Mapping probabilities of extreme continental water storage changes from space gravimetry

J. Kusche[1], A. Eicker[1], E. Forootan[1,2], A. Springer[1], and L. Longuevergne[3]

Corresponding author: J. Kusche, Institute of Geodesy and Geoinformation, University of Bonn, Nussallee 17, D-53115 Bonn, Germany. (kusche@uni-bonn.de)

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

[2]School of Earth and Ocean Sciences, Cardiff University, Cardiff, UK.

[3]CNRS, Geosciences Rennes, University or Rennes 1, Rennes, France.

Key Points.

1. From 12 years of GRACE data, we derive statistically robust 'hotspot' regions of high probability of peak anomalous water storage and flux.
2. Comparison to ERA-Interim reanalysis reveals good agreement of these regions to GRACE, with most exceptions located in the Tropics.
3. Provided GRACE will be succeeded in time by GRACE-FO, by around year 2020 we will be able to detect changes in the frequency of peak total flux.

Using data from the Gravity Recovery and Climate Experiment (GRACE) mission, we derive statistically robust 'hotspot' regions of high probability of peak anomalous – i.e. with respect to the seasonal cycle – water storage (of up to 0.7 m one-in-five-year return level) and flux (up to 0.14 m/mon). Analysis of, and comparison to, up to 32 years of ERA-Interim reanalysis fields reveals generally good agreement of these hotspot regions to GRACE results, with most exceptions located in the Tropics. A simulation experiment reveals that dif-

ferences observed by GRACE are statistically significant. Further error analysis suggests that, provided we will have a continuation of GRACE by its follow-up GRACE-FO, we will likely be able around year 2020 to detect temporal changes in the frequency of extreme total fluxes (i.e. combined effects of mainly precipitation and floods) for at least one tenth to one fifth of the continental area.

1. Introduction

Due to its memory effect, terrestrial water storage contains information on antecedent rainfall and runoff conditions that, to some extent, control future drought and flood occurence and severity. However, at the time of writing, the NASA/DLR GRACE twin-satellite mission represents the only platform that observes terrestrial water storage with global coverage. GRACE has provided an unprecedented record of more than 14 years of monthly terrestrial water storage anomaly maps. The GRACE satellites show signs of ageing, but with its successor GRACE-FO set for launch in 2017 [*Flechtner et al.*, 2016], it appears possible that we will soon have an almost uninterrupted observational record of terrestrial water storage over three decades. The primary observable of GRACE, time-variable changes in the Earth's geopotential measured via precise intersatellite ranging, has provided a new view on the ongoing patterns of mass redistribution at the planet's surface, in particular related to the terrestrial and oceanic hydrological cycle.

Several researchers quantified the variability of water storage in form of groundwater, soil moisture, and surface water [*Forootan et al.*, 2014], snowpack and ice [*Velicogna et al.*, 2014], and mass-driven sea level [*Rietbroek et al.*, 2016], on different timescales from interannual to days. Observed variability in groundwater storage has been attributed to episodic events such as droughts and floods, to 'natural' variability related to modes of the climate system [*Phillips et al.*, 2012], and to anthropogenic effects such as depletion [*Döll et al.*, 2014] and land use change.

Ogawa et al. [2011] have shown how GRACE data can be related to total terrestrial water flux, the sum of precipitation, evapotranspiration, and runoff, and *Springer et al.* [2014]

suggested its use to validate the water cycle in atmospheric reanalyses. In these applications, numerical differentiation schemes are applied to total water storage time series in order to derive flux. At longer time scales, anomalies of total flux with respect to a mean state can be linked to the sum of (i) modifications of the land boundary conditions and the resulting climate forcing, (ii) the direct and indirect impact of anthropogenic activitites, and (iii) the hydrological response of the system [*Eicker et al.*, 2016]. Important, at shorter time scales and at grid scale, GRACE data relate to lateral water redistribution: water storage increase (flux has a positive sign) corresponds to precipitation plus upstream river flow, while storage decrease (flux has a negative sign) corresponds to evapotranspiration plus river discharge. However, inferring lateral transports from GRACE is difficult since month-to-month variability in GRACE data is contaminated by stronger noise and limited in spatial resolution, when compared to longer timescales.

Estimating the frequency or probability of future events based on time-limited records represents an established concept in hydrology and hydrological engineering (e.g. *Beard* [1962], *Stedinger et al.* [1992]). Drought and flood indicators can be expressed as percentiles in reference to their historical frequency of occurrence. For example, the U.S. Drought Monitor combines several short-term and long-term indices and indicators in this way for each location and time of year. Since only little information on deep soil moisture and groundwater enters common drought indices, GRACE data are being assimilated in the Catchment Land Surface Model, and assimilated fields are converted into soil moisture and groundwater percentiles [*Houborg et al.*, 2012].

Few studies so far (e.g. *Moore and Williams* [2014], *Humphrey et al.* [2016]) have attempted to look directly at the statistical behaviour of 'anomalous' GRACE signals, i.e. beyond the dominating seasonal cycle and beyond episodic drought and flood events, and no study is known to us that quantifies occurrence frequency and expected return levels of such changes in a probabilistic sense.

In particular interesting would be, in the light a hypothesized intensification of the water cycle (*Huntington* [2006], or [*Durack et al.*, 2012]), whether and if, after what time, changes in the occurrence frequency of extremes in storage and flux, including floods and droughts, can be observed with space gravimetry. In a probabilistic view, changes in the mean and variance of the distribution underlying a climate variable affect the severity and occurrence frequency of extremes [*Folland et al.*, 2001]. Return times of events of a given magnitude, or return levels for a given return time as considered here, are sensitive indicators to increases in magnitude in the tails of the underlying distribution [*Allen and Ingram*, 2002]. For example, the CMIP5 analysis by *Yoon et al.* [2015] projects an increase in the sliding-window variance of California top-1m soil moisture that equates to an increase of intense droughts and excessive floods by at least 50% towards the end of the twenty-first century. Validating such studies using GRACE/GRACE-FO would be of tremendous significance.

