Global time-variations of hydrological signals from GRACE satellite gravimetry

G. Ramillien, A. Cazenave, O. Brunau

Geophys. J. Int. (2004) **158,** 813–826 doi: 10.1111/j.1365-246X.2004.02328.x

Global time variations of hydrological signals from GRACE satellite gravimetry

G. Ramillien, A. Cazenave and O. Brunau

LEGOS-CNES, UMR 5566*/Observatoire Midi-Pyrénées,* 18, *Avenue Edouard Belin, Toulouse* 31401*, Cedex* 04*, France. E-mail: ramillie@notos.cst.cnes.fr*

Accepted 2004 April 16. Received 2004 April 5; in original form 2003 June 20

SUMMARY

Successfully launched in 2002 mid-March, the goal of the Gravity Recovery And Climate Experiment (GRACE) satellite mission is to measure the spatio-temporal variations of the gravity field of the Earth to high accuracy (∼1 cm in terms of geoid height) and a spatial resolution of ∼200–300 km, for a nominal period of 5 yr. The unprecedented precision of the GRACE mission will enable us to detect tiny time variations of the gravity field related to global redistributions of water and air mass inside fluid envelops at the surface of the Earth. In this paper, we present a new approach based on linear inverse methods to separate the different contributions of the main surface fluid reservoirs (oceans, atmosphere, total continental water storage including snow, soil wetness, ground water and ice caps) from monthly synthetic GRACE geoids. The synthetic geoids were computed from outputs of global models of different climatic fields. Because of the non-uniqueness of the classical inverse problems in gravimetry, independent information was added before inverting the synthetic geoids. Geoid solutions associated with each fluid contribution were then converted into water-equivalent thickness maps. Validation of the continental water storage solutions was performed by comparing total soil water estimates [soil moisture (SM) plus groundwater] with predictions of a global hydrological model in 71 different drainage basins of the world. Analysis of the *a posteriori* errors of the solutions suggests that the inversion method developed in this study allows recovering monthly water mass changes with a cm precision.

Key words: global hydrology, GRACE satellite, inverse problem, time-variable gravimetry.

1 INTRODUCTION

After a successful launch on the 2002 March 17, the Gravity Recovery And Climate Experiment (GRACE) mission developed by NASA in the USA and Deutsches Zentrum fur Luft und Raumfahrt in Germany is currently mapping the variations of the gravity field of the Earth over its 5-yr lifetime. The gravity GRACE mission scheduled for ∼5 yr consists of two identical spacecrafts flying approximately 220 km apart in a polar orbit at an altitude of 400–500 km. GRACE detects time variations of the global gravity field by making accurate measurements of the varying relative speed between the two satellites using a microwave Doppler system. For the very first time, this satellite technique will provide an efficient way to measure time variations of the gravity field with an accuracy better than ∼1 cm rms in terms of geoid height and an unprecedented resolution (200–400 km spatially and 1 month temporally). This will represent crucial information about the distribution and flow of mass within the Earth system and specially inside the surface fluid envelops. The gravity variations that GRACE can detect include: changes as a result of surface and deep currents in the oceans; changes in soil and ground water storage on land; mass changes of the ice sheets and glaciers; air and water vapor mass change within the atmosphere; and variations of mass within the solid Earth. The promising results from GRACE will make a huge contribution to global climate change studies. The potentials of the GRACE mission to recover hydrological signals are presented in the National Research Council (NRC) report (Dickey *et al.* 1997) and ESA (2000).

However, one disadvantage is that GRACE products are values of integrated mass over a vertical column of matter, thus it is not possible to distinguish between the different sources of geoid anomaly. Numerically, separating each components is a problem that requires supplementary information, e.g. other type of satellite data, *in situ* measurements or forecasts of hydrological variables based on global climate models.

