An Independent Component Analysis filtering approach for estimating continental hydrology in the GRACE gravity data

Frédéric Frappart (1), Guillaume Ramillien (2), Marc Leblanc (3), Sarah O. Tweed (3), Marie-Paule Bonnet (1), Philippe Maisongrande (4)

(1) Université de Toulouse, UPS, OMP, LMTG, 14 Avenue Edouard Belin, 31400 Toulouse, France (frederic.frappart@lmtg.obs-mip.fr, marie-paule.bonnet@ird.fr)

(2) Université de Toulouse, UPS, OMP, DTP, 14 Avenue Edouard Belin, 31400 Toulouse, France (guillaume.ramillien@dtp.obs-mip.fr)

(3) Hydrological Sciences Research Unit, School of Earth and Environmental Sciences, James Cook University, Cairns, Queensland, Australia (marc.leblanc@jcu.edu.au, sarah.tweed@jcu.edu.au)

(4) Université de Toulouse, UPS, OMP, LEGOS, 14 Avenue Edouard Belin, 31400 Toulouse, France (philippe.maisongrande@cnes.fr)

Submitted to Remote Sensing of Environment in August 2010

Abstract:

An approach based on Independent Component Analysis (ICA) has been applied on a combination of monthly GRACE satellite solutions computed from official providers (CSR, JPL and GFZ), to separate useful geophysical signals from important striping undulations. We pre-filtered the raw GRACE Level-2 solutions using Gaussian filters of 300, 400, 500-km of radius to verify the non-gaussianity condition which is necessary to apply the ICA. This linear inverse approach ensures to separate components of the observed gravity field which are statistically independent. The most energetic component found by ICA corresponds mainly to the contribution of continental water mass change. Series of ICA-estimated global maps of continental water storage have been produced over 08/2002-07/2009. Our ICA estimates were compared with the solutions obtained using other post-processings of GRACE Level-2 data, such as destriping and Gaussian filtering, at global and basin scales. Besides, they have been validated with *in situ* measurements in the Murray Darling Basin. Our computed ICA grids are consistent with the different approaches. Moreover, the ICA-derived time-series of water masses showed less north-south spurious gravity signals and improved filtering of unrealistic hydrological features at the basin-scale compared with solutions obtained using other filtering methods.

1. Introduction

Continental water storage is a key component of global hydrological cycles and plays a major role in the Earth's climate system via controls over water, energy and biogeochemical fluxes. In spite of its importance, the total continental water storage is not well-known at regional and global scales because of the lack of *in situ* observations and systematic monitoring of the groundwaters (Alsdorf and Lettenmaier, 2003).

The Gravity Recovery and Climate Experiment (GRACE) mission provides a global mapping of the time-variations of the gravity field at an unprecedented resolution of ~400 km and a precision of ~1 cm in terms of geoid height. Tiny variations of gravity are mainly due to redistribution of mass inside the fluid envelops of the Earth (i.e., atmosphere, oceans and continental water storage) from monthly to decade timescales (Tapley et al., 2004).

Pre-processing of GRACE data is made by several providers (University of Texas, Centre for Space Researh - CSR, Jet Propulsion Laboratory - JPL, GeoForschungsZentrum - GFZ and Groupe de Recherche en Géodésie Spatiale - GRGS) which produce residual GRACE spherical harmonic solutions that mainly represent continental hydrology as they are corrected from known mass transfers using *ad hoc* oceanic models (*i.e.*, Toulouse Unstructured Grid Ocean model 2D - T-UGOm 2D) and atmospheric reanalyses from National Centers for Environmental Prediction (NCEP) and European Centre for Medium Weather Forecasting (ECMWF). Unfortunately these solutions suffer from the presence of important north-south striping due to orbit resonance in spherical harmonics determination and aliasing of short-time phenomena which are geophysically unrealistic.

Since its launch in March 2002, the GRACE terrestrial water storage anomalies have been increasingly used for large-scale hydrological applications (see Ramillien et al., 2008; Schmitt et al., 2008 for reviews). They demonstrated a great potential to monitor extreme hydrological events (Andersen et al., 2005; Seitz et al., 2008; Chen et al., 2009), to estimate water storage variations in the soil (Frappart et al., 2008), the aquifers (Rodell et al., 2007; Strassberg et al., 2007; Leblanc et al., 2009) and the snowpack (Frappart et al., 2006; in press), and hydrological fluxes, such as basin-scale evapotranspiration (Rodell et al., 2004a; Ramillien et al. , 2006a) and discharge (Syed et al. , 2009).

Because of this problem of striping that limits geophysical interpretation, different post-processing approaches for filtering GRACE geoid solutions have been proposed to extract useful geophysical signals (see Ramillien et al., 2008; Schmidt et al., 2008 for reviews). These include the classical isotropic Gaussian filter (Jekeli, 1981), various optimal filtering

decorrelation of GRACE errors (Han et al., 2005; Seo and Wilson, 2005; Swenson and Wahr, 2006; Sasgen et al., 2006; Kusche, 2007; Klees et al., 2008), as well as statistical constraints on the time evolution of GRACE coefficients (Davis et al., 2008) or from global hydrology models (Ramillien et al., 2005). However, these filtering techniques remain imperfect as they require input non-objective *a priori* information which are most of the time simply tuned by hand (*e.g.* choosing the cutting wavelength while using the Gaussian filtering) or based on other rules-of-thumb.

We propose another post-processing approach of the Level-2 GRACE solutions by considering completely objective constraints, so that the gravity component of the observed signals is forced to be uncorrelated numerically using an Independent Component Analysis (ICA) technique. This approach does not require *a priori* information except the assumption of statistical independence of the elementary signals that compose the observations, *i.e.*, geophysical and spurious noise. The efficiency of ICA to separate gravity signals and noise from combined GRACE solutions has previously been demonstrated on one month of Level-2 solutions (Frappart et al., 2010). In this paper, we use this new statistical linear method to derive complete time series of continental water mass change.

The first part of this article presents the datasets used in this study: the monthly GRACE solutions to be inverted by ICA and to be used for comparisons, and the *in situ* data used for validation of our estimates over the Murray Darling drainage basin (~ 1 million of km²). This region has been selected for validation because of available dense hydrological observations. The second part outlines the three steps of the ICA methodology. Then the third and fourth parts present results and comparisons with other post-processed GRACE solutions at global and regional scales, and *in situ* measurements for the Murray Darling Basin respectively. Error balance of the ICA-based solutions is also made by considering the effect of spectrum truncation, leakage and formal uncertainties.

2. Datasets

2.1. The GRACE data

The GRACE mission, sponsored by National Aeronautics and Space Administration (NASA) and Deutsches Zentrum für Luft- und Raumfahrt (DLR), has been collecting data since mid-2002. Monthly gravity models are determined from the analysis of GRACE orbit perturbations in terms of Stokes spherical harmonic coefficients, *i.e.*, geopotential or geoid heights. The geoid is an equipotential surface of the gravity field that coincide with mean sea

level. For the very first time, monthly global maps of the gravity time-variations can be derived from GRACE measurements, and hence, to estimate the distribution of the change of mass in the Earth's system.

2.1.1. The Level-2 raw solutions

The Level-2 raw data consist of monthly estimates of geo-potential coefficients adjusted for each 30-day period from raw along-track GRACE measurements by different research groups (i.e., CSR, GFZ and JPL). These coefficients are developed up to a degree 60 (or spatial resolution of 333 km) and corrected for oceanic and atmospheric effects (Bettadpur, 2007) to obtain residual global grids of ocean and land signals corrupted by a strong noise. These data are available at: ftp://podaac.jpl.nasa.gov/grace/.

2.1.2. The destriped and smoothed solutions

The monthly raw solutions (RL04) from CSR, GFZ, and JPL were destriped and smoothed by Chambers (2006) for hydrological purposes. These three datasets are available for several averaging radii (0, 300 and 500 km on the continents and 300, 500 and 750 km on the oceans) at ftp://podaac.jpl.nasa.gov/tellus/grace/monthly.

In this study, we used the Level-2 RL04 raw data from CSR, GFZ and JPL, that we filtered with a Gaussian filter for radii of 300, 400 and 500 km, and the destriped and smoothed solutions for the averaging radii of 300 and 500 km over land.

2.2. The hydrological data for the Murray Darling Basin

In the predominantly semiarid Murray Darling Basin, most of the surface water is regulated using a network of reservoirs, lakes and weirs (Kirby et al., 2006) and the surface water stored in these systems represent most of the total surface water (SW) present across the basin. A daily time series of the total surface water storage in the network of reservoirs, lakes, weirs and in-channel storage was obtained from the Murray-Darling Basin Commission and the state governments from January 2000 to December 2008.

In the Murray Darling Basin, we derived monthly soil moisture (SM) storage values for the basin from January 2000 to December 2008 from the NOAH land surface model (Ek et al.,

153 2003), with the NOAH simulations being driven (parameterization and forcing) by the Global
154 Land Data Assimilation System (Rodell et al., 2004b). The NOAH model simulates surface
155 energy and water fluxes/budgets (including soil moisture) in response to near-surface
156 atmospheric forcing and depending on surface conditions (*e.g.*, vegetation state, soil texture
157 and slope) (Ek et al., 2003). The NOAH model outputs of soil moisture estimates have a 1°
158 spatial resolution and, using four soil layers, are representative of the top 2 m of the soil.
159 *In situ* estimates of annual changes in the total groundwater storage (GW) across the drainage
160 basin were obtained from an analysis of groundwater levels observed in government
161 monitoring bores from 2000 to 2008. Compared to earlier estimates by Leblanc et al. (2009),
162 the groundwater estimates presented in this paper provide an update of the *in situ* water level
163 and a refinement of the distribution of the aquifers storage capacity.
164 Assuming that (1) the shallow aquifers across the Murray-Darling drainage basin are
165 hydraulically connected and that (2) at a large scale the fractured aquifers can be assimilated
166 to a porous media, changes in groundwater storage across the area can be estimated from
167 observations of groundwater levels (*e.g.*, Rodell et al., 2007; Strassberg et al., 2007).
168 Variations in groundwater storage (ΔS_{GW}) were estimated from *in situ* measurements as:

169 $$\Delta S_{GW} = S_y \Delta H \qquad (1)$$

170 where S_y is the aquifer specific yield (%) and H is the groundwater level (L^{-1}) observed in
171 monitoring bores. Groundwater level data (H) were sourced from State Government
172 departments that are part of the Murray-Darling Basin (QLD; Natural Resources and Mines;
173 NSW; Department of Water and Energy; VIC; Department of Sustainability and
174 Environment; and SA; Department of Water Land and Biodiversity Conservation). Only
175 government observation bores (production bores excluded) with an average saturated zone
176 <50 m from the bottom of the screened interval were selected. Deeper bores were excluded as
177 they can reflect processes occurring on longer time scales (Fetter, 2001). A total of 6183
178 representative bores for the unconfined aquifers across the Murray Darling Basin were
179 selected on the basis of construction and monitoring details obtained from the State
180 departments. ~85% (5075) of the selected monitoring bores have a maximum annual standard
181 deviation of the groundwater levels below 2 m for the study period (2000-2008) and were
182 used to analyze the annual changes in groundwater storage during the period 2000 to 2008.
183 The remaining 15% of the observation bores, with the highest annual standard deviation, were
184 discarded as possibly under the immediate influence of local pumping or irrigation. The
185 potential influence of irrigation on some of the groundwater data is limited because during

this period of drought, irrigation is substantially reduced across the basin. Changes in groundwater levels across the basin were estimated using an annual time step as most monitoring bores have limited groundwater level measurements in any year (50% of bores with 5 measures per year). The annual median of the groundwater level was first calculated for each bore and change at a bore was computed as the difference of annual median groundwater level between two consecutive years. For each year, a spatial interpolation of the groundwater level change was performed across the basin using a kriging technique. Spatial averages of annual groundwater level change were computed for each aquifer group.

The Murray Darling drainage basin comprises several unconfined aquifers that can be regrouped into 3 categories according to their lithology: a clayey sand aquifer group (including the aquifers Narrabri (part of), Cowra, Shepparton, and Murrumbidgee (part of)); a sandy clay aquifer group (including the aquifers Narrabri (part of), Parilla, Far west, and Calivil (part of)); and a fractured rock aquifer group comprising metasediments, volcanics and weathered granite (including the aquifers Murrumbidgee (part of), North central, North east, Central west, Barwon, and Queensland boundary). The specific yield is estimated to range from 5 to 10% for the clayey sand unconfined aquifer group (Macumber, 1999; Cresswell et al., 2003; Hekmeijer and Dawes, 2003a; CSIRO, 2008); from 10 to 15% for the shallow sandy clay unconfined aquifer group (Macumber, 1999; Urbano et al., 2004); and from 1 to 10% for the fractured rock aquifer group (Cresswell et al., 2003; Hekmeijer and Dawes, 2003b; Smitt et al., 2003; Petheram et al., 2003). *In situ* estimates of changes in GW storage are calculated using the spatially averaged change in annual groundwater level across each type of unconfined aquifer group and the mean value of the specific yield for that group using (Eq. 1); while the range of possible values for the specific yield was used to estimate the uncertainty.

Groundwater changes in the deep, confined aquifers (mostly GAB and Renmark aquifers) are either due to: 1) a change in groundwater recharge at the unconfined outcrop; 2) shallow pumping at the unconfined outcrop or 3) deep pumping in confined areas for farming (irrigation and cattle industry). GRACE TWS estimates accounts for all possible sources of influence, while GW *in situ* estimates only include those occurring across the outcrop. Total pumping from the deep, confined aquifers was estimated to amount to -0.42 $km^3 .yr^{-1}$ in 2000 (Ife and Skelt, 2004), while groundwater pumping across the basin was -1.6 km^3 in 2002–2003 (Kirby et al., 2006). To allow direct comparison between TWS and in-situ GW estimates, pumping from the deep aquifers was added to the in-situ GW time series assuming the -0.42 $km^3 .yr^{-1}$ pumping rate remained constant during the study period.