Here, we analyse annual peak high and low levels of water storage and in storage change (total flux), observed by GRACE, for their recurrence frequency and levels in a probibilistic framework. Recurrence frequency and return level are equivalent once the underlying distribution is known, so we express all findings as return levels. We outline hotspot regions where large anomalies are to be expected, which differ, to some extent, from what is expected based on 32 years of ECMWF ERA-Interim reanalysis. Based on a realistic simulation, we then discuss the probability of detecting temporal changes in recurrence frequency of total water flux with the future combined GRACE/GRACE-FO data record.

2. Data and methods

Total water storage (TWS) represents aggregated variations in the terrestrial water content with respect to a long-term mean; thus reflecting the combined effect of changes in groundwater volume, soil moisture, root and canopy water content, and lake, river and reservoir levels. We use GRACE data to derive TWS as follows: Monthly spherical harmonic coefficients (University of Texas, release 5) for the 2003.0-2015.0 timespan, augmented by geocenter, c_{20}, and glacial isostatic adjustment (GIA) corrections and decorrelated/smoothed through the DDK3 method as in *Eicker et al.* [2016], are mapped to 1° grids. Finally, in order to focus on departures from the large average seasonal water storage modes, we first remove a six-parameter model (mean, rate, annual and semi-annual waves), and then, subsequently, the monthly residual TWS climatology from the grids.

The instantanous rate of change of TWS corresponds, according to mass conservation, to the sum of precipitation, evapotranspiration and runoff, and we denote this quantity as total water flux (TWF) here. From the GRACE coefficients, TWF grids are derived following methods outlined in *Eicker et al.* [2016]. Since TWF exhibits more noise due to temporal differentiation, we chose to apply slightly more agressive spatial filtering (DDK2) as compared to TWS grids.

It is not clear whether any current hydrological or land surface model captures the full storage capacity in all soil and groundwater layers, and lends itself for providing a reference for GRACE-derived extreme TWS and TWF under either stationarity or non-stationarity assumptions. In this work, both for comparison and multi-decadal simulation purposes, TWF grids are derived directly from ERA-Interim reanalysis fields [*Dee et al.*, 2011] of precipitation, evapotranspiration and runoff. We realize that limitations of the underlying land surface model (H-TESSEL) exist, but results in *Eicker et al.* [2016] suggest that ERA-Interim and GRACE data fit well at shorter timescales. To enforce spectral consistency, these fields are first converted to spherical harmonic representation, filtered using the same procedures as with GRACE, and converted back to grids as in *Eicker et al.* [2016].

From this point on, our method is as follows (see supporting information): We decimate all grid time series first to annual maximum and minimum anomalous storage and flux, and compute mean, standard deviation and skewness (up to 2.6-2.9) for these peak series. A Generalized Extreme Values (GEV) distribution is then fitted using a moment method [*Martins and Stedinger*, 2000], and return levels (i.e. expected maximum or minimum after N years) are computed. This has the advantage that occurrence frequencies can be represented through a single, physically interpretable value (the N-year return level in m or m/mon) per grid point. Since the GRACE record is rather short compared to precipitation, discharge or sea level data where GEV analysis is common, we restrict ourselves in this study to one-in-five-year return levels, and no attempt is made to extrapolate return level curves to more infrequent extremes.

3. Probabilities of anomalously high or low water storage

Following this approach, Fig. 1 illustrates return levels of annual anomalously high (a) and low (b) water storage from GRACE, with respect to the monthly TWS climatology. Expected one-in-five-year peak water levels reach up to 0.70 m, with dominating regions being the Central Amazon and the Mississippi (related to catastrophic 2011 floods) basins, and a range of regions at the 0.2 - 0.4 m level; such as the South America Parana basin, Central Africa (including the Zambezi), India, Northern Australia, Turkey, and North-East China (it is important to understand that expected N-year levels can be larger than those actually observed within any N-year period). Measured by the latitude-weighted RMS, land-averaged return levels amount to 0.14 m. As an aside, we note that one-in-ten-year levels are generally found about 25% larger than one-in-five year levels.

It must be understood that, at GRACE temporal and spatial resolution, hydro-meteorological extreme events are difficult to relate to common flood or rainfall peak levels or return intervals. For example, the 2011 Mississippi 500-year flood inundated an area of several thousand km^2 by the order of meters, as a result of rainfall rates of 50 cm per week but concentrated within few days. Though GRACE results are typically expressed in metric 'equivalent water height', due to its measurement principle the mission observes water mass (or 'equivalent volume') which is difficult to scale to observable water levels. As a result, GRACE-derived extreme events always refer to monthly large-scale averages and may miss, or average out, 'real' extreme events that are focused in space and time by nature.

In contrast to annual maxima, one-in-five-year levels of exceptional low (i.e. below climatology) water storage are found reaching 0.55 m in the Amazon, and on average over land masses, less than 0.14 m. It is interesting to note that extreme levels in TWS are not symmetric; some regions affected by floodings (e.g. Mississippi basin, Lake Victoria) feature prominently in Fig. 1.a while some others show up only (Amazonas river mouth) in Fig. 1.b, but overall the maps are quite similar. 14-year minimum TWS events have been analysed in *Humphrey et al.* [2016], and their Fig. 14 of maximum average storage deficit and year of maximum resembles our Fig. 1.b. This confirms that levels of high probability of low water storage in our Fig. 1.b are typically related to the occurrence of two or three strong droughts (some of which may not have been described in literature, *Humphrey et al.* [2016]).

