005, Scientific technical report-data, GFZ, Helmholtz-Zentrum, Potsdam, Potsdam.

Dee, D.P. *et al.*, 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.,* **137**(656), 553–597.

Dill, R., 2008. Hydrological model LSDM for operational Earth rotation and gravity field variations, GFZ Scientific Technical Report-STR08/09, GFZ, Helmholtz-Zentrum, Potsdam, Potsdam.

Dill, R. & Dobslaw, H., 2013. Numerical simulations of global-scale high-resolution hydrological crustal deformations, *J. geophys. Res.,* **118**(9), 5008–5017.

Dobslaw, H., Flechtner, F., Bergmann-Wolf, I., Dahle, C., Dill, R., Esselborn, S., Sasgen, I. & Thomas, M., 2013. Simulating high-frequency

atmosphere-ocean mass variability for dealiasing of satellite gravity observations: AOD1B RL05, *J. geophys. Res.,* **118,** 3704–3711.

Döll, P., Kaspar, F. & Lehner, B., 2003. A global hydrological model for deriving water availability indicators: model tuning and validation, *J. Hydrol.,* **270**(1–2), 105–134.

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, *Surv. Geophys.,* **35**(6), 1285–1309.

Famiglietti, J.S. & Rodell, M., 2013. Water in the balance, *Science,* **340**(6138), 1300–1301.

Flechtner, F., Watkins, M., Morton, P. & Webb, F., 2014. Status of the GRACE Follow-on Mission, in *Proceedings of the International Association of Geodesy Symposia Gravity, Geoid and Height System (GGHS2012),* vol. IAGS-D-12-00141, Venice, Italy.

Frappart, F., Seoane, L. & Ramillien, G., 2013. Validation of GRACE-derived terrestrial water storage from a regional approach over South America, *Remote Sens. Environ.,* **137,** 69–83.

Gudmundsson, L., Wagener, T., Tallaksen, L.M. & Engeland, K., 2012. Evaluation of nine large-scale hydrological models with respect to the seasonal runoff climatology in Europe, *Water Resour. Res.,* **48**(11), W11504.

Gunkel, A. & Lange, J., 2011. New insights into the natural variability of water resources in the lower jordan river basin, *Water Resour. Manage.,* **26**(4), 963–980.

Güntner, A., Stuck, J., Werth, S., Döll, P., Verzano, K. & Merz, B., 2007. A global analysis of temporal and spatial variations in continental water storage, *Water Resour. Res.,* **43**(5), W05416, doi:10.1029/2006WR005247.

Hagemann, S. & Gates, L., 2001. Validation of the hydrological cycle of ECMWF and NCEP reanalyses using the MPI hydrological discharge model, *J. geophys. Res.,* **106,** 1503–1510.

Hagemann, S. & Gates, L., 2003. Improving a subgrid runoff parameterization scheme for climate models by the use of high resolution data derived from satellite observations, *Clim. Dyn.,* **21**(3–4), 349–359.

Hollmann, R. *et al.*, 2013. The ESA climate change initiative: satellite data records for essential climate variables, *Bull. Am. Meteorol. Soc.,* **94,** 1541–1552.

Houborg, R., Rodell, M., Li, B., Reichle, R. & Zaitchik, B.F., 2012. Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations, *Water Resour. Res.,* **48**(7), W07525, doi:10.1029/2011WR011291.

Jacob, T., Wahr, J., Pfeffer, W.T. & Swenson, S., 2012. Recent contributions of glaciers and ice caps to sea level rise, *Nature,* **481,** 514–518.

Jungclaus, J.H. *et al.*, 2013. Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean compartment of the MPI-Earth system model, *JAMES,* **5**(2), 422–446.

Klees, R., Zapreeva, E.A., Winsemius, H.C. & Savenije, H.H.G., 2007. The bias in GRACE estimates of continental water storage variations, *Hydrol. Earth Syst. Sci.,* **11**(4), 1227–1241.

Kusche, J., 2007. Approximate decorrelation and non-isotropic smoothing of time-variable GRACE-type gravity field models, *J. Geod.,* **81**(11), 733–749.

Kusche, J., Schmidt, R., Petrovic, S. & Rietbroek, R., 2009. Decorrelated GRACE time-variable gravity solutions by GFZ, and their validation using a hydrological model, *J. Geod.,* **83**(10), 903–913.

Landerer, F.W. & Swenson, S.C., 2012. Accuracy of scaled GRACE terrestrial water storage estimates, *Water Resour. Res.,* **48**(4), W04531, doi:10.1029/2011WR011453.