3. Methodology

3.1. ICA-based filter

ICA is a powerful method for separating a multivariate signal into subcomponents assuming their mutual statistical independence (Comon, 1994; De Lathauwer et al., 2000). It is commonly used for blind signals separation and has various practical applications (Hyvärinen and Oja, 2000), including telecommunications (Ritsaniemi and Joutsensalo, 1999; Cristescu et al., 2000), medical signal processing (Vigário, 1997; van Hateren and van der Schaaf, 1998), speech signal processing (Stone, 2004), and electrical engineering (Gelle et al., 2001; Pöyhönen et al., 2003).

Assuming that an observation vector y collected from N sensors is the combination of P ($N \geq P$) independent sources represented by the source vector x, the following linear statistical model can be considered:

$$y = Mx \qquad (2)$$

where M is the mixing matrix whose elements m_{ij} ($1 \leq i \leq N$, $1 \leq j \leq P$) indicate to what extent the j^{th} source contribute to the i^{th} observation. The columns $\{m_j\}$ are the mixing vectors.

The goal of ICA is to estimate the mixing matrix M and/or the corresponding realizations of the source vector x, only knowing the realizations of the observation vector y, under the assumptions (De Lathauwer et al., 2000):

1) the mixing vectors are linearly independent,
2) the sources are statistically independent.

The original sources x can be simply recovered by multiplying the observed signals y with the inverse of the mixing matrix also known as the "unmixing" matrix:

$$x = M^{-1}y \qquad (3)$$

To retrieve the original source signals, at least N observations are necessary if N sources are present. ICA remains applicable for square or over-determined problems. ICA proceeds by maximizing the statistical independence of the estimated components. As a condition of applicability of the method, non-Gaussianity of the input signals has to be checked. The central limit theorem is then used for measuring the statistical independence of the components. Classical algorithms for ICA use centering and whitening based on eigenvalue decomposition (EVD) and reduction of dimension as main processing steps. Whitening ensures that the input observations are equally treated before dimension reduction.

ICA consists of three numerical steps. The first step of ICA is to centre the observed vector, i.e., to substract the mean vector $m = E\{y\}$ to make y a zero mean variable. The second step consists in whitening the vector y to remove any correlation between the components of the observed vector. In other words, the components of the white vector $\tilde{y}$ have to be uncorrelated and their variances equal to unity. Letting $C = E\{yy^t\}$ be the correlation matrix of the input data, we define a linear transform B that verifies the two following conditions:

$$\tilde{y} = By \qquad (4)$$

and:

$$E\{\tilde{y}\tilde{y}^t\} = I_P \qquad (5)$$

where I_P is identity matrix of dimension P x P.

This is easily accomplished by considering:

$$B = C^{-\frac{1}{2}} \qquad (6)$$

The whitening is obtained using an EVD of the covariance matrix C:

$$C = EDE^t \qquad (7)$$

where E is the orthogonal matrix of the eigenvectors of C and D is the diagonal matrix of its eigenvalues. D=diag(d_1,…,d_P) as a reduction of the dimension of the data to the number of independent components (IC) P is performed, discarding the too small eigenvalues.

For the third step, an orthogonal transformation of the whitened signals is used to find the separated sources by rotation of the joint density. The appropriate rotation is obtained by maximizing the non-normality of the marginal densities, since a linear mixture of independent random variables is necessarily more Gaussian than the original components.

Many algorithms of different complexities have been developed for ICA (Stone, 2004). The *FastICA* algorithm, a computationally highly efficient method for performing the estimation of ICA (Hyvärinen and Oja, 2000) has been considered to separate satellite gravity signals. It uses a fixed-point iteration scheme that has been found to be 10 to 100 times faster than conventional gradient methods for ICA (Hyvärinen, 1999).

We used the *FastICA* algorithm (available at http://www.cis.hut.fi/projects/ica/fastica/) to unravel the IC of the monthly gravity field anomaly in the Level-2 GRACE products. We previously demonstrated, on a synthetic case, that land and ocean mass anomalies are statistically independent from the north-south stripes using information from land and ocean models and simulated noise (Frappart et al., 2010). Considering that the GRACE Level-2 products from CSR, GFZ and JPL are different observations of the same monthly gravity

anomaly and, that the land hydrology and the north-south stripes are the independent sources, we applied this methodology to the complete 2002-2009 time series. The raw Level-2 GRACE solutions present Gaussian histograms which prevent the successful application of the ICA method. To ensure the non-Gaussianity of the observations, the raw data have been preprocessed using Gaussian filters with averaging radii of 300, 400 and 500 km as in Frappart et al. (2010).

3.2. Time-series of basin-scale total water storage average

For a given month *t*, the regional average of land water volume $\delta V(t)$ (or height $\delta h(t)$) over a given river basin of area *A* is simply computed from the water height δh_j, with *j=1, 2, ...* (expressed in terms of mm of equivalent-water height) inside *A*, and the elementary surface $R_e^2\, \delta\lambda\, \delta\theta \sin\theta_j$:

$$\delta V(t) = R_e^{\,2} \sum_{j \in A} \delta h_j(\theta_j, \lambda_j, t) \sin\theta_j \delta\lambda \delta\theta \qquad (8)$$

$$\delta h(t) = \frac{R_e^{\,2}}{A} \sum_{j \in A} \delta h_j(\theta_j, \lambda_j, t) \sin\theta_j \delta\lambda \delta\theta \qquad (9)$$

where θ_j and λ_j are co-latitude and longitude of the jth point, $\delta\lambda$ and $\delta\theta$ are the grid steps in longitude and latitude respectively (generally $\delta\lambda=\delta\theta$). In practice, all points of *A* used in (Eq. 8 and 9) are extracted for the eleven drainage basins masks at a 0.5° resolution provided by Oki and Sud (1998), except for the Murray Darling Basin where we used basin limits from Leblanc et al. (2009).

3.3. Regional estimates of formal error

As ICA provides separated solutions which have Gaussian distributions, the variance of the regional average for a given basin is:

$$\sigma^2_{formal} = \frac{\sum_{k=1}^{L} \sigma_k^2}{L^2} \qquad (10)$$

where σ_{formal} is the regional formal error, σ_k is the formal error at a grid point number *k*, and *L* is the number of points used in the regional averaging.

If the points inside the considered basin are independent, this relation is slightly simplified:

$$\sigma_{formal} = \frac{\sigma_k}{\sqrt{L}} \qquad (11)$$

3.4. Frequency cut-off error estimates

Error in frequency cut-off represents the loss of energy in the short spatial wavelength due to the low-pass harmonic decomposition of the signals that is stopped at the maximum degree N_1. For the GRACE solution separated by ICA; N_1=60, thus the spatial resolution is limited and stopped at ~330 km by construction. This error is simply evaluated by considering the difference of reconstructing the remaining spectrum between two cutting harmonic degrees N_1 and N_2, where $N_2 > N_1$ and N_2 should be large enough compared to N_1 (*e.g.*, N_2=300 in study):

$$\sigma_{truncation} = \sum_{n=0}^{N_2} \xi_n - \sum_{n=0}^{N_1} \xi_n = \sum_{n=N_1}^{N_2} \xi_n \qquad (12)$$

using the scalar product $$\xi_n = \sum_{m=0}^{n} \left(C_{nm} A_{nm} + S_{nm} B_{nm} \right) \qquad (13)$$

where A_{nm} and B_{nm} are the harmonic coefficients of the considered geographical mask, and C_{nm} and S_{nm} are the harmonic coefficients of the water masses.

3.5. Leakage error estimates

We define « leakage » as the portion of signals from outside the considered geographical region that pollutes the region's estimates. By construction, this effect can be seen as the limitation of the geoid signals degree in the spherical harmonics representation. For each basin and at each period of time, leakage is simply computed as the average of outside values by using an « inverse » mask, which is 0 and 1 in and out of the region respectively, developed in spherical harmonics and then truncated at degree 60. This method of computing leakage of continental water mass has been previously proposed for the entire continent of Antarctica (Ramillien et al., 2006b), which revealed that the seasonal amplitude of this type of error can be quite important (*e.g.* up to 10% of the geophysical signals). In case of no leakage, this average should be zero (at least, it decreases with the maximum degree of decomposition). However, the maximum leakage of continental hydrology remains in the order of the signals magnitude itself.

4. Results and discussion

4.1. ICA-filtered land water solutions

The methodology presented in Frappart et al. (2010) has been applied to the Level-2 RL04 raw monthly GRACE solutions from CSR, GFZ and JPL, preprocessed using a Gaussian filter with a radius of 300, 400 and 500 km, over the period July 2002 to July 2009. The results of this filtering method is presented in Fig. 1 for four different time periods (March and

353 September 2006, March 2007 and March 2008) using the GFZ solutions Gaussian-filtered
354 with a radius of 400 km. Only the ICA-based GFZ solution is presented since, for a specific
355 radius, the ICA-based CSR, GFZ, and JPL solutions only differ from a scaling factor for each
356 specific component. The ICA-filtered CSR, GFZ, and JPL solutions are obtained by
357 multiplying the j^{th} IC with the j^{th} mixing vector (2). As the last two modes correspond to the
358 north–south stripes, we present their sum in Fig. 1.
359 The first component is clearly ascribed to terrestrial water storage with variations in the range
360 of ±450 mm of Equivalent Water Thickness (EWT) for an averaging radius of 400 km. The
361 larger water mass anomalies are observed in the tropical regions, *i.e.*, the Amazon, the Congo,
362 the Ganges and the Mekong Basins, and at high latitudes in the northern hemisphere. The
363 components 2 and 3 correspond to the north-south stripes due to resonances in the satellite's
364 orbits. They are smaller than the first component by a factor of 3 or 4 as previously found
365 (Frappart et al., 2010).
366 The FastICA algorithm was unable to retrieve realistic patterns and/or amplitudes of TWS-
367 derived from GRACE data preprocessed using a Gaussian filter with a radius 300 km for
368 several months (02/2003, 06 to 11/2004, 02/2005, 07/2005, 01/2006, 01/2007, 02/2009).
369 Some of these dates, such as the period between June and November 2004, correspond to
370 deep resonance between the satellites caused by an almost exact repeat of the orbit,
371 responsible for a significantly poorer accuracy of the monthly solutions (Chambers, 2006). As
372 ICA is based on the assumption of independence of the sources, if the sources exhibit similar
373 statistical distribution, the algorithm is unable to separate them.
374 A classical measure of the peakiness of the probability distribution is given by the kurtosis.
375 The kurtosis K_y is dimensionless fourth moment of a variable y and classically defined as:

376 $$K_y = \frac{E\{y^4\}}{E\{y^2\}^2} \qquad (14)$$

377 If the probability density function of y is purely Gaussian, its kurtosis has the numerical value
378 of 3. In the followings, we will consider the excess of kurtosis (K_y-3) and refer to the kurtosis
379 as it is commonly done. So a variable y will be Gaussian if its kurtosis remains close to 0.
380 The time series of the kurtosis of the sources separated using ICA are presented in Fig. 2 for
381 different radii of Gaussian filtering (300, 400 and 500 km) of GRACE mass anomalies. The
382 kurtosis of the sum of the 2nd and 3rd ICs, corresponding to the north-south stripes, is most of
383 the time, close to 0; that is to say that the meridian oriented spurious signals is almost
384 Gaussian. Almost equal values of the kurtosis for the 1st IC and the sum of the 2nd and 3rd ICs

can be observed for several months. Most of the time, they correspond to time steps where the algorithm is unable to retrieve realistic TWS (02/2003, 08/2004, 11/2004, 02/2005, 01/2006, 01/2007, 02/2009).

We also observed that the number of time steps with only one IC (the outputs are identical to the inputs, *i.e.*, no independent sources are identified and hence no filtering was performed) increases with the radius of the Gaussian filter (none at 300 km, 2 at 400 km, 7 at 500 km).

In the following, as the ICA-derived TWS with a Gaussian prefiltering of 300 km, exhibits an important gap of 6 months in 2004, we will only consider the solutions obtained after a preprocessing with a Gaussian filter for radii of 400 and 500 km (ICA400 and ICA500).

4.2. Global scale comparisons

Global scale comparisons have been achieved with commonly-used GRACE hydrology preprocessings: the Gaussian filter (Jekeli, 1981) and the destriping method (Swenson and Wahr, 2006) for several smoothing radii.

4.2.1. ICA versus Gaussian-filtered solutions

Advantages of extracting continental hydrology using ICA after a simple Gaussian filtering have to be demonstrated for the complete period of availability of the GRACE Level-2 dataset, as it was for one period of GRACE Level-2 data in Frappart et al., (2010). Numerical tests of comparisons before and after ICA have been made to show full utility of considering a post-treatment by ICA for signals separation. For the period 2003-2008, we present correlation (Fig. 3) and RMS (Fig. 4) maps computed between the ICA400 (respectively ICA500) and Gaussian-filtered solutions with a radius of 400 km (500 km), named in the followings G400 (G500). High correlation coefficients are generally observed over land (greater than 0.9), increasing with the smoothing radius, especially over areas with large hydrological signals, *i.e.*, Amazon, Congo and Ganges basins, boreal regions. On the contrary, low correlation values, structured as stripes, are located over arid and semi-arid regions (southwest of the US, Sahara, Saudi Arabia, Gobi desert, centre of Australia), especially for GFZ and JPL solutions. These important RMS differences between Gaussian and ICA-based solutions reveal that the GRACE signals still contains remaining stripes after the Gaussian filtering. This justifies that extracting the useful continental hydrology signals requires a further processing. For this purpose, ICA succeeds in isolating this noise in its second and third components (as illustrated in Figure 1).

In Fig. 4, we observe that the spatial distribution of the RMS between ICA solutions and Gaussian solutions presents north-south stripes with values generally lower than 30 mm, except for some spots between ±(20 – 30)° of latitudes on the GFZ (Fig. 4c and d) and JPL (Fig. 4e and e) solutions. The ICA approach allows the filtering of remnant stripes present in the Gaussian solutions, especially for GFZ and JPL (Fig. 4c to f). These unrealistic structures (stripes and spots), which correspond to resonances in the orbit of the satellites, are clearly filtered out using the ICA approach (compare with Fig. 1). A more important smoothing due to a larger radius caused a decrease of the RMS between ICA and Gaussian solutions (Fig 4a, c and e). The RMS can reach 100 mm between ICA400 and G400 and only 65 mm between ICA500 and G500 for the GFZ solutions.