As expected, these hotspot regions of extreme annual anomalous storage broadly correspond to regions where seasonal water storage amplitudes are large (see Fig. S1), but they also reflect that GRACE picks up anomalous floods and droughts for Southern Australia, the Parana basin (where the large groundwater response to climate variability was shown in *Chen et al.* [2010]), or North East China regions where the annual signal is less prominent. On the contrary, Fig. 1 does not prominently feature part of the Amazon, Alaska coastal glaciers, and the Ganges-Brahmaputra delta where surface or snow-equivalent water loads are huge but mostly follow the seasonal cycle (Fig. S1).

Figure 1 (c, d, g, and h) show time series (black dots), annual maximum TWS levels (red dots), observed frequency of maximum TWS (red bars) and the fitted GEV distribution (grey) for the two locations Central Amazon (c,d), Parana (g, h) indicated in Fig. 1 (a), while figures on the right-hand side show the same for minimum TWS levels (Cuvelai-

Etosha e, f, Northern Australia i, j). Locations have been chosen to display different behaviour: Central Amazon where annual signals are among the largest on Earth, with a wide spread of both annual maxima and minima of anomalous TWS, while for the Cuvelai-Etosha basin [*Eicker et al.*, 2016] a multiannual oscillation appears to be present. For the Parana basin, again the seasonal signal is weak but extreme levels peak every two to three years, likely related to ENSO (*Chen et al.* [2010], *Phillips et al.* [2012], *Eicker et al.* [2016]). We note that, with the exception of the Central Amazon location, the GEV distribution appears quite suitable for fitting to observed extreme levels of storage. As will be shown later, our GEV fits are less sensible with respect to record length compared to, e.g. Gaussian fits.

4. Probabilities of anomalous increase or decrease of water storage

Analysis of the time-differentiated GRACE record reveals a number of regions of increased probability of maximum (Fig.2.a) and minimum (Fig.2.b) water flux that broadly correspond to those of anomalous TWS but in general follow rainfall patterns such as the monsoon. One-in-five-year levels of annual peak flux (Fig.2.a) amount up to 0.14 m/mon for the Central Amazon region, with an overall land-average weighted RMS of 0.033 m/mon. We remind that annual extremes in TWF relate to the fastest increase (linked to extreme precipitation) or decrease of total water storage per given year; and peak maxima in the figures have to be interpreted as levels of storage increase or decrease that statistically occur once every five year. Peak one-in-five year decrease (Fig.2.b) reaches up to 0.14 m/mon, with a land-average RMS of 0.031 m/mon.

In ERA-Interim (Fig.2.c), tropical precipitation extremes dominate total flux and contribute to one-in-five-year maximum levels up to 0.31 m/mon over Tropical Northern Australia and South-East Asia, with RMS close to 0.040 m/mon. Minimum levels (Fig.2.c) are up to 0.27 m/mon, land-averaged to RMS 0.036 m/mon. In fact, ERA-Interim identifies many regions outside the Tropics that closely correspond to GRACE-derived extreme levels, with some exceptions (Southern Europe, US/Canada West coast, East Europe/Russia). Overall, we find an average difference between GRACE-derived TWF and ERA-Interim reanalysis fields of only RMS 0.023 m/mon (max) and 0.021 m/mon (min).

Humphrey et al. [2016] found significant positive correlation between GRACE high-frequency anomalies and ERA-Interim precipitation over many regions, that we identify here as having high probability of maximum water flux: the Amazon and Parana basins, Northern Australia, South/Central Africa, Northern India, South-East Europe, parts of the U.S.. This supports our hypothesis that extreme levels of TWS increase (positive TWF) are likely driven by precipitation.

5. Stationarity with respect to climate modes: ENSO

It is possible that our results are influenced by the occurrence of climate modes within the analysis time frame. In fact, *Phillips et al.* [2012] and *Eicker et al.* [2016] have shown that GRACE-derived water storage is correlated with ENSO, and other authors have identified correlations e.g. with the Pacific Decadal Oscillation (PDO, e.g. *Seoane et al.* [2013]). While these studies generally focused on identification of modes and problems in estimation of trend and accelerations, we here focus on the imprint of ENSO on the occurrence probability of extreme storage and flux.

In what follows, we repeat our previous experiments but we exclude either (1) years 2003, 2009, and 2010, or (2) 2007, 2008, 2010 and 2011 from our analysis. These years were, according to the Ocean Niño Index (ONI, a three-month running mean of sea surface temperature anomalies in the Niño3.4 region) categorized as (1) moderate or stronger El Niño years (ONI>1) or (2) moderate or stronger La Niña years (ONI< -1).

Results are shown in Fig. S2, and can be compared to Fig.1 which shows return levels derived from the full time series including ENSO years. We find that, overall, our results are surprisingly robust. Largest differences can be observed for South America. In case of (1), excluding El Niño years, maximum one-in-five-year water levels are reduced from 0.70 m (full 12-year period) to 0.50 m with land RMS 0.013 m, while for (2) excluding La Niña years leads to up to 0.53 m with RMS 0.014 m. Yet, removing El Niño years does not lead to a general smoothing, and for some regions one-in-five-year levels slightly increase. In contrast, minimum five-year water levels (0.55 m for 12-year period) increase by (1) removing El Niño years to up to 0.65 m with RMS 0.014 m, while for (2) excluding La Niña years leads to up to 0.63 m with RMS 0.013 m.