In this paper, we present a new method based on an inverse approach using least-squares criteria to unravel the different contributions of the main surface fluid reservoirs: atmosphere, oceans and total continental water storage. In this study, we ignore the ice sheets and the solid Earth contributions. In order to stabilize the inversion procedure, independent information derived from outputs of the global atmospheric, hydrological and oceanic models is used. First, we describe the direct and inverse procedures used for recovering the harmonic coefficients of the main fluid contributions from the gravity field. In the following section, we apply the proposed inverse

method to unravel long-time series of synthetic GRACE geoids that are constructed using gravity data derived by combining outputs of climatic models, because GRACE geoid data are not yet available to the scientific community. In order to assess the ability of GRACE for monitoring global hydrological signals such as total soil water and snow cover over continents, we extract time-variable maps of the solutions in some of the largest drainage basins of the globe. We particularly focus on South American basins: Amazon, Paraná, Orinoco, Araguaia, São Francisco and Uruguay basins. Finally, *a posteriori* errors of inversion are evaluated by analysing the global differences between the estimates of the four fluid contributions (atmosphere, oceans, soil water and snow cover) and the ones from the starting models.

2 PREVIOUS STUDIES ON THE PREPARATION OF GRACE DATA ANALYSIS

Several authors have already demonstrated the ability of GRACE to monitor water storage variability over continental areas of several hundreds of km or larger, anywhere in the world (Dickey *et al.* 1997; Wahr *et al.* 1998; Rodell & Famiglietti 1999, 2002, 2001; Velicogna *et al.* 2001; Swenson & Wahr 2002). Wahr *et al.* (1998) previously used outputs from global atmospheric, oceanographic and hydrological models to estimate the variability of the gravity field resulting from those sources. These authors proposed a forward method for using the gravity information from GRACE to infer time changes in surface mass and estimate the corresponding accuracy at monthly intervals. Their method is based on the filtering of spherical harmonic coefficients for constructing global maps of surface mass variations from synthetic GRACE data and taking the deformation of the surface of the Earth into account (i.e. introducing elastic Love numbers in the evaluation). Time changes of GRACE geoid coefficients δC_{nm} and δS_{nm} (of degree n and order m) are constructed as spatial averaging of surface mass coefficients using an arbitrary weighting function W, such as the gaussian operator proposed by Jekeli (1981). According to Wahr *et al.* (1998), using an operator with a large enough radius of averaging (i.e. hundreds of km) should reduce the leakage of energy between the oceanic and continental domains while recovering water mass variations on land. Leakage oscillations at the continent/ocean boundaries would represent an important source of errors in water mass recovery, especially for small river basins. This early work largely inspired subsequent studies addressing the accuracy of recovering water mass signals in continental basins from synthetic GRACE geoids (Rodell & Famiglietti 1999, 2001, 2002; Velicogna *et al.* 2001; Swenson & Wahr 2002). The results of these simulations suggest that final accuracy of the recovered soil water would increase with the size of the monitored river basin at the monthly, seasonal, annual and interannual timescales. Comparing variations of soil moisture (SM) and ground water (GW) contents from a network of stations in Illinois ($\sim$145 800 km^2), Rodell & Famiglietti (1999, 2001) concluded that SM and GW signals would be detectable by GRACE for basin sizes greater than 200 000 km^2 with a precision of a few mm in water height. According to these authors, the uncertainty of the mean variations of GRACE-derived GW in the High Plain aquifer, that represents $\sim$450 000 km^2, is approximately 8.7 mm, whereas the amplitudes of water height is 19.8 and 45.2 mm for annual and 4-yr periods, respectively (Rodell & Famiglietti 2000). Latest developments by Swenson & Wahr (2002) addressed the need of using an averaging kernel to cancel the effects of the short oscillations of high amplitude while recovering water storage variations. However, this stabilizing procedure has the disadvantage of leakage as a result of the gravitational influence of masses outside the studied area that yields underestimated prediction. Another source of errors resulting from high-frequency atmospheric perturbations has been identified by Velicogna *et al.* (2001) while extracting the GRACE signals over continental areas by direct subtraction of the atmospheric contribution. Removing the effects of the atmosphere needs the 30-day mean surface pressure to be known to an accuracy of $\sim$1 mbar (i.e. 1 cm of equivalent water height) or less, whereas short-period errors in the pressure field would be aliased in the GRACE data and can potentially contribute to errors of $\sim$1 cm in terms of water thickness.