4.2.2. ICA versus destriped and smoothed solutions

Similar to Fig. 3 and 4, we present correlation (Fig. 5) and RMS (Fig. 6) maps over the period 2003-2008; between ICA400 (ICA500 respectively) and destriped and smoothed solutions with radii of 300 km (DS300) and 500 km (DS500), made available by Chambers (2006). The correlation maps between ICA400 (respectively ICA500) and DS300 (DS500) exhibit very similar patterns compared with those presented in Fig. 3, except in the McKenzie Basin and Nunavut (northern Canada) where low correlation or negative correlations for GFZ and JPL solutions (Fig. 5c to f) are found.

The RMS differences between ICA and destriped and smoothed solutions present low values above 30° (or below -30°) of latitude (< 30 mm of EWT), with the exceptions of Nunavut in northern Canada and extreme values (up to 100 mm) in the tropics. This is clearly due to the low performance of the destriping method in areas close to Equator as previously mentioned by Swenson and Wahr (2006) and Klees et al. (2008). These extrema are especially present in the DS300 GFZ and JPL solutions. Some secondary maxima (up to 80 mm) are also noticed for the GFZ and the JPL solutions. These important differences correspond to north-south stripes that still appear in the destriped and smoothed solutions despite the filtering process (and that can be filtered out by applying an ICA approach – see the results for DS300 GFZ solution of March 2006 in Fig. 7). For an averaging radius of 500 km, the RMS between ICA500 and DS500 is lower than 20 to 30 mm, except for the Nunavut (Fig. 6b, d, f), along the Parana stream (40 to 50 mm in the CSR solutions – Fig. 6b), and along the Amazon and Parana streams (60 to 80 mm in the GFZ solutions – Fig. 6d).

These low spatial correlations suggest Gaussian-ICA provides at the least equivalent results on the continents to the smoothing-destriping method. Besides, it is interesting that both approaches are based on a pre-Gaussian filtering. Short wavelength differences between the maps obtained separetely using ICA and destriping reveals the limitation of the destriping which generates artefacts in the tropics.

4.2.3. Trend comparisons

The effects of the stripes on the trends estimated using GRACE-based TWS is supposed to be significant. From the time-series of TWS anomaly grids derived from GRACE (using equation 1), the temporal trend, seasonal and semi-annual amplitudes were simultaneously fitted by least-square adjustment at each grid point over the period 2003-2008 (see Frappart et al., in press, for details). We present in Fig. 8 trends of TWS for ICA400 (Fig. 8a) and ICA500 (Fig. 8b), for G400 (Fig. 8c) and G500 (Fig. 8d), and for DS300 (Fig. 8e) and DS500 (Fig. 8f).

The trend estimates exhibit large differences in spatial patterns and the amplitude of the signal, especially between ICA and other processing methods. The most significant differences over land (except Antarctica) are located at high latitudes, over Scandinavia, and the Laurentide region, in the northeast of Canada. The ICA400 and ICA500 solutions present negative trends of TWS (Fig. 8a and b), whereas the G400 and G500 and the DS300 and DS500 present large positive trends (Fig. 8c to f). These two zones are strongly affected by the post-glacial rebound (PGR) which has a specific signature in the observed gravity field. This effect accounts for positive trends in these regions. According to the PGR models developed by Peltier (2004) and Paulson et al. (2007), its intensity can be greater in these regions than the trends measured by GRACE. For example, in the Nelson Basin, Frappart et al. (in press) found a trend of TWS from GRACE of (4.5±0.2) mm/yr over 2003-2006, while model-based estimates of PGR represents 18.8 mm/yr in this region. As the space and time characteristics of the glacial isostatic adjustement (GIA) is different than the one from continental hydrology, the ICA approach may have separated it from the signals only related to the redistribution of water masses. Unfortunately, trends computed on the 2nd and 3rd ICs, and their sum does not exclusively correspond to GIA, but to a mixture of geophysical remaining signals and noise. It is also worth noticing that the gravity signature of the Sumatra event in December 2004, which is clearly apparent on the Gaussian-filtered solutions (Fig. 8c and d), is not visible in the 1st mode of the ICA solutions (Fig. 8a and b), but is present in the

sum of the 2nd and 3rd ICs. This confirms that the ICA is able to isolate a pure hydrological mode in the GRACE products.

A second important difference between ICA solutions and Gaussian-filtered and destriped and smoothed solutions concerns the impact of the filtering radius on the trends estimate. An increase of the filtering radius causes a smoothing of the solutions, and, consequently, a decrease of the intensity of the trends on Gaussian-filtered and destriped and smoothed solutions (Fig. 8c to f). On the contrary, the intensity of the trends increases with radius of prefiltering on the ICA solutions. The location of the extrema is also shifted. This change in the location is a side-effect of the prefiltering with the Gaussian filter. A better location of the trends is observed when a Gaussian-filter of 400 km of radius is used instead of 500 km (see for instance the trends pattern in the Amazon and Orinoco Basins in Fig. 8a and b).

4.3. Basin scale comparisons

Changes in total water volume were estimated for 27 drainage basins whose locations are shown in Fig. 9 (the corresponding areas are given in Table 1). ICA400 and ICA500 were used to compute regional TWS averages versus time (Eq. 8 and 9). These times-series were compared to the G400 and G500, and the DS300 and DS500 respectively. Examples of the Amazon, the Ob and the Mekong basins for ICA400, G400 and DS300 (GFZ solutions) are presented in Fig. 10. TWS exhibit very similar temporal patterns for all the types of filtering radii and basins. The correlation coefficients between the ICA and the Gaussian-filtered, or the ICA and the destriped and smoothed time series of TWS are greater than 0.9 for 21 out of 27 basins. For five other basins (Amur, Colorado, Hwang Ho, Parana, St Lawrence), most of the correlation coefficients are greater than 0.8 or 0.9, and the others (generally the correlation coefficients between ICA and destriped and smoothed GFZ solutions) greater than 0.65 or 0.7. The only exception is the McKenzie Basin where all the correlation coefficients between ICA and destriped and smoothed are lower than 0.75 ($r_{(ICA400,DS300)\ JPL}$=0.45 and $r_{(ICA500,DS500)\ JPL}$=0.51). For the GFZ solutions, some unrealistic peaks are present in the destriped and smoothed TWS time series for some periods (Fig. 10a, 10b and 10c). These peaks only appear on the Gaussian-filtered solutions for the smallest basins, such as the one of the Mekong river (Fig. 10c), but are not present in the ICA-filtered solutions. RMS difference between ICA-filtered solutions and the other type of solutions are generally lower than 30 mm of equivalent water height and logically decrease with the radius of filtering. Differences with ICA solutions are generally larger for the GFZ-based destriped solutions, especially in tropical

521 regions where the performances of the destriping are not the best and the hydrological signal
522 the largest (Fig. 11).
523 We present an analysis of possible sources of error on the computation of regional averages
524 versus time using the 300, 400, 500-km pre-filtered ICA solutions. This task is made on the
525 longest available period of time for each center (CSR, JPL, GFZ), and for the 27 drainage
526 basins (Table 1).
527 The formal error decreases with the number of points (see Eq. 10 and 11), the surface of the
528 considered region, and the value of error σ_k at each grid point. In the case of the Amazon
529 Basin (~6 millions of km²), the formal error is only of 4.5 mm when σ_k = 100 mm of
530 equivalent-water height. For the Dniepr river basin that represents the smallest surface of our
531 chosen basins (~0.52 million of km²), the formal error reaches 12.3 mm. For the series of
532 basins, the maximum values of formal errors on the regional average are in the range of 10-15
533 km^3 of water volume per hundred of mm of error on the gridded points.
534 To estimate the frequency cut-off error or "omission error", we made statistics of the
535 numerical tests were performed to see what maximum error can be reached using Eq. 12. We
536 computed this residual quantity for 300, 400 and 500 km-filtered ICA solutions and for each
537 hydrological basin for N_1=60 and N_2=300. The maximum error is always less than 1 km^3, as
538 shown previously (Ramillien et al., 2006a), and it decreases with the filtering wavelength of
539 the pre-processing. In other words, this error simply increases with the level of noise in the
540 data. For the 300 km and 400 km pre-filtered solutions, the maximum values are found for the
541 Amur River: 0.3 km^3 (CSR), 0.1 km^3 (JPL), and 0.8 km^3 (GFZ), and 0.04 km^3 (CSR), 0.09
542 km^3 (JPL), 0.08 km^3 (GFZ) respectively. While using the 500-km filtered solutions, the error
543 of truncation is less than 0.005 km^3.
544 The leakage error on the ICA solutions was computed per drainage basin (see section 3.5).
545 The results are presented in Fig. 12 for the different centres and radii of Gaussian prefiltering
546 of 300, 400 and 500 km. We observed that the leakage decreases with radius of filtering and
547 is generally lower in the JPL solutions, which presents lowest peak to peak amplitudes. This
548 leakage error is logically greater in areas where several basins with large hydrological signal
549 are close, *i.e.*, South America with the Amazon, the Parana, the Orinoco and the Tocantins, or
550 tropical Asia with the Ganges, the Brahmaputra and the Mekong rivers. The leakage error is
551 also greater for basins with small hydrological signals close to a predominant basin having
552 important hydrological variations, *i.e.*, the Okavango and the Zambezi with the Congo.
553
554 *4.4. Basin scale validation*

GRACE observations were used to estimate variations in TWS over the Murray Darling Basin (see Fig. 13 for the location of the basin) at an interannual scale from January 2003 to December 2008. The annual variations in TWS from GRACE for different types of solutions were compared to the annual TWS computed as the sum of *in situ* observations (SW+GW) and NOAH outputs (SM). In the Murray-Darling River basin, GLDAS-NOAH simulations of SM range from 5 to 29% (in volumetric water content) across the basin for the study period, and are within typical values for monthly means at 1° resolution (Lawrence and Hornberger, 2007). Besides, Leblanc et al. (2009) show that, between 2003 and 2007, the linear rate of water changes for the GRACE TWS time series is similar to that observed for the annual total water storage from *in situ* observations and modeling. The combined annual anomalies of surface water, groundwater and soil moisture are highly correlated with the annual GRACE TWS (R = 0.94 and mean absolute difference =13 km^3 for the 2003–2007 period). These latter interannual field-based data are considered as the reference in the following analysis. The results are presented on Fig. 14 for the GFZ solutions. The ICA solutions generally present a temporal pattern closer to the so-called reference (*i.e.*, the sum of *in situ* data for SW and GW and model outputs for SM). This temporal pattern and the associated dynamics do not change with the radius of filtering, which is not the case with the Gaussian filtered, and the destriped and smoothed solutions. None of the solutions are able to retrieve the minimum observed in 2007 in the reference dataset. It is important to notice that the largest part of the interannual variations of TWS comes from the GW reservoir. For this hydrological component, the water storage is derived from *in situ* measurements through Eq. (1), and is highly dependent on the specific yield coefficient; where averages of local measurements are used to determine ranges of regional-scale estimates. For the Murray Darling Basin, composed of several aquifers, this leads to a broad range of spatial variability for the GW estimates. This variability is around or greater than 20 km^3 for 2002, 2007 and 2008 and around 10 km^3 for 2003, 5 km^3 in 2004 and 2006, 2.5 km^3 in 2005. We present yearly deviations to the reference for the different considered solutions (Fig. 15) over 2003-2006, where the range of variability of the GW is lower. Generally, the absolute deviation of ICA-derived TWS to the reference is lower than 5 km^3, especially for solutions filtered at 400 km. At 500 km, the smoothing due to the preprocessing using a Gaussian filter is more important and explains the slight increase of the absolute deviation as the spatial resolution is degraded. The destriping method and the Gaussian filtering also exhibit good performances, even if

588 important deviations are sometimes observed (particularly with the Gaussian filter for a radius
589 of 400 km).
590
591 **5. Conclusion**
592
593 The ICA-based approach is a very efficient method for successfully separating TWS from
594 noise in the GRACE Level-2 data. We demonstrated that this method is more robust than
595 classical filtering methods, such as the Gaussian filtering or the destriping. Comparisons at a
596 global-scale showed that the ICA-based solutions present less north-south stripes than
597 Gaussian and destriped solutions on the land, and more realistic hydrological structures than
598 the destriped solutions in the tropics. Trend maps over 2003-2008 have also been computed.
599 The corresponding trend maps present more realistic trend patterns than those obtained with
600 other types of solutions (see for example over the Amazon and Orinoco Basins with the 400
601 km radius of prefiltering). ICA filtering seems to allow the separation of the GIA from the
602 TWS as negative trends were found over the Laurentides and Scandinavia. Unfortunately, this
603 important geophysical parameter does not appear clearly in an ICA mode yet. The major
604 drawback of this approach is that it can not directly be applied to the GRACE Level-2 raw
605 data, as a first step of prefiltering is required. In this study, we applied a Gaussian filtering
606 which deteriorates the location of the important water mass patterns. This aspect of the pre-
607 treatment to be improved and highlights the necessity to replace the Gaussian filter used for
608 preprocessing the GRACE Level-2 raw data by one more suited to improve the quality of the
609 GRACE-derived TWS and thus obtain trustworthy estimate of the trends.
610 At the basin-scale, the ICA-based solutions allowed us to filter out the unrealistic peaks
611 present in the time-series of TWS obtained using classical filtering for basins with areas lower
612 than one million km². Among the ICA-based solutions, the JPL solutions are less affected by
613 leakage compared with other solutions. JPL solutions also exhibit the lowest peak-to-peak
614 amplitudes. The error balance of the GRACE-derived TWS is dominated by the effect of the
615 leakage.
616 Validation with *in situ* measurements was performed in the Murray Darling Basin where
617 broad networks of *in situ* measurements of SW and GW are available. The ICA-based
618 solutions are in better agreement with *in situ* data compared with the other types of solutions,
619 especially where prefiltering at a 400 km of radius. The maximum deviations are lower by a
620 factor two or three compared with the other filtering methods.
621
622 **Acknowlegdements**

This work was partly supported by the fondation Sciences et Techniques pour l'Aéronautique et l'Espace (STAE) in the framework of the CYMENT Project (post-doctoral grant for Frédéric Frappart) and the Australian Research Council grant DP0666300 (Marc Leblanc and Sarah Tweed). The authors also wish to thank Dr. Matthew Rodell for his help during the revision process.