Results for total flux (not reported here) point in the same direction. In line with expectations, we conclude that ENSO, to some extent, influences extreme high water volumes and less so extremely low levels (storage deficit events). It is also interesting to note that El Niño and La Niña do not appear to have a symmetric effect on water surplus; although due to the reduced sample size such comparisons are problematic and need to be repeated once we have longer data records.

6. Detecting an intensification from a future combined GRACE and GRACE-FO record

With 14 years of GRACE data at the time of writing, and the GRACE Follow-On mission (GRACE-FO) on track for launch in late 2017, it is reasonable to ask whether, and after what time, a continuous multi-decadal data set of TWS and TWF will enable us to detect temporal changes in the frequency of extreme water storage and water flux events. In order to answer this question, we conduct a twin experiment: (1) We derive the ocurrence frequency of peak total water flux in ERA-Interim, when analyzed over varying time frames from 32 years to 12 years, all ending January 2015. (2) We create a synthetic, composite 32-years GRACE/GRACE-FO data set which is then analyzed for peak TWF frequency over varying analysis intervals. This data set is derived from the 'truth' ERA-Interim data by adding realistic, spatially anisotropic GRACE-errors. In a conservative approach, we assume that GRACE-FO will have the same error characteristics as GRACE (both GRACE and GRACE-FO errors are synthesized from fully populated, monthly covariance matrices from real-data GRACE analysis over 2013 and 2014, as in *Landerer and Swenson* [2012], but in random permutations over all 32 years).

In Fig. 3, we represent ERA-Interim-derived one-in-five-year peak water flux (left) and the same statistics derived from synthetic GRACE/GRACE-FO (i.e. 'GRACE-perturbed' ERA-Interim, Fig. 3 right), for 2003.0-2015.0 (top), 1991.0-2015.0 (center), and 1983.0-2015.0 (bottom). All results are summarized in Table 1.

Differences between ERA-Interim peak TWF return levels over differing time spans, yet referring to the same seasonal model, are small (cf. Table 1) but can be identified for some regions (Orinoco basin, North-West India, Siberia). Such differences may occur due to either problems in fitting the GEV distribution to small samples (comparing moment and maximum likelihood (ML) estimation, or using the information matrix from the ML approach [*Hosking*, 1985], suggests that 1σ values may be at the 0.01 - 0.02 m/mon level from 12 years) or owing to real nonstationarity. But in our twin experiment we will assume they represent the 'target signal' to be detected from the gravity mission records. As expected, differences grow with decreasing analysis window. As a reference, for the same 12-year period for which we analysed real GRACE data, reanalysis-derived return levels differ from those derived from 32 years by about 0.01 m/mon RMS and up to 0.09 m/mon for the above mentioned regions. We note that it is of course possible that ERA-Interim fails to capture real nonstationarity; in this case our GRACE/GRACE-FO simulation is biased towards stationarity and conclusions on detectability may be too conservative.

Differences between one-in-five-year levels of peak TWF from simulated GRACE/GRACE-FO data and the corresponding 'truth' ERA-Interim derived levels (right column in Table 1) vary from 0.007 m (32 years) to 0.011m (8 years); they depend on the data record length but much less compared to Gaussian statistics such as applied in trend estimates. Our twin experiment simulates that for the 12 year period analysed in this study from real GRACE data, average RMS errors may be slightly above the 0.01 m level. In fact, GRACE-like errors amplify near-zero peak levels over the Sahara desert (Fig. 3 b, d, f) to about 0.02 m/mon which is almost exactly what we observe from the real GRACE data (Fig. 2 a and b); the visual correspondence is striking, suggesting that our error model may be quite close to the real noise. With the above, we can conclude that we find

ERA-Interim return levels above the noise level for nearly 90% of the total land area.

Finally, the RMS fit of GRACE/GRACE-FO peak levels from differing analysis windows to the 32-year ERA-interim five-in-one-year return levels (middle column in Table 1) suggest that temporal variability (or sampling error) of 'true' water flux and the effect of GRACE errors add up virtually independent. This metric tells how good the satellite-derived return levels from limited observational records would reconstruct 'true' return levels under the assumption of stationarity.

The main conclusion from the twin experiment is that GRACE/GRACE-FO errors will allow to derive return level statistics for water flux extremes at the 0.010 - 0.012 m/mon land average error level within a 12 year temporal window (Table 1), only slighly less accurate from 8-year windows, and more accurate down to 0.007 m/mon from 32 years. In the light of our 12-year comparison of GRACE and ERA-Interim discussed in the previous section, this means that the differences observed by GRACE (Fig. 3 (a) vs. (c) and (b) vs. (d)) exceed a noise level of 0.01 m/mon for about 40% of the land area (and, conservatively, a noise level of 0.03 m/mon for still more than 9 % of the area).

This suggests that we apply a moving-window approach to a near-future GRACE/GRACE-FO time series in order to isolate temporal changes in extreme water flux frequency. Our error estimates can be compared against such changes in the ERA-Interim record, to understand for which part of the landmass changes would be detected with statistical significance. Yet, identifying such an 'intensification', leading to changes in the mean, variability, or skewness of extremes and therefore to a change in e.g. one-in-five-year levels, requires to contrast a certain window against a reference period of at least the same duration. With this in mind, we deduce that for a 24-year data record with two 12-year windows (i.e. corresponding to 2027, likely within the GRACE-FO lifetime), we could have detected about 13 - 18% of the frequency changes in ERA-Interim (this is the share of land area where changes exceed the noise). Surprisingly, for a 32-year data record, divided into two windows each of the about the lifetime of GRACE, only for 10 - 16 % of area such changes in ERA-Interim would be detected; this is since the longer timeframe despite allowing for better GEV fitting tends to temporally average out changes in peak frequency. On the contrary, our results suggest that already for a 16-year record (i.e. 2019) for 13 - 21 % of land area those changes in frequency that were captured in the recent 8-year period in ERA-Interim with respect to the previous one would be detectable.