Long, D., Longuevergne, L. & Scanlon, B.R., 2015. Global analysis of approaches for deriving total water storage changes from GRACE satellites, *Water Resour. Res.,* **51**(4), 2574–2594.

Mayer-Gürr, T., Zehentner, N., Klinger, B. & Kvas, A., 2014. ITSG-GRACE2014: A new GRACE gravity field release computed in Graz, GRACE Sci. Team Meet., Potsdam.

Müller Schmied, H., Eisner, S., Franz, D., Wattenbach, M., Portmann, F.T., Flörke, M. & Döll, P., 2014. Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration, *Hydrol. Earth Syst. Sci.,* **18**(9), 3511–3538.

Petit, G. & Luzum, B., 2010. IERS Conventions (2010), IERS Technical Note; 36, Bundesamt für Kartographie und Geodäsie, Frankfurt am Main.

Raddatz, T.J. *et al.*, 2007. Will the tropical land biosphere dominate the climate-carbon cycle feedback during the twenty-first century?, *Clim. Dyn.,* **29**(6), 565–574.

Reager, J.T., Thomas, B.F. & Famiglietti, J.S., 2014. River basin flood potential inferred using GRACE gravity observations at several months lead time, *Nat. Geosci.,* **7**(8), 588–592.

Rodell, M., Houser, P., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K. & Cosgrove, B., 2004. The Global Land Data Assimilation System, *Bull. Am. Meteorol. Soc.,* **85,** 381–394.

Sasgen, I., Dobslaw, H., Martinec, Z. & Thomas, M., 2010. Satellite gravimetry observation of Antarctic snow accumulation related to ENSO, *Earth planet Sci. Lett.,* **299**(3–4), 352–358.

Savcenko, R. & Bosch, W., 2012. EOT11a - Empirical Ocean Tide Model from Multi-Mission Satellite Altimetry, DGFI Report No. 89, Deutsches Geodätisches Forschungsinstitut (DGFI), München.

Stacke, T. & Hagemann, S., 2012. Development and evaluation of a global dynamical wetlands extent scheme, *Hydrol. Earth Syst. Sci.,* **16**(8), 2915–2933.

Steffen, H., Denker, H. & Müller, J., 2008. Glacial isostatic adjustment in Fennoscandia from GRACE data and comparison with geodynamical models, *J. Geodyn.,* **46**(3–5), 155–164.

Stevens, B. *et al.*, 2013. Atmospheric compartment of the MPI-M Earth System Model: ECHAM6, *JAMES,* **5**(2), 146–172.

Swenson, S. & Wahr, J., 2006. Post-processing removal of correlated errors in grace data, *Geophys. Res. Lett.,* **33**(8), L08402, doi:10.1029/2005GL025285.

Swenson, S., Chambers, D. & Wahr, J., 2008. Estimating geocenter variations from a combination of GRACE and ocean model output, *J. geophys. Res.,* **113,** 8410, doi:10.1029/2007JB005338.

Syed, T.H., Famiglietti, J.S., Rodell, M., Chen, J. & Wilson, C.R., 2008. Analysis of terrestrial water storage changes from GRACE and GLDAS, *Water Resour. Res.,* **44**(2), W02433, doi:10.1029/2006WR005779.

Tapley, B.D., Bettadpur, S., Watkins, M. & Reigber, C., 2004. The gravity recovery and climate experiment: mission overview and early results, *Geophys. Res. Lett.,* **31**(9), L09607, doi:10.1029/2004GL019920.

Vörösmarty, C., Fekete, B., Meybeck, M. & Lammers, R., 2000. A simulated topological network representing the global system of rivers at 30-minute spatial resolution (STN-30), *Glob. Biogeochem. Cycle,* **14,** 599–621.

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 Resour. Res.,* **49**(2), 904–914.

Wahr, J., Molenaar, M. & Bryan, F., 1998. Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE, *J. geophys. Res.,* **103**(B12), 30205–30229.

Wahr, J., Swenson, S. & Velicogna, I., 2006. Accuracy of GRACE mass estimates, *J. geophys. Res.,* **33**(6), L06401, doi:10.1029/2005GL025305.

Weedon, G.P. *et al.*, 2011. Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century, *J. Hydrometeorol.,* **12,** 823–848.

Werth, S., Güntner, A., Schmidt, R. & Kusche, J., 2009. Evaluation of GRACE filter tools from a hydrological perspective, *Geophys. J. Int.,* **179**(3), 1499–1515.

Zhang, L., Dobslaw, H., Dahle, C., Sasgen, I. & Thomas, M., 2015. Validation of MPI-ESM Decadal Hindcast experiments with terrestrial water storage variations as observed by the GRACE satellite mission, *Meteorol. Z.,* doi:10.1127/metz/2015/0596.