References

Alsdorf, D. E., & Lettenmaier, D.P. (2003). Tracking fresh water from space. *Science, 301,* 1492-1494.

Andersen, O. B., Seneviratne, S.I., Hinderer, J. & Viterbo, P. (2005). GRACE-derived terrestrial water storage depletion associated with the 2003 European heat wave, *Geophysical Research Letters, 32(18)*, L18405.

Bettadpur, S. (2007). CSR level-2 processing standards document for level-2 product release 0004, GRACE 327–742, Rev. 3.1.

Chambers, D.P. (2006). Evaluation of new GRACE time-variable gravity data over the ocean. *Geophysical Research Letters, 33*, LI7603, doi:10.1029/2006GL027296.

Chen, J. L., Wilson, C. R., Tapley, B. D., Yang, Z. L. & Niu G.Y. (2009). 2005 drought event in the Amazon River basin as measured by GRACE and estimated by climate models, *Journal of Geophysical Research, 114*, B05404, doi:10.1029/2008JB006056.

Comon, P. (1994). Independent component analysis: a new concept? *Signal Processing, 36*, 287-314.

Cresswell, R.G., Dawes, W.R., Summerell, G.K., & Walker G.R. (2003). Assessment of salinity management options for Kyeamba Creek, New South Wales: Data analysis and groundwater modelling. CSIRO Land and Water Technical Report 26/03.

Cristescu, R., Joutsensalo, J., & Ristaniemi, T. (200). Fading Channel Estimation by Mutual Information Minimization for Gaussian Stochastic Processes, *Proceedings of IEEE International Conference on Communications (ICC2000)*, New Orleans, USA, June 18-22, 2000, 56-59.

CSIRO (2008). Water availability in the Loddon-Avoca. A report to the Australian Government from the CSIRO Murray-Darling Basin Sustainable Yields Project. CSIRO, Australia. 123pp.

Davis, J.L., Tamisiea, M.E., Elósegui, P., Mitrovica, J.X., & Hill, E.M. (2008). A statistical filtering approach for Gravity Recovery and Climate Experiment (GRACE) gravity data. *Journal Geophysical Research, 113*, B04410, doi:10.1029/2007JB005043.

De Lathauwer, L., De Moor, B., & Vandewalle, J. (2000). An introduction to independent component analysis. *Journal of Chemometrics, 14*, 123-149.

671
672 Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G. &
673 Tarpley, J.D. (2003). Implementation of Noah land surface model advances in the National
674 Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.,*
675 *108(D22)*, 8851, doi:10.1029/2002JD003296.
676
677 Fetter, C. W. (2001). *Applied Hydrogeology*. 4th ed. Prentice-Hall, New Jersey. 598 p.
678
679 Frappart, F., Ramillien, G., Biancamaria, S., Mognard, N.M. & Cazenave A. (2006).
680 Evolution of high-latitude snow mass derived from the GRACE gravimetry mission (2002-
681 2004), *Geophysical Research Letters*, *33*, L02501.
682
683 Frappart, F., Papa, F., Famiglietti, J.S., Prigent, C., Rossow, W.B. & Seyler F. (2008).
684 Interannual variations of river water storage from a multiple satellite approach: A case study
685 for the Rio Negro River basin, *Journal of Geophysical Research, 113*, D21104,
686 doi:10.1029/2007JD009438.
687
688 Frappart, F., Ramillien, G., Maisongrande, P., & Bonnet, M-P. (2010). Denoising satellite
689 gravity signals by Independent Component Analysis. *IEEE Geosciences and Remote Sensing*
690 *Letters*, *7*(3), 421-425, doi:10.1109/LGRS.2009.2037837.
691
692 Frappart, F., Ramillien, G. & Famiglietti, J.S. (in press). Water balance of the Arctic drainage
693 system using GRACE gravimetry products. *International Journal of Remote Sensing,* doi:
694 10.1080/01431160903474954.
695
696 Gelle, G., Colas, M., & Serviere, C., (2001). Blind source separation: a tool for rotating
697 machine monitoring by vibration analysis ? *Journal of Sound and Vibration, 248*, 865-885.
698
699 Han, S-C., Shum, C.K., Jekeli, C., Kuo, C-Y., Wilson, C., & Seo K-W. (2005). Non-isotropic
700 filtering of GRACE temporal gravity for geophysical signal enhancement. *Geophysical*
701 *Journal International, 163*, 18–25.
702
703 van Hateren, J.H., & van der Schaaf, A. (1998). Independent component filters of natural
704 images compared with simple cells in primary visual cortex. *Proceedings of the Biological*
705 *Society, 265*, 359–366.
706
707 Hekmeijer P., & Dawes, W. (2003a). Assessment of salinity management options for South
708 Loddon Plains, Victoria: Data analysis and groundwater modeling. CSIRO Land and Water
709 Technical Report 24/03.
710
711 Hekmeijer P., & Dawes, W. (2003b). Assessment of salinity management options for Axe
712 Creek, Victoria: Data analysis and groundwater modelling. CSIRO Land and Water Technical
713 Report 22/03, MDBC Publication 08/03, 40pp.
714
715 Hyvärinen, A. (1999). Fast and Robust Fixed-Point Algorithms for Independent Component
716 Analysis. *IEEE Transactions on Neural Networks*, *10*, 626-634.
717
718 Hyvärinen, A., & Oja, E. (2000). Independent Component Analysis: Algorithms and
719 Applications. *Neural Networks, 13*, 411-430.
720

Ife, D. & Skelt, K. (2004). Murray Darling Basin Groundwater Status 1990-2000. Murray Darling Basin Commission, publication 32/04, Canberra. ISBN 1876830948. Available online at http://www.mdbc.gov.au

Jekeli, C. (1981). Alternative methods to smooth the Earth's gravity field. Tech. Rep., Department of Geodetic Science, Ohio State University, Colombus, Ohio.

Kirby, M., Evans, R., Walker, G., Cresswell, R., oram, S. Khan, Z. Paydar, M. Mainuddin, N. McKenzie and S. Ryan (2006). The shared water resources of the Murray-Darling Basin. Murray-Darling Basin Commission, Publication 21/06, Canberra. ISBN 192103887X. Available online at http://www.mdbc.gov.au.

Klees, R., Revtova, E.A., Gunter, B.C., Ditmar, P., Oudman, E., Winsemius, H.C., & Savenije, H.H.G. (2008). The design of an optimal filter for monthly GRACE gravity models. *Geophysical Journal International, 175*, 417-432, doi: 10.1111/j.1365-246X.2008.03922.x.

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

Lawrence, J. E., & Hornberger, G.M. (2007). Soil moisture variability across climate zones, *Geophysical Research Letters, 34*, L20402, doi:10.1029/2007GL031382.

Leblanc, M. J., Tregoning, P., Ramillien, G., Tweed, S.O., & Fakes, A. (2009). Basin‐scale, integrated observations of the early 21st century multiyear drought in southeast Australia. *Water Resources Research*, *45*, W04408, doi:10.1029/2008WR007333.

Macumber, P.G. (1999). Groundwater flow and resource potential in the Bridgewater and Salisbury West GMAs. Phillip Macumber Consulting Services, Melbourne, 88p.

Oki, T., & Sud, Y.C. (1998). Design of Total Runoff Integrating Pathways (TRIP) - A global river channel network. *Earth Interactions, 2* (1), 1-37.

Paulson, A., Zhong, S. & Wahr, J. (2007). Inference of mantle viscosity from GRACE and relative sea level data. *Geophysical Journal International, 171 (2)*, 497-508.

Peltier, W.R. (2004). Global Glacial Isostasy and the Surface of the Ice-Age Earth: The ICE-5G(VM2) model and GRACE. *Annual Review of Earth and Planetary Sciences,* 32, 111-149.

Petheram, C., Dawes, W., Walker, G., Grayson, R.B. (2003). Testing in class variability of groundwater systems: local upland systems. *Hydrological Processes, 17*, 2297-2313.

Pöyhönen, S., Jover, P., & Hyötyeniemi, H. (2003). Independent Component Analysis of vibrations for fault diagnosis of an induction motor. *Proceedings of the IASTED International Conference on Circuits, Signal and Systems (CSS 2003)*, Cancun, Mexico, 19-21 May 2003, 1, 203-208.

Ramillien, G., Frappart, F., Cazenave, A., & Güntner, A. (2005). Time variations of land water storage from the inversion of 2-years of GRACE geoids. *Earth Planetary Science Letters, 235*, 283-301, doi:10.1016/j.epsl.2005.04.005.

771 Ramillien, G., Frappart, F., Güntner, A., Ngo-Duc, T., Cazenave, A. & Laval, K. (2006a).
772 Time variations of the regional evapotranspiration rate from Gravimetry Recovery And
773 Climate Experiment (GRACE) satellite gravimetry. *Water Resources Research, 42*, W10403,
774 doi:10.1029/2005WR004331.
775
776 Ramillien, G., Lombard, A., Cazenave, A., Ivins, E.R., Llubes, M., Remy, F. & Biancale, R.
777 (2006b). Interannual variations of the mass balance of the Antarctica and Greenland ice sheets
778 from GRACE. *Global* and *Planetary Change, 53(3)*, 198–208.
779
780 Ramillien, G., Famiglietti, J.S., & Wahr, J. (2008). Detection of continental hydrology and
781 glaciology signals from GRACE: a review. *Surveys in Geophysics, 29*, 361-374, doi:
782 10.1007/s10712-008-9048-9.
783
784 Ristaniemi, T., & Joutsensalo, J. (1999). On the Performance of Blind Symbol Separation in
785 CDMA Downlink. *Proceedings of the International Workshop on Independent Component*
786 *Analysis and Signal Separation (ICA'99)*, Aussois, France, January 11-15, 1999, 437-442.
787
788 Rodell, M., Famiglietti, J.S., Chen, J., Seneviratne, S.I., Viterbo, P., Holl S. & Wilson C.R.
789 (2004a), Basin scale estimates of evapotranspiration using GRACE and other observations
790 *Geophysycal Research Letters, 31(20)*, L20504.
791
792 Rodell, M., Houser, P. R., Jambor, U. & Gottschalck, J. (2004b). The Global Land Data
793 Assimilation System. *Bulletin of the American Meteoroogical Society, 85(3)*, 381– 394,
794 doi:10.1175/BAMS-85-3-381.
795
796 Rodell, M., Chen, J., Kato, H., Famiglietti, J.S., Nigro, J. & Wilson, C. (2007). Estimating
797 groundwater storage changes in the Mississippi River basin (USA) using GRACE.
798 *Hydrogeology Journal, 15*(1), 159-166.
799
800 Sasgen, I., Martinec, Z., & Fleming, K. (2006). Wiener optimal filtering of GRACE data.
801 *Studia Geophysica Et Geodaetica,50,* 499 – 508.
802
803 Schmidt, R., Flechtner, F., Meyer, U., Neumayer, K.-H., Dahle, Ch., Koenig, R., & Kusche, J.
804 (2008). Hydrological Signals Observed by the GRACE Satellites.
805 *Surveys in Geophysics, 29,* 319 – 334, doi: 10.1007/s10712-008-9033-3.
806
807 Seitz, F., Schmidt, M. & Shum, C.K. (2008). Signals of extreme weather conditions in Central
808 Europe in GRACE 4-D hydrological mass variations, *Earth and Planetary Science Letters,*
809 *268(1-2)*, 165-170.
810
811 Seo, K-W., & Wilson, C.R. (2005). Simulated estimation of hydrological loads from GRACE,
812 *Journal of Geodesy, 78*, 442–456.
813
814 Smitt, C., Doherty, J., Dawes, W. & Walker, G. (2003). Assessment of salinity management
815 options for the Brymaroo catchment, South-eastern Queensland. CSIRO Land and Water
816 Technical Report 23/03.
817
818 Stone, J.V. (2004). *Independent Component Analysis: A Tutorial Introduction*, MIT Press:
819 Bradford Book.
820

Strassberg, G., Scanlon, B. R., & Rodell, M. (2007) Comparison of seasonal terrestrial water storage variations from GRACE with groundwater-level measurements from the High Plains Aquifer (USA). *Geophysical Research Letters, 34*, L14402, doi:10.1029/2007GL030139.

Swenson, S., & Wahr, J. (2006). Post-processing removal of correlated errors in GRACE data, *Geophysycal Research Letters*, *33, L08402*, doi:10.1029/2005GL025285.

Syed, T.H., Famiglietti, J.S., Chambers, D. (2009). GRACE-based estimates of terrestrial freshwater discharge from basin to continental scales, *Journal of Hydrometeorology, 10(1)*, doi: 10.1175/2008JHM993.1.

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

Tregonning, P., Ramillien, G., McQueen, H. & Zwartz, D. (2009). Glacial isostatic adjustement observed by GRACE. *Journal of Geophysical Research, 114*, B06406, doi:10.1029/2008JB006161.

Urbano, L. D., Person, M., Kelts, K. & Hanor, J. S. (2004). Transient groundwater impacts on the development of paleoclimatic lake records in semi-arid environments. *Geofluids, 4*, 187–196.

Vigário, R. (1997). Extraction of ocular artifacts from EEG using independent component analysis, *Encephalography and Clinical Neurophysiology, 103*(3), 395-404.

Tables

Table 1: The 29 drainage basins considered in this study sorted by decreasing area.

Figures

Figure 1: GRACE water storage from GFZ filtered with a Gaussian filter of 400 km of radius. (Top) First ICA component corresponding to land hydrology and ocean mass. (Bottom) Sum of the second and third components corresponding to the north–south stripes. (a) March 2006, (b) September 2006, (c) March 2007, (d) March 2008. Units are millimeters of EWT.

Figure 2: Time series of the kurtosis of the mass anomalies detected by GRACE after Gaussian filtering for radii of a) 300 km, b) 400 km, c) 500 km.