7. Conclusions

Climate variability as well as a range of direct and indirect anthropogenic modifications of the water cycle cause land-atmosphere water fluxes and surface runoff to depart from the regional climatology on a range of timescales. Such anomalous total water flux and total water storage signals can be observed with the GRACE satellite mission and, with its successor GRACE-FO hopefully launched in 2017, we may have soon a multi-decade observational record that can be used to inform model simulations.

Here we have focused on the occurrence of extreme, annual maximum or minimum anomalous fluxes and storages in the GRACE record. From 144 months of GRACE data, we quantify and map return levels (expected anomalous flux or storage once in N years) of these extremes, with good statistical significance. We find that most hotspot regions correspond to regions of known large storage amplitudes due to groundwater variability or seasonal flooding or inundation, but the situation is more complex and not symmetric. Few studies so far have aimed at lateral water redistribution using GRACE due to its comparably low resolution, but here we show that this is largely possible when focusing on extreme events.

The current GRACE data set has been used before to isolate and study the response of total water storage and of groundwater and river discharge to extreme events such as heat-waves and heavy-precipitation years. Yet it is too short to derive conclusions on changes in the probability of such events. But provided the GRACE mission will be succeeded in time by GRACE-FO, we conclude that around year 2020 we will be able to detect changes in the frequency of extreme total fluxes for at least one tenth to one fifth of the continental area, when assuming the magnitude of such changes corresponds to what we observed in the ERA-Interim reanalysis over the past decades. We anticipate that such changes may occur along with an intensified water cycle due to global warming as the combined effects of precipitation and floods. Yet, there is no consensus on what exactly may happen in the future, and where, and we suggest that a combined GRACE/GRACE-FO record may provide a useful additional observational data sets in order to test hypotheses regarding the changing water cycle.

Acknowledgments. The University of Texas GRACE solutions are available via the GFZ Potsdam Information Systems and Data Center (ISDC, isdc.gfz-potsdam.de) as well as the JPL Physical Oceanography Distributed Active Archive Center (PODAAC, www.grace.jpl.nasa.gov). The European Center for Midrange Weather Forecasting provided the ERA-Interim data, which are publicly available.

References

Allen, M. and W. Ingram, Constraints on future changes in climate and the hydrologic cycle. *Nature*, *419*, 224-229

Beard, L. (1962), Statistical methods in hydrology. *U.S. Army Corps of Engineers*, Sacramento, CA

Chen, J.L., C. Wilson, B. Tapley, L. Longuevergne, Z. Yang, and B. Scanlon (2010), Recent La Plata basin drought conditions observed by satellite gravimetry, *J. Geophys. Res. (Atmospheres)*, *115*, D22108

Dee, D.P., and 35 authors (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, *137*(656), 553-597

Döll, P., H. Müller Schmied, C. Schuh, F. T. Portmann, and A. Eicker (2014), Global-scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites, *Water Resour. Res.*, *50*, 56985720

Durack, P., S. Wijffels, and R. Matear (2012), Ocean salinities reveal strong global water cycle intensification during 1950 to 2000, *Science*, *336*(6080), 455-458

Eicker, A., E. Forootan, A. Springer, L. Longuevergne, and J. Kusche (2016), Does GRACE see the terrestrial water cycle 'intensifying'?, *J. Geophys. Res. (Atmospheres)*, *121*, 733–745.

Flechtner, F., K.-H.Neumayer, C. Dahle, H. Dobslaw, E. Fagiolini, J.-C. Raimondo, and A. Güntner (2016), What can be expected from the GRACE-FO laser ranging interferometer for Earth science applications?, *Surv. Geophys.*, *37*(2), 453-470

Folland, C., T. Karl, J. Christy, R. Clarke, G. Gruza, J. Jouzel, M. Mann, J. Oerlemans, M. Salinger, and S.-W. Wang, Observed climate variability and change, In: Houghton, J. et al. (eds.), Climate Change 2001: The Scientific Basis, *Cambridge Univ. Press*, 99-181

Forootan, E., R. Rietbroek, J. Kusche, M. Sharifi, J. Awange, M. Schmidt, P. Omondi, and J. Famiglietti (2014), Separation of large-scale water storage patterns over Iran using GRACE, altimetry and hydrological data, *Remote Sens. Environ.*, *140*, 580-595

Houborg R., M. Rodell, B. Li, R. Reichle, and B. Zaitchik, Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations, *Water Resour. Res.*, *48*, W07525

Humphrey, V., L. Gudmundsson, and S. Seneviratne (2016), Assessing global water storage variability from GRACE: Trends, seasonal cycle, subseasonal anomalies and extremes, *Surv. Geophys.*, *37*, 357-395

Huntington, T. (2006), Evidence for intensification of the global water cycle: Review and synthesis, *J. Hydrol.*, *319*, 83-95

Hosking, J.R.M. (1985), Maximum likelihood estimation of the parameters of the generalized extreme value distribution, *Appl. Stat.*, *34*, 301-310

Landerer, F., and S. Swenson (2012), Accuracy of scaled terrestrial water storage estimates, *Water Resour. Res.*, *48*, W04531

Martins, E.S., and J.R. Stedinger (2000), Generalized maximum-likelihood generalized extreme value quantile estimators for hydrologic data, *Water Resour. Res.*, *36*(3), 737–744.

Moore, P., and S. Williams (2014), Integration of altimetric lake levels and GRACE gravimetry over Africa: Inferences for terrestrial water storage change 2003-2011, *Water Resour. Res.*, *50*(12), 9696-9720

Phillips, T., R. Nerem, B. Fox-Kemper, J. Famiglietti, and B. Rajagopalan (2012), The influence of ENSO on global terrestrial water storage using GRACE. *Geophys. Res. Lett.*, *39*(16), L16705.

Rietbroek, R., S.-E. Brunnabend, J. Kusche, J. Schröter, and C. Dahle (2016), Revisiting the contemporary sea level budget on global and regional scales, *Proc. Nat. Acad. Sci. U.S.A.*, *113*(6), 15041509.

Seoane, L., G. Ramillien, F. Frappart, and H. Leblanc (2013), Regional GRACE-based estimates of water mass variations over Australia: validation and interpretation, *Hydrol. Earth Syst. Sci.*, *17*, 4925-4939.

Springer, A., J. Kusche, K. Hartung, C. Ohlwein, and L. Longuevergne (2014), New estimates of variations in water flux and storage over Europe based on regional (Re) analyses and multisensor observations, *J. Hydrometeorol.*, *15*(6), 23972417.

Stedinger, J., R. Vogel, and E. Foufula-Georgiou (1992), Frequency analysis of extreme events, In: Maidment, R. (ed.) Handbook of Hydrology *McGraw-Hill*, New York, NY

Ogawa, R., B. F. Chao, and K. Heki (2011), Acceleration signal in GRACE time-variable gravity in relation to interannual hydrological changes, *Geophys. J. Int.*, *184*(2), 673679.

Velicogna, I., T. Sutterley, and M. van den Broeke (2014), Regional acceleration in ice mass loss from Greenland and Antarctica using GRACE time-variable gravity data, *Geophys. Res. Lett.*, *41*(22), 8130-8137.

Table 1. Differences of one-in-five year levels from ERA-Interim and simulated GRACE/GRACE-FO derived TWF with respect to ERA-Interim, 1983.0-2015.0 (latitude-weighted land-only RMS, m/mon)

	years	ERA-I vs. 1983.0-2015.0	GRACE/GRACE-FO vs. ERA-I 1983.0-2015.0	GRACE/GRACE-FO vs. ERA-I, same timeframe
2007.0-2015.0	8.0	0.012	0.015	0.011
2003.0-2015.0	12.0	0.009	0.012	0.009
1999.0-2015.0	16.0	0.008	0.011	0.009
1991.0-2015.0	24.0	0.004	0.008	0.008
1983.0-2015.0	32.0	-	0.007	0.007

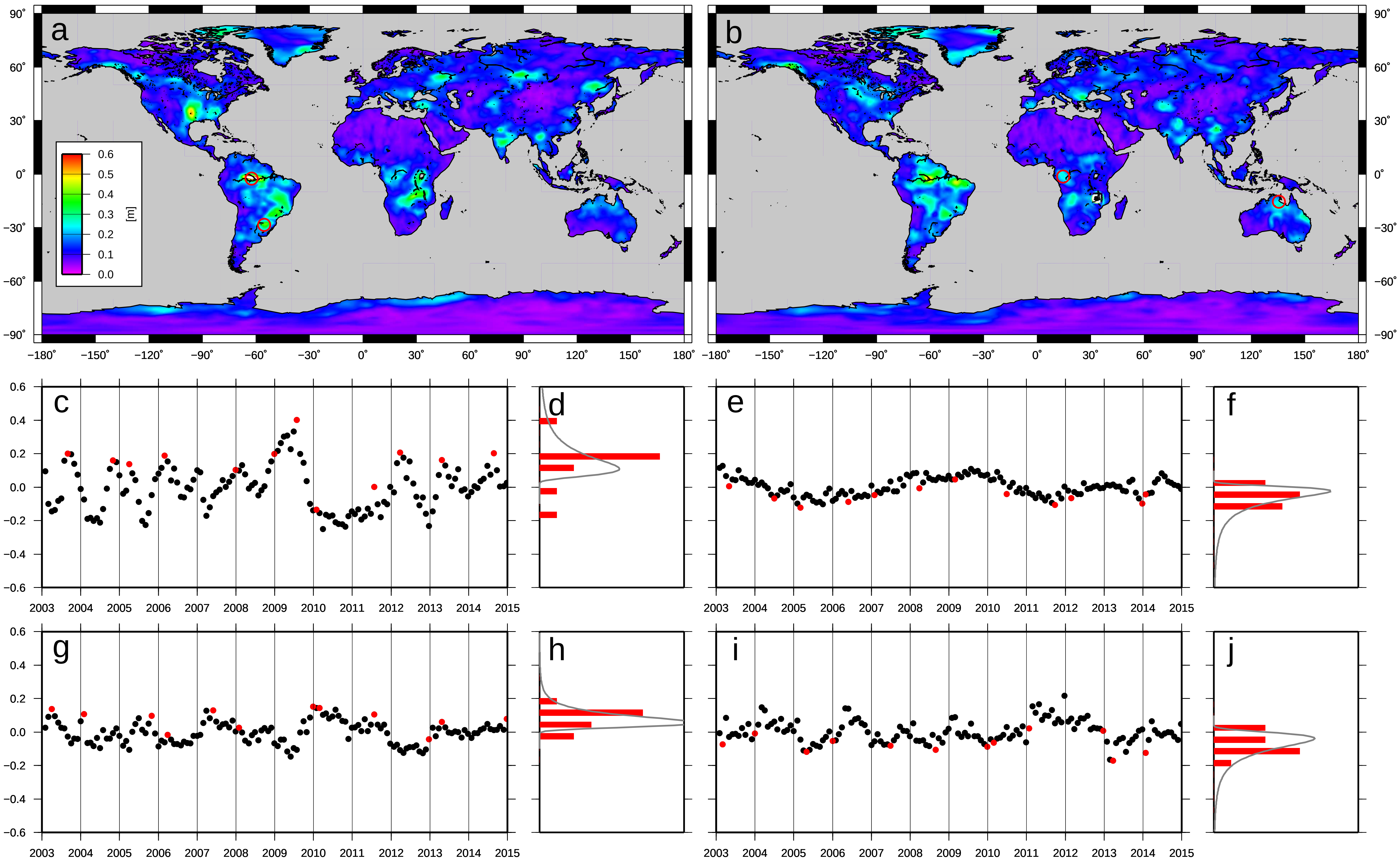
Yoon, J.-H., S. Wang, R. Gilles, B. Kravitz, L. Hipps, and P. Rasch (2015), Increasing water cycle extremes in California and in relation to ENSO cycle under global warming, *Nat. Commun.*, *6:8657*, doi:10.1038/ncomms9657.