Figure 3: Correlation maps over the period 2003-2008 between the ICA-filtered TWS and the Gaussian-filtered TWS. Left column: ICA400-G400 (a: CSR, c: GFZ, e: JPL). Right column: ICA500-G500 (b: CSR, d: GFZ, f: JPL).

Figure 4: RMS maps over the period 2003-2008 between the ICA-filtered TWS and the Gaussian-filtered TWS. Left column: ICA400-G400 (a: CSR, c: GFZ, e: JPL). Right column: ICA500-G500 (b: CSR, d: GFZ, f: JPL).

Figure 5: Correlation maps over the period 2003-2008 between the ICA-filtered TWS and the destriped and smoothed TWS. Left column: ICA400-DS300 (a: CSR, c: GFZ, e: JPL). Right column: ICA500-DS500 (b: CSR, d: GFZ, f: JPL).

Figure 6: RMS maps over the period 2003-2008 between the ICA-filtered TWS and the destriped and smoothed TWS. Left column: ICA400-DS300 (a: CSR, c: GFZ, e: JPL). Right column: ICA500-DS500 (b: CSR, d: GFZ, f: JPL).

Figure 7: GRACE water storage from GFZ destriped and smoothed with a Gaussian filter of 300 km of radius for March 2006. (Top) First ICA component corresponding to land hydrology. (Bottom) Sum of the second and third components corresponding to the north–south stripes.

Figure 8: Trend maps over the period 2003-2008 of TWS using the GFZ solutions a) ICA400, b) ICA500, c) G400, d) G500, e) DS300 and f) DS500.

Figure 9: Location of the 27 drainage basins chosen in this study. See Table 1 for the correspondence between basins and numbers.

Figure 10: Time series of TWS (mm) derived from ICA400 (black), G400 (blue), DS300 (red) for GFZ solutions over the Amazon (a), Ob (b) and Mekong (c) basins.

Figure 11: RMS between ICA400 and G400 for CSR (dark blue), GFZ (red), and JPL (dark green) solutions and between ICA400 and DS300 for CSR (light blue), GFZ (orange), and JPL (green) solutions per basin (sorted by decreasing area of drainage basin) over the period October 2002 – July 2009.

Figure 12: Standard deviation of the leakage error (mm) per basin (sorted by decreasing area of drainage basin) over the period October 2002 – July 2009 for the ICA solutions.

Figure 13: Map of the Murray Darling drainage basin in Australia. Cumulative rainfall deficit across the Murray Darling Basin for the 2001–2006 period and location of the shallow groundwater monitoring bores.

Figure 14: Comparison of GRACE TWS annual anomalies with hydrological estimates from *in situ* measurements (SW and GW) and modeling (SM) for the period 2003–2008. The grey curves correspond to *in situ* + model TWS, the blue to ICA-filtered solutions, the green to Gaussian filtered, and the red to destriped GFZ solutions at a) 400 km of filtering (300 km for the destriped solutions) and b) 500 km.

Figure 15: Yearly deviation over 2003-2006 from *in situ* + model TWS of GRACE-derived TWS filtered with different approaches at a) 400 km of filtering (300 km for the destriped solutions) and b) 500 km. In blue, the ICA solutions (dark blue: CSR, blue: GFZ, light blue: JPL), in green, the Gaussian solutions (dark green: CSR, green: GFZ, light green: JPL), and the destriped and smoothed solutions (red: CSR, orange: GFZ, yellow: JPL).

Tables

Table 1: The 29 drainage basins considered in this study sorted by decreasing area.

Number	River basin	Area (10^6 km²)
1	Amazon	6.20
2	Amur	1.88
3	Brahmaputra	0.65
4	Colorado	0.63
5	Congo	3.81
6	Danube	0.81
7	Dniepr	0.52
8	Euphrates	0.81
9	Ganges	0.97
10	Hwang Ho	0.74
11	Lena	2.45
12	McKenzie	1.73
13	Mekong	0.77
14	Mississippi	3.32
15	Niger	2.18
16	Nile	3.16
17	Ob	2.84
18	Okavango	0.83
19	Orinoco	0.87
20	Parana	2.98
21	St Lawrence	1.12
22	Tocantins	0.86
23	Volga	1.42
24	Yangtze	1.78
25	Yenisey	2.56
26	Yukon	0.82
27	Zambezi	1.39

Figures

Figure 1: GRACE water storage from GFZ filtered with a Gaussian filter of 400 km of radius for March 2006. (Top) First ICA component corresponding to land hydrology and ocean mass. (Bottom) Sum of the second and third components corresponding to the north–south stripes. (a) March 2006, (b) September 2006, (c) March 2007, (d) March 2008. Units are millimeters of EWT.

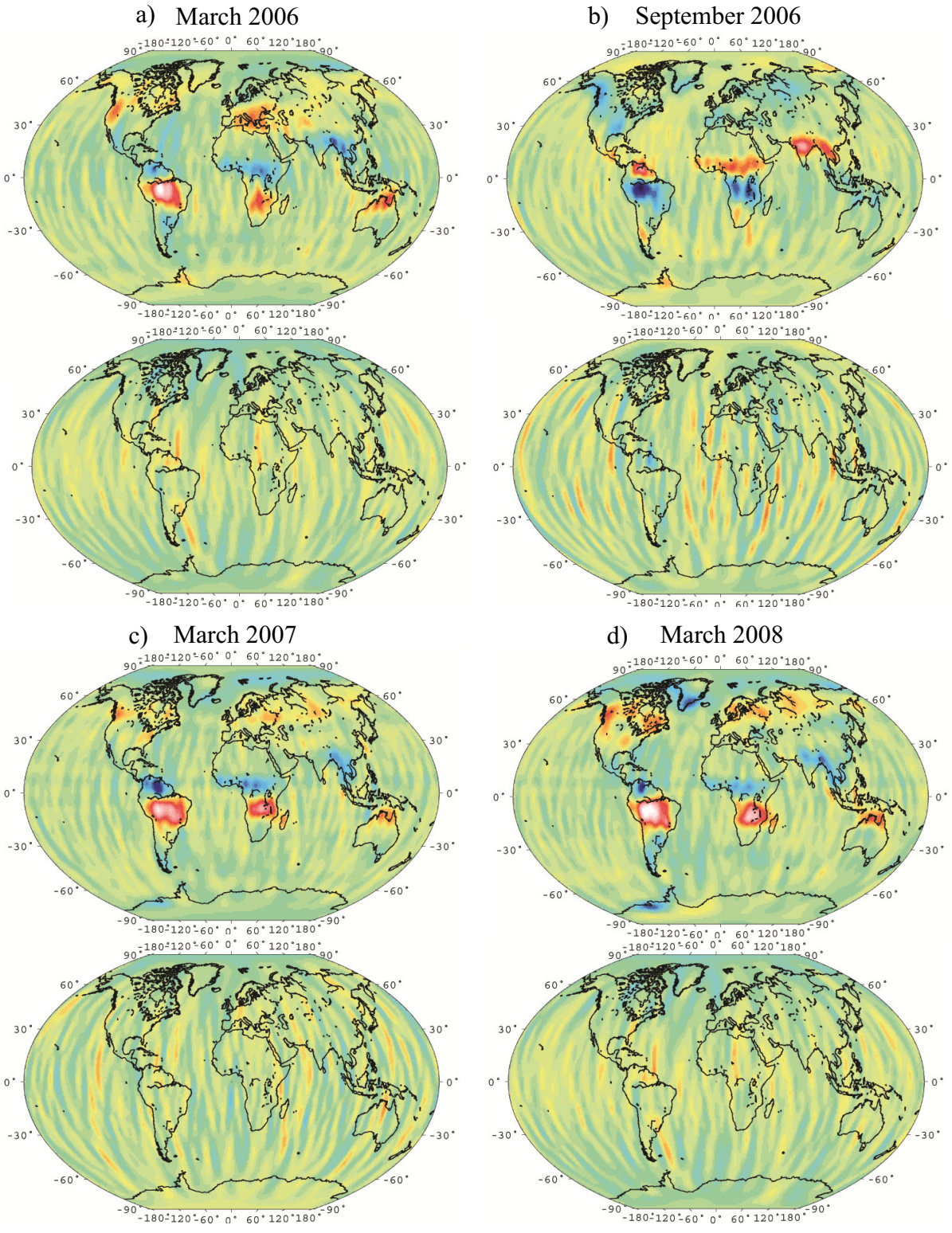


Figure 2: Time series of the kurtosis of the independent components (1st and sum of 2nd and 3rd) of mass anomalies detected by GRACE for different radii of Gaussian filtering: a) 300 km, b) 400 km, c) 500 km.

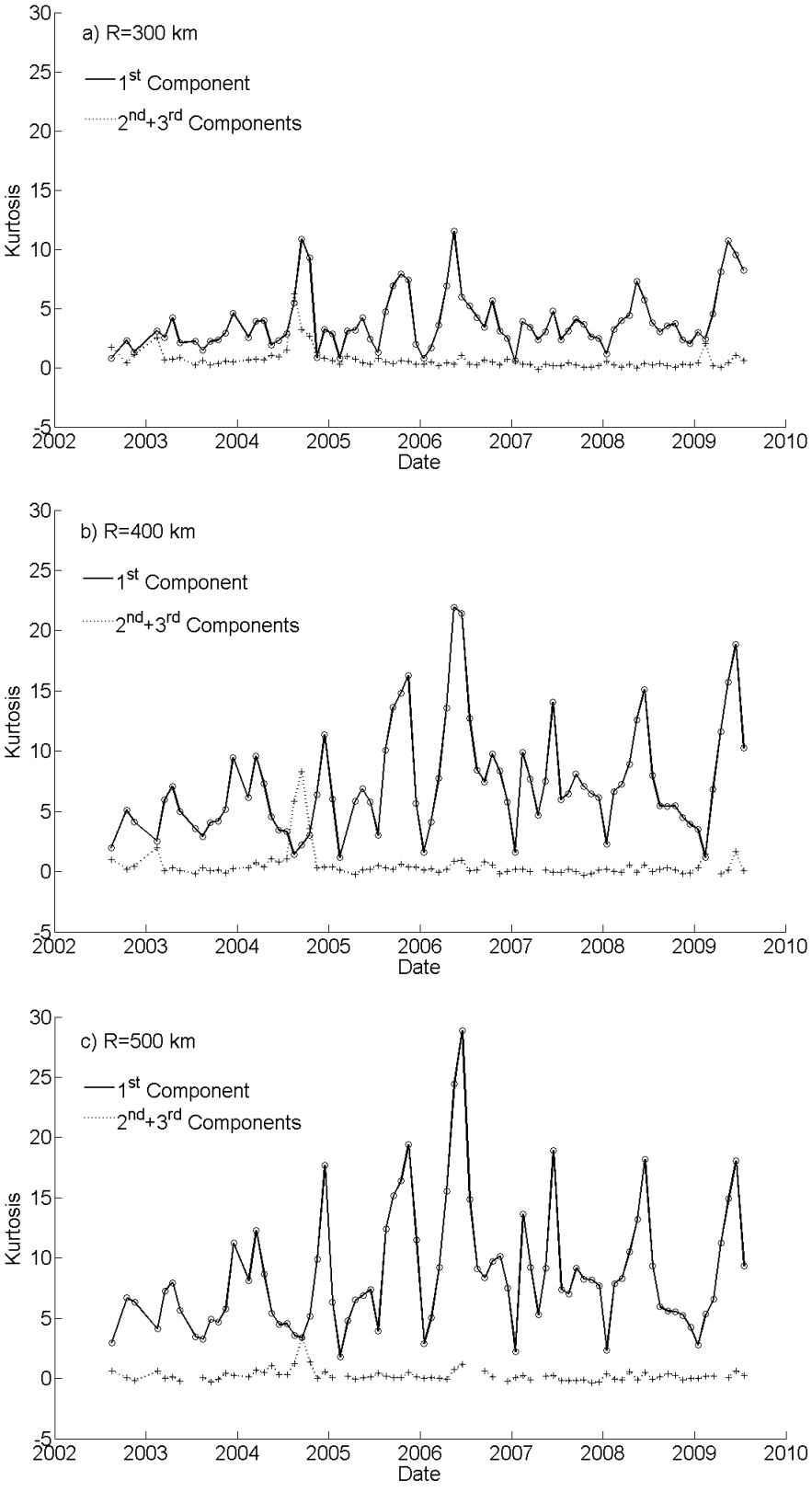


Figure 3: Correlation maps over the period 2003-2008 between the ICA-filtered TWS and the Gaussian-filtered TWS. Left column: ICA400-G400 (a: CSR, c: GFZ, e: JPL). Right column: ICA500-G500 (b: CSR, d: GFZ, f: JPL).

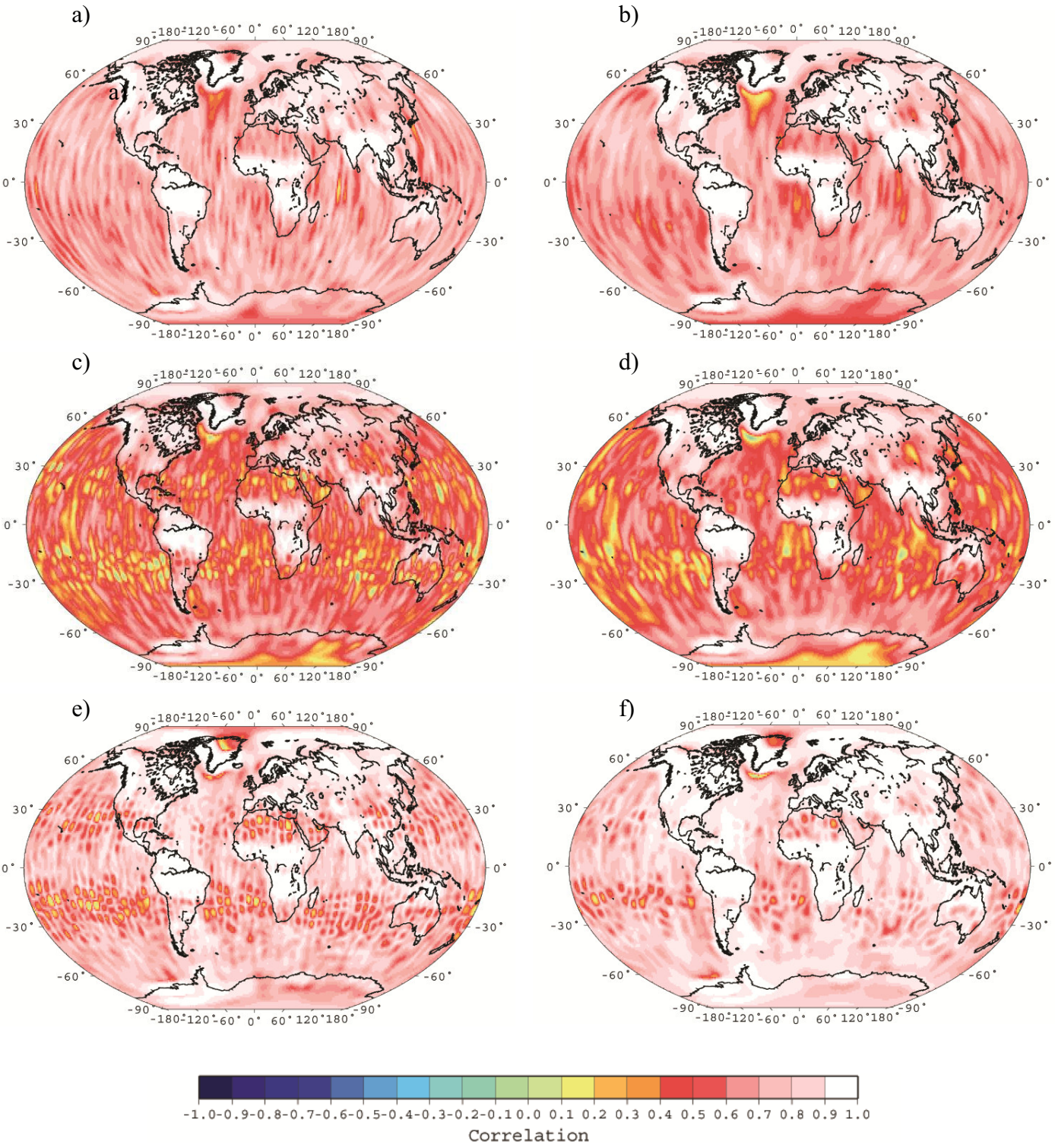


Figure 4: RMS maps over the period 2003-2008 between the ICA-filtered TWS and the Gaussian-filtered TWS. Left column: ICA400-G400 (a: CSR, c: GFZ, e: JPL). Right column: ICA500-G500 (b: CSR, d: GFZ, f: JPL).

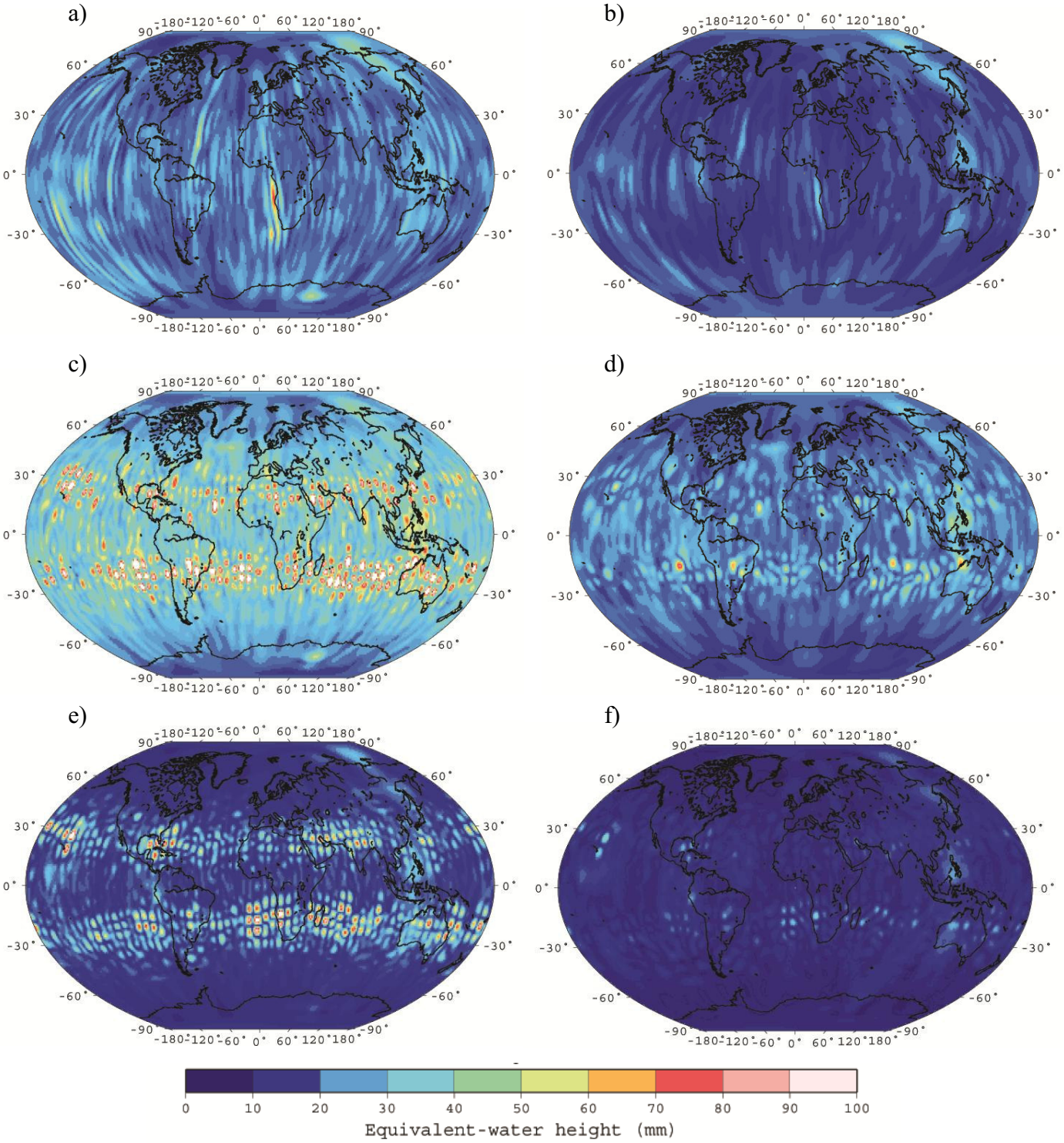


Figure 5: Correlation maps over the period 2003-2008 between the ICA-filtered TWS and the destriped and smoothed TWS. Left column: ICA400-DS300 (a: CSR, c: GFZ, e: JPL). Right column: ICA500-DS500 (b: CSR, d: GFZ, f: JPL).

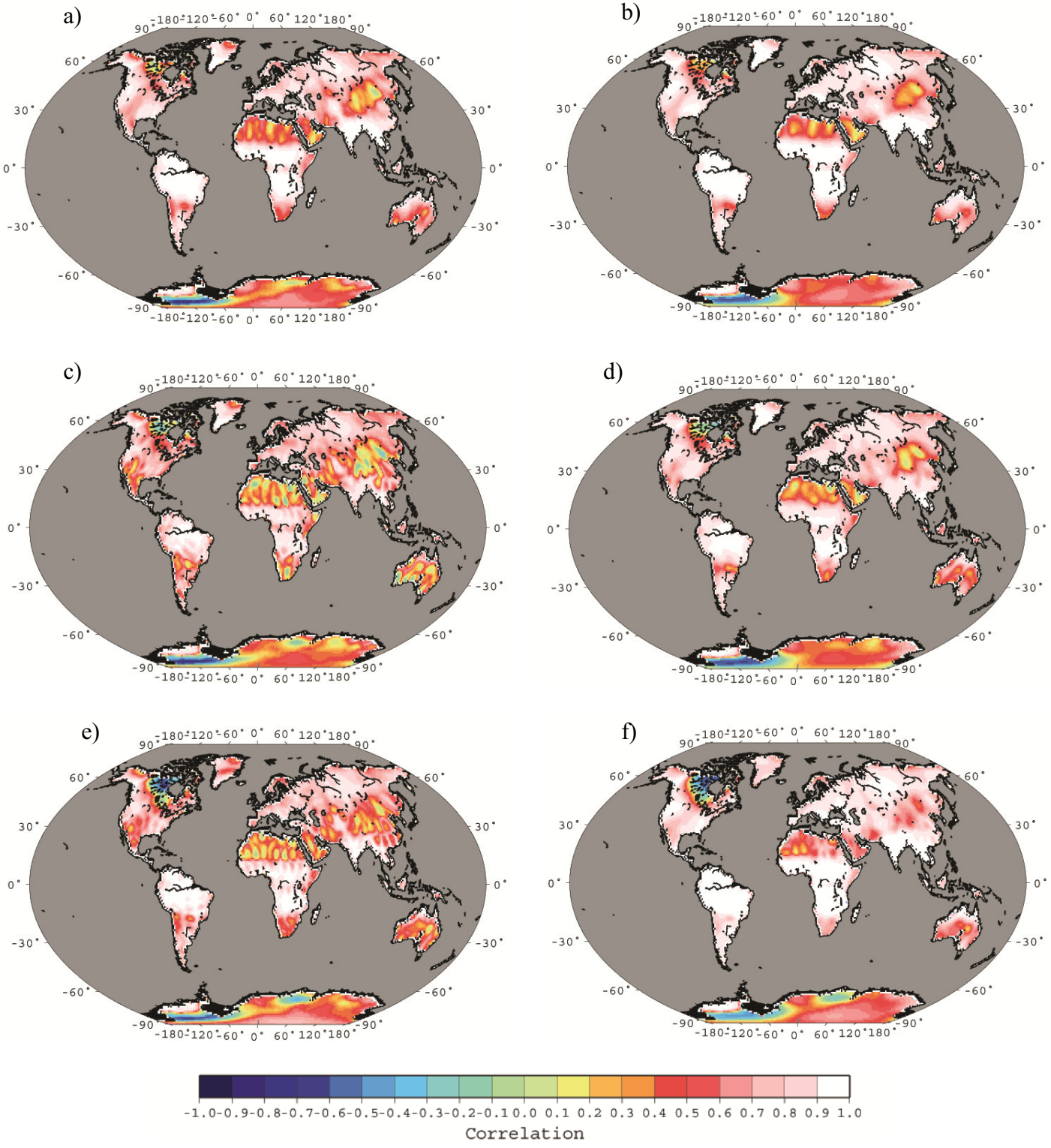


Figure 6: RMS maps over the period 2003-2008 between the ICA-filtered TWS and the destriped and smoothed TWS. Left column: ICA400-DS300 (a: CSR, c: GFZ, e: JPL). Right column: ICA500-DS500 (b: CSR, d: GFZ, f: JPL).