Figure 1. (a) One-in-five-year levels of anomalously high total water storage (TWS) with respect to climatology, as observed by the GRACE satellite mission (2003.0-2015.0), (b) one-in-five-year levels of anomalously low TWS from GRACE, (c, e, g and j) TWS time series for locations indicated by red circles) (d and h) corresponding empirical and fitted probability density functions for anomalously high TWS, (f and j) corresponding empirical and fitted probability density functions for anomalously low TWS. c and d: Central Amazon, g and h: Parana, e and f: Cuvelai-Etosha, i and j: Northern Australia

Figure 2. (a) One-in-five-year levels of anomalously high total water flux (TWF) as observed by the GRACE satellite mission (2003.0-2015.0), (b) one-in-five-year levels of anomalously low TWF from GRACE, (c) one-in-five-year levels of anomalously high TWF from ERA-Interim reanalysis, (d) one-in-five-year levels of anomalously low TWF from ERA-Interim reanalysis

Figure 3. (a) One-in-five-year levels of anomalously high TWF from ERA-Interim, 2003.0-2015.0, (b) same, from simulated GRACE data (ERA-Interim plus GRACE correlated noise model), 2003.0-2015.0, (c) one-in-five- year levels of anomalously high TWF from ERA-Interim, 1991.0-2015.0, (d) same, from simulated GRACE data (ERA-Interim plus GRACE correlated noise model), 1991.0-2015.0, (e) one-in-five- year levels of anomalously high TWF from ERA-Interim, 1983.0-2015.0, (f) same, from simulated GRACE data (ERA-Interim plus GRACE correlated noise model), 1983.0-2015.0

Figure 1. Figure



a
b
c
d
e
f
g
h
i
j
[m]
0.0
0.1
0.2
0.3
0.4
0.5
0.6
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015