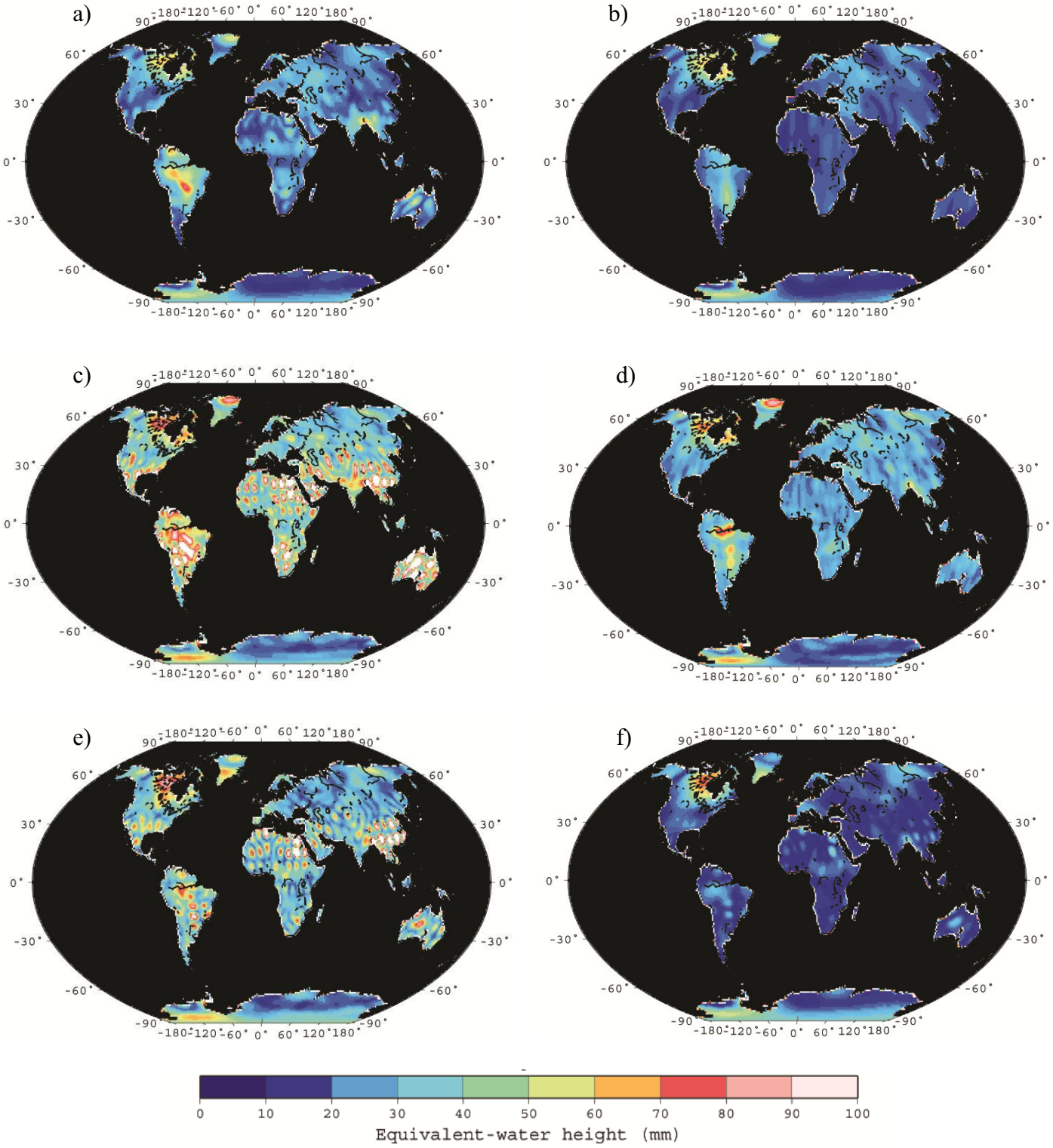
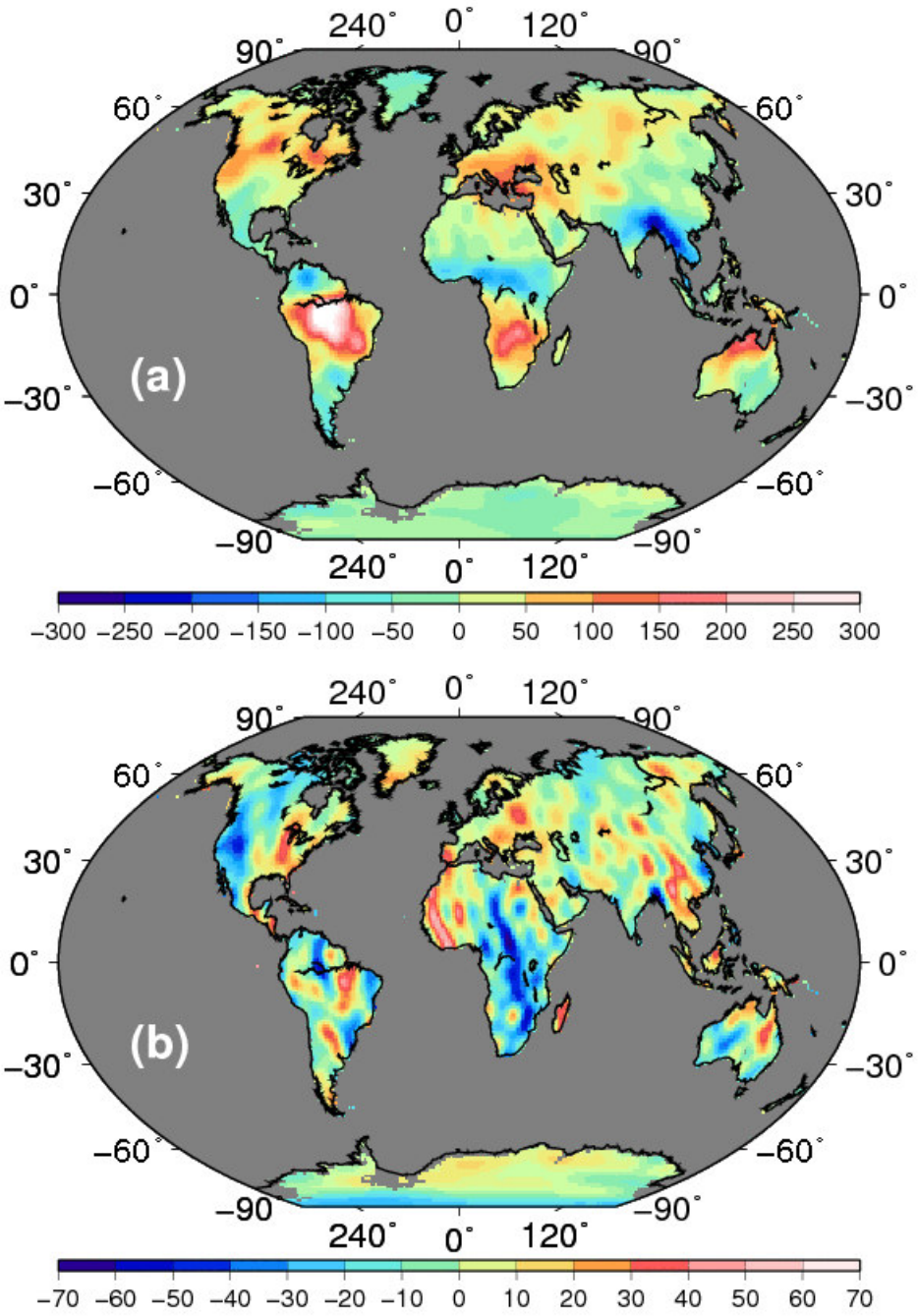


Figure 7: GRACE water storage from GFZ destriped and smoothed with a Gaussian filter of 300 km of radius for March 2006. (Top) First ICA component corresponding to land hydrology. (Bottom) Sum of the second and third components corresponding to the north–south stripes.



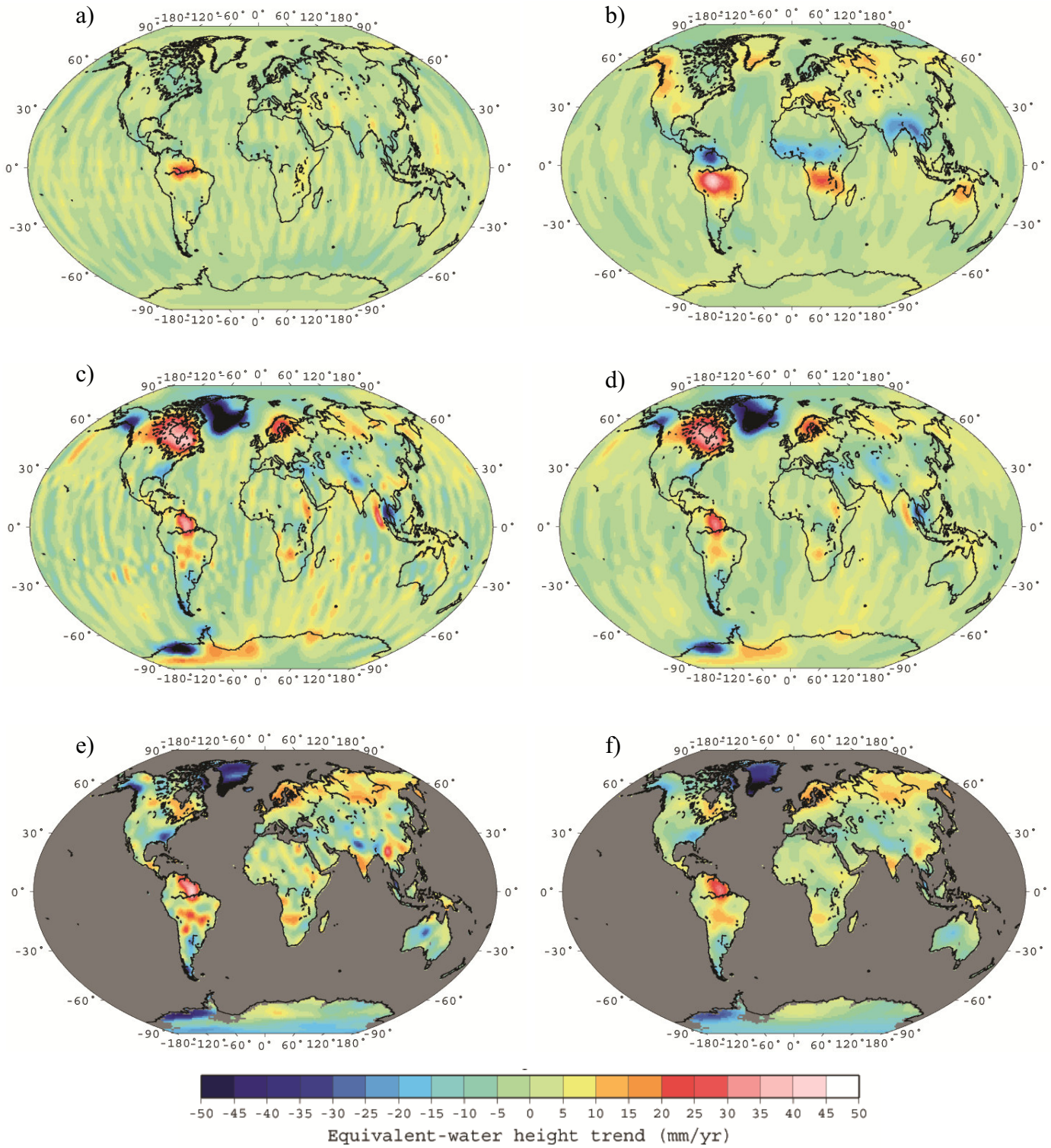


Figure 8: Trend maps over the period 2003-2008 of TWS using the GFZ solutions a) ICA400, b) ICA500, c) G400, d) G500, e) DS300 and f) DS500.

Figure 9: Location of the 27 drainage basins chosen in this study. See Table 1 for the correspondence between basins and numbers.

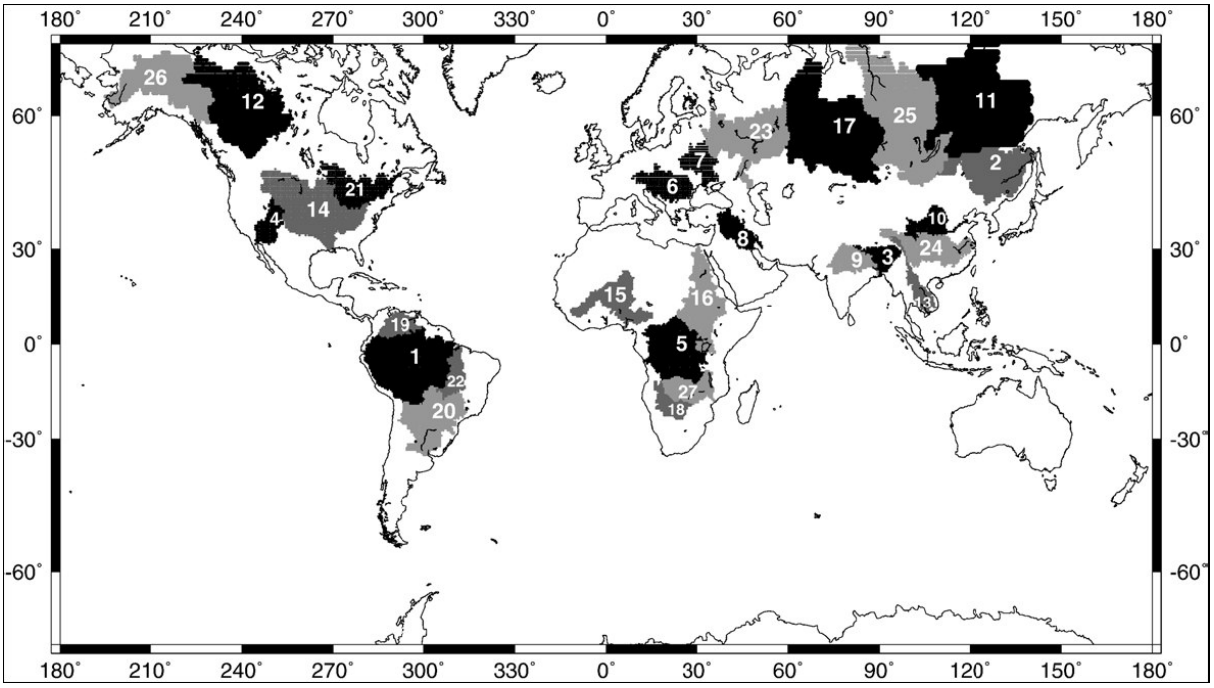


Figure 10: Time series of TWS (mm) derived from ICA400 (black), G400 (blue), DS300 (red) for GFZ solutions over the Amazon (a), Ob (b) and Mekong (c) basins.

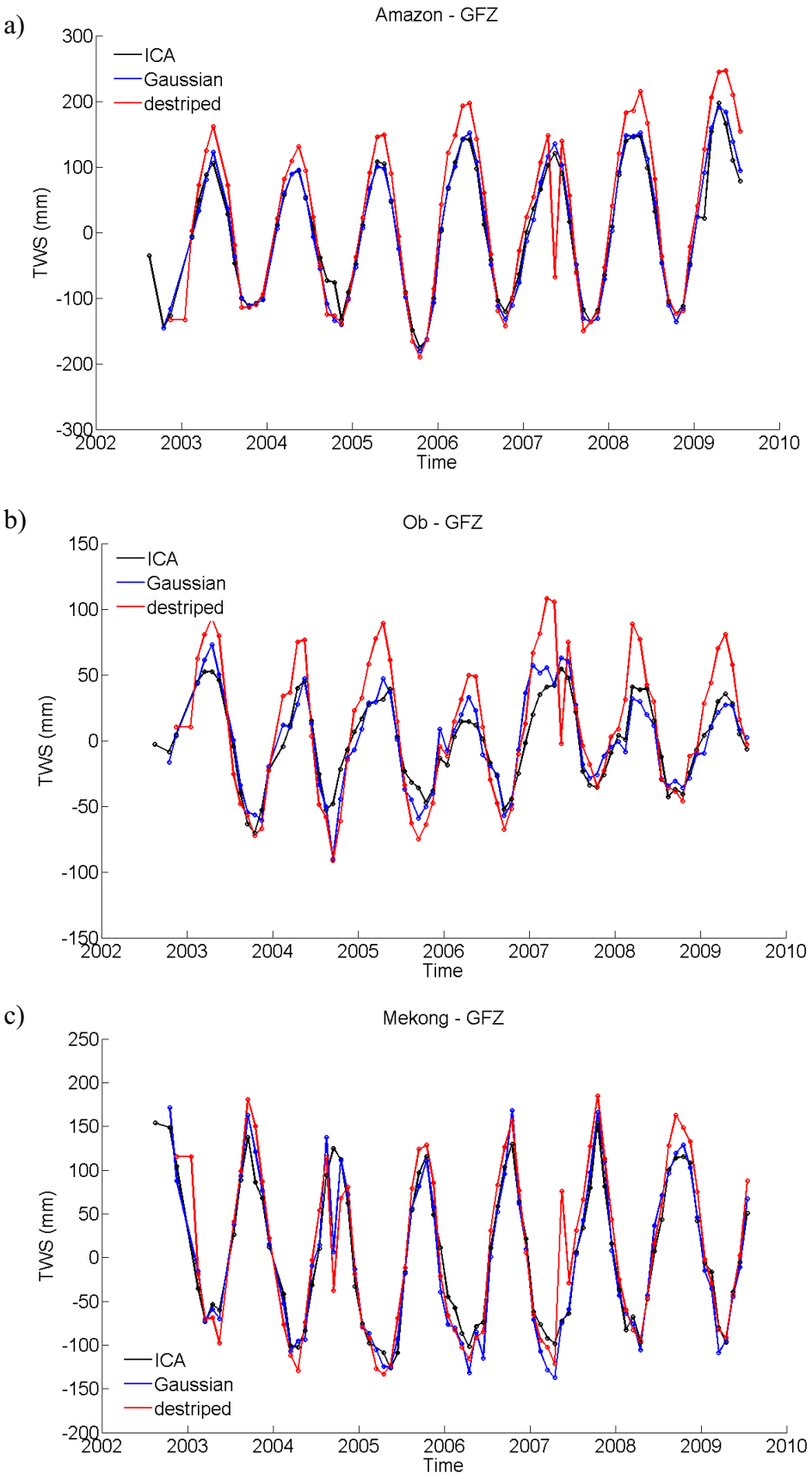


Figure 11: RMS between ICA400 and G400 for CSR (dark blue), GFZ (red), and JPL (dark green) solutions and between ICA400 and DS300 for CSR (light blue), GFZ (orange), and JPL (green) solutions per basin (sorted by decreasing area of drainage basin) over the period October 2002 – July 2009.

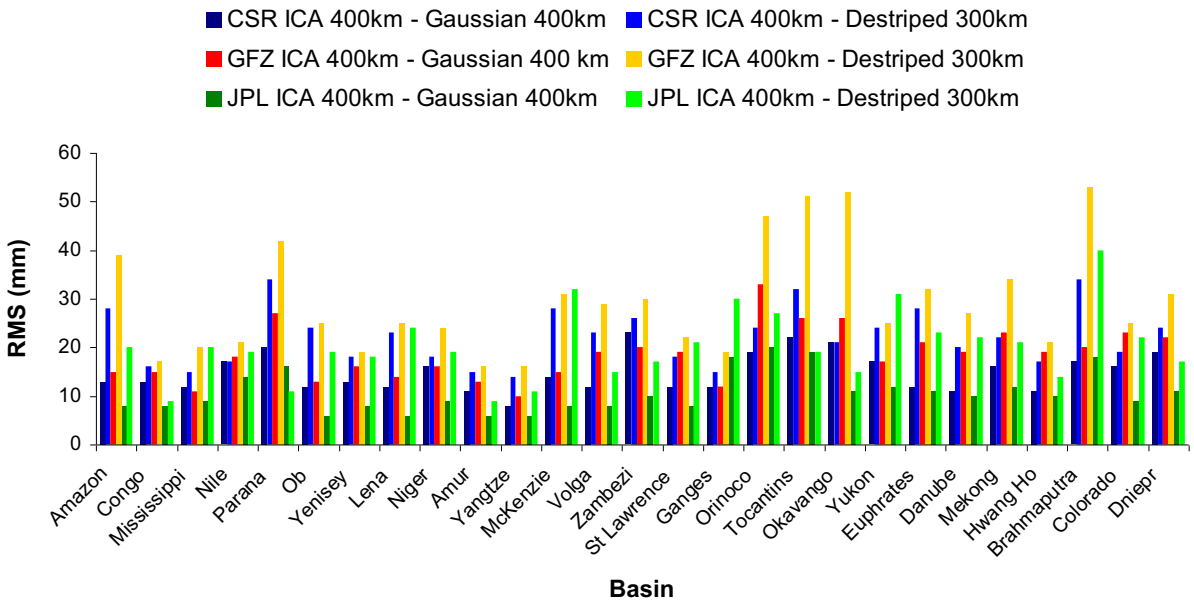


Figure 12: Standard deviation of the leakage error (mm) per basin (sorted by decreasing area of drainage basin) over the period October 2002 – July 2009 for the ICA solutions.

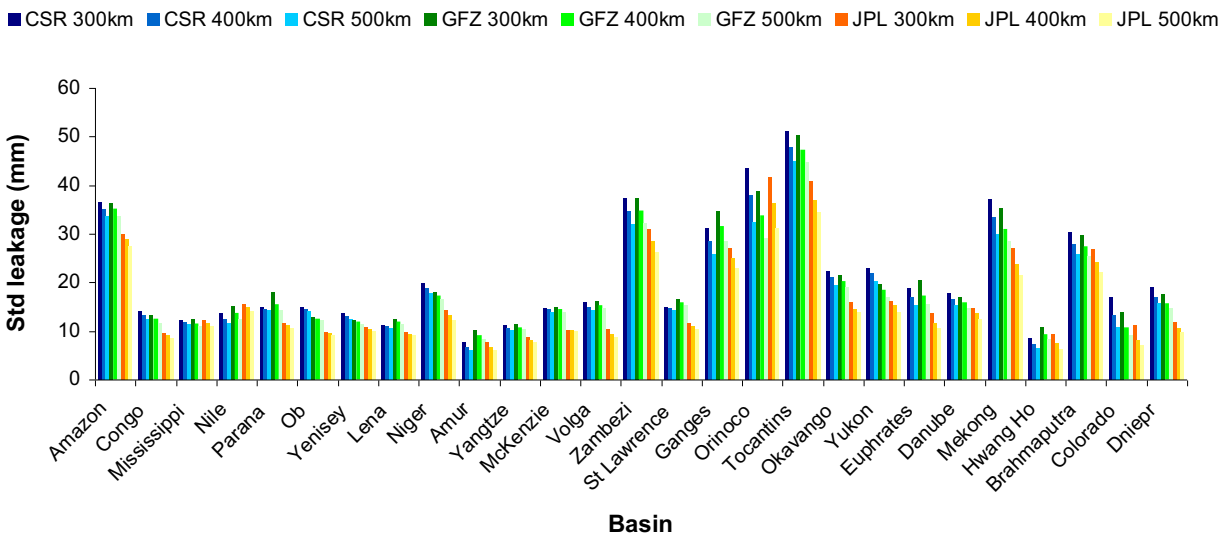


Figure 13: Map of the Murray Darling drainage basin in Australia. Cumulative rainfall deficit across the Murray Darling Basin for the 2001–2006 period and location of the shallow groundwater monitoring bores.

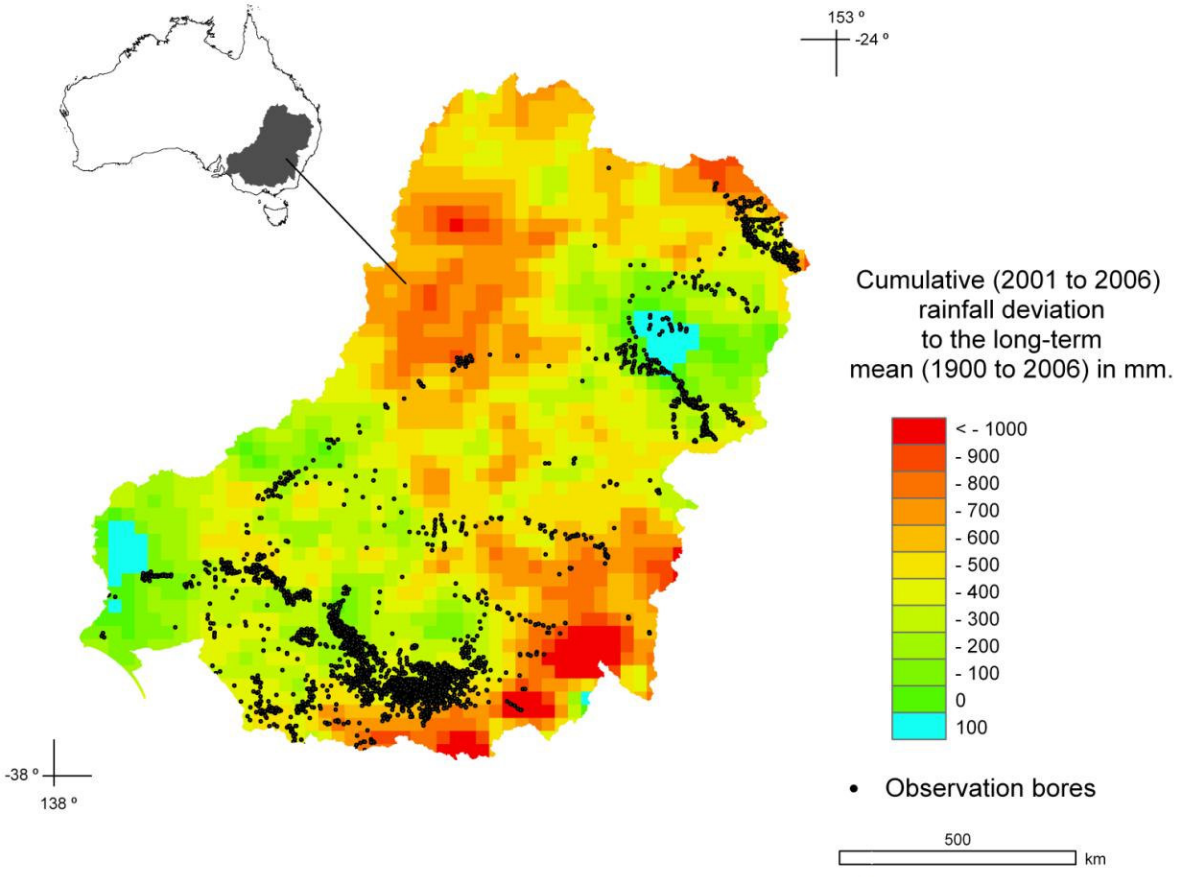


Figure 14: Comparison of GRACE TWS annual anomalies with hydrological estimates from *in situ* measurements (SW and GW) and modeling (SM) for the period 2003–2008. The grey curves correspond to *in situ* + model TWS, the blue to ICA-filtered solutions, the green to Gaussian filtered, and the red to destriped GFZ solutions at a) 400 km of filtering (300 km for the destriped solutions) and b) 500 km.

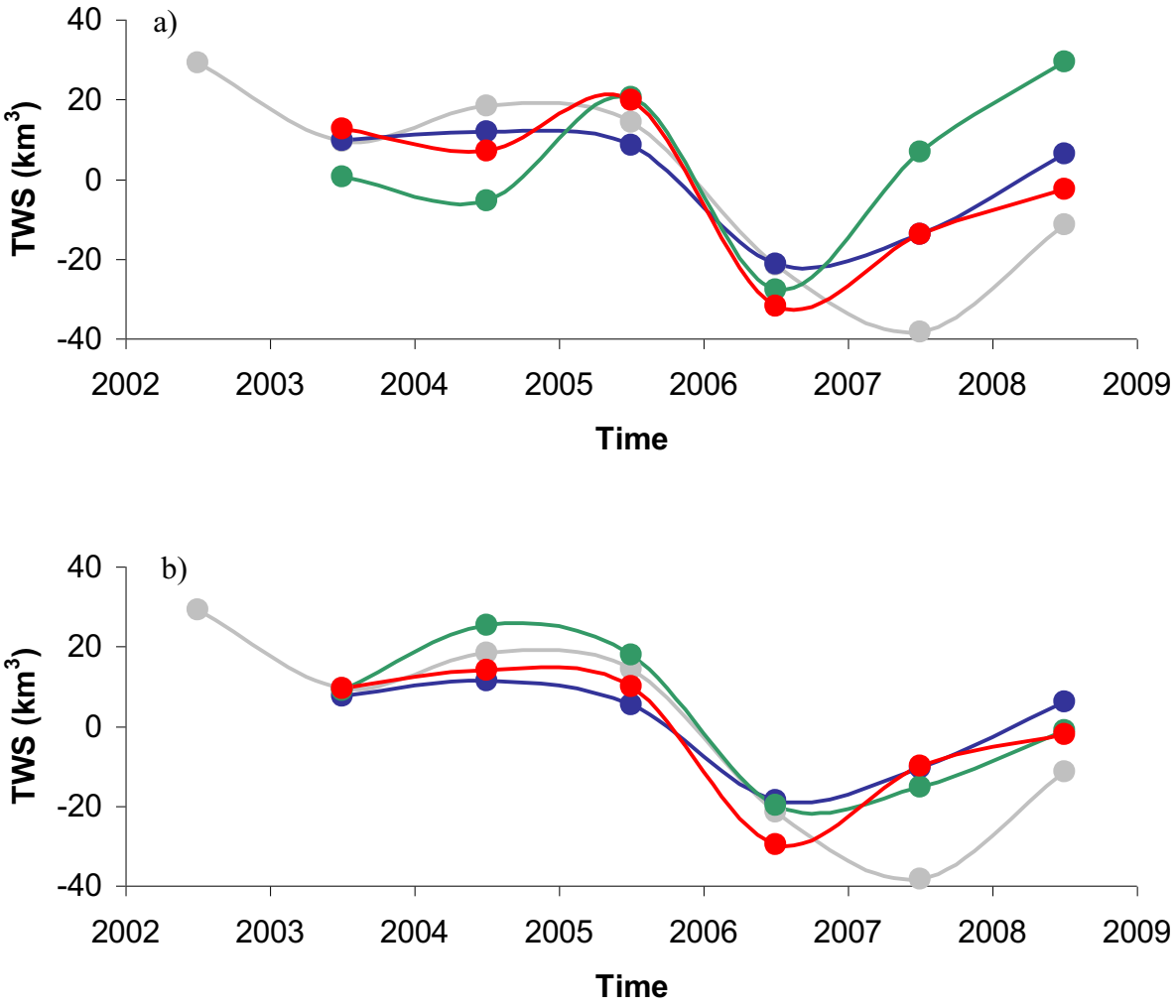


Figure 15: Yearly deviation over 2003-2006 from *in situ* + model TWS of GRACE-derived TWS filtered with different approaches at a) 400 km of filtering (300 km for the destriped solutions) and b) 500 km. In blue, the ICA solutions (dark blue: CSR, blue: GFZ, light blue: JPL), in green, the Gaussian solutions (dark green: CSR, green: GFZ, light green: JPL), and the destriped and smoothed solutions (red: CSR, orange: GFZ, yellow: JPL).