Figure 2. Figure

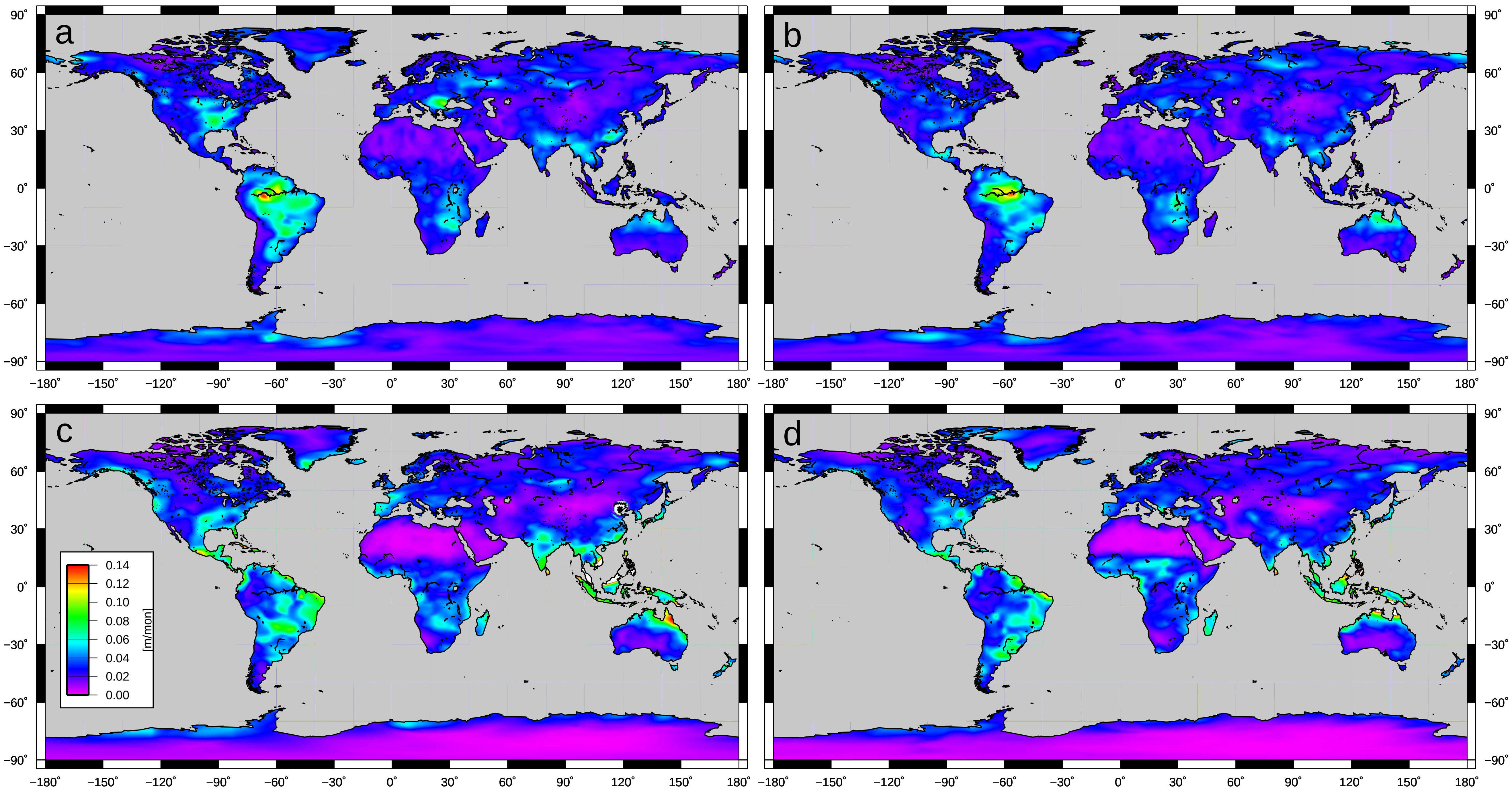
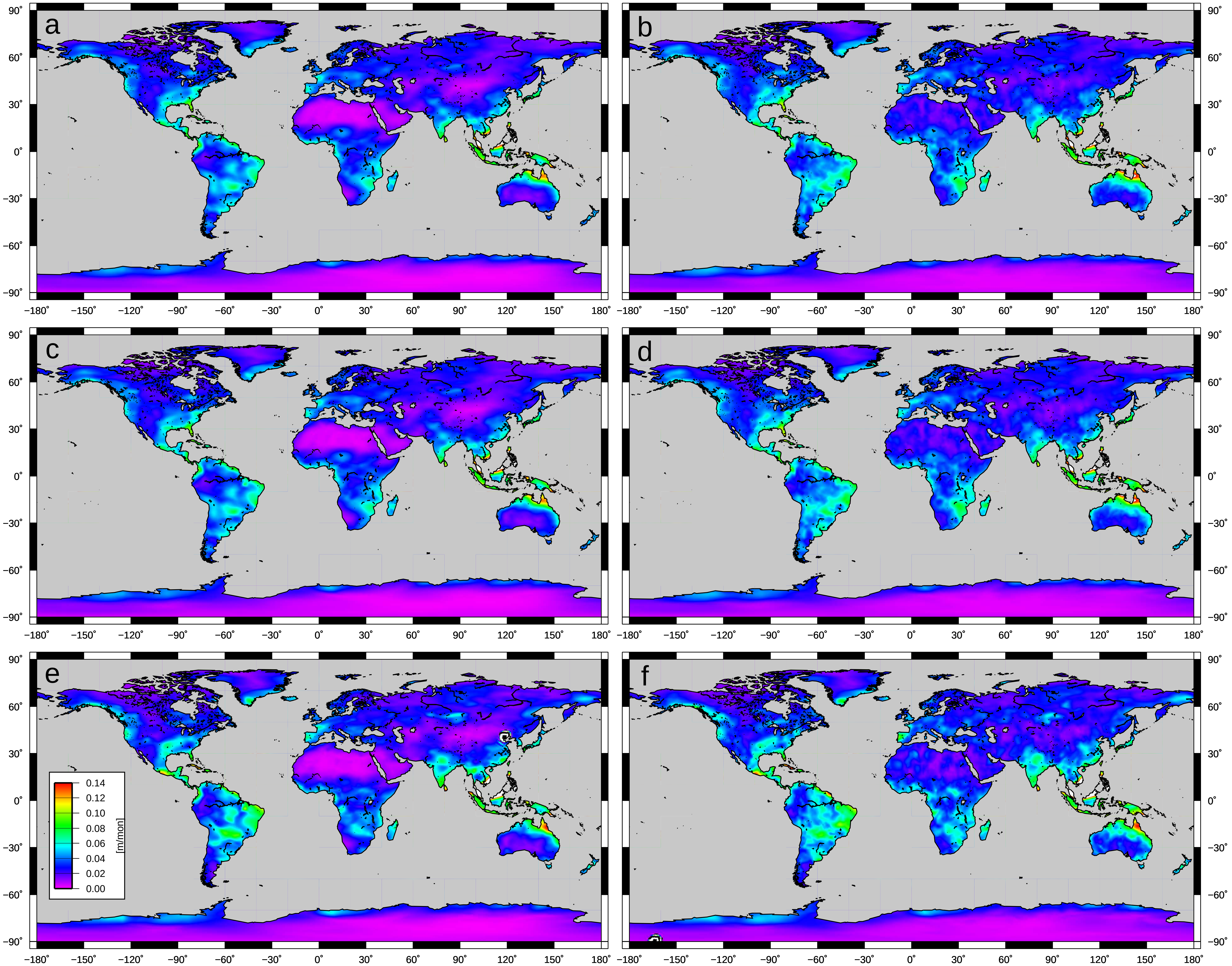

a
b
c
d
90°
60°
30°
0°
-30°
-60°
-90°
-180°
-150°
-120°
-90°
-60°
-30°
0°
30°
60°
90°
120°
150°
180°
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
[m/mon]

Figure 3. Figure


a
b
c
d
e
f
90°
60°
30°
0°
-30°
-60°
-90°
-180°
-150°
-120°
-90°
-60°
-30°
0°
30°
60°
90°
120°
150°
180°
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
[m/mon]