3 METHOD

3.1 The direct problem: from water mass distribution to geoid anomaly maps

The basic data set provided to GRACE users is expected to be time-series of spherical harmonic coefficients, up to degree $N = 100$ at monthly intervals. Thus, the first step for studying the fluid contributions to the gravity field will consist of extracting monthly variations of each contribution.

The stationary component of the geoid G_0 that characterizes the main contribution to the gravity field (nearly 99 per cent) as a result of the solid part of the Earth, is merely evaluated by computing the mean geoid map from several geoids $G(t)$ that concern a long enough period of time, such as several years of GRACE observations.

The monthly time-variable geoid $\delta G(t)$ corresponds to the difference between the monthly geoid $G(t)$ measured by GRACE and the static mean component G_0:

$$\delta G(t) = G(t) - G_0, \tag{1}$$

whose $\delta C_{nm}(t)$ and $\delta S_{nm}(t)$ are the normalized Stokes coefficients expressed in terms of millimeters of geoid, which depend upon time t. It will be these coefficients that will be provided to the GRACE users. n and m are the harmonic degree and order respectively, so that the corresponding harmonic coefficients can be summed up to a given maximum degree N for constructing the map of the geoid anomaly:

$$\delta G(t) = \sum_{n=1}^{N} \sum_{m=0}^{n} (\delta C_{\rm nm}(t) \cos m\lambda + \delta S_{nm}(t) \sin m\lambda) P_{\rm nm}(\cos\theta), \tag{2}$$

where θ is the co-latitude, λ is longitude and P_{nm} is the normalized associated Legendre function, which is dimensionless (see Heiskanen & Moritz 1967).

In the following, we consider $\delta G(t)$ corresponding to the sum of the $k = 1, 2, \ldots, K$ contributions $\delta G_k(t)$ of different independent fluid reservoirs, but neglecting those from mass variations inside the solid Earth induced by tectonics such as earthquakes, postglacial rebound and mantle convection. Moreover, we assume that the fluid contributions are not correlated in time and space in order to simplify the modelling approach:

$$\delta G(t) = \sum_{k=1}^{K} \delta G_k(t) = A \delta G_k, \tag{3}$$

where A is the separating matrix formed by a column of identity matrices (see below).

For a given surface density of mass $\delta h(t)$ at the surface of the Earth, expressed in terms of equivalent-water height, the

Table 1. Characteristics of the available global data sets used in the inversion.

Reservoir	Model	Parameters	Period
Atmosphere	ECMWF	Surface pressure from global re-analysis	1979–1993
	NCEP	of ground observations	1979–1996
Oceans	POCM	Variations of ocean bottom	1979–1997
	OPA	pressure predicted by global	1992–1999
	MIT	ocean circulation model	1985–1996
Land waters	LaD	soil water and ground water	1981–1998
	GSWP		1987–1988
	H96	soil water	1979–1998

corresponding geoid anomaly coefficients can be predicted from the the surface density ones $\delta C^h_{nm}(t)$ and $\delta\delta S^h_{nm}(t)$, using a fast linear filtering:

$$\begin{Bmatrix} \delta C_{nm}(t) \\ \delta S_{nm}(t) \end{Bmatrix} = W^0_n \begin{Bmatrix} \delta C^h_{nm}(t) \\ \delta S^h_{nm}(t) \end{Bmatrix}, \tag{4}$$

where W^0_n is an isotropic (i.e. not depending upon order m) and stationary function that weights the harmonic coefficients. Its simple expression is as follows (e.g. Ramillien 2002):

$$W^0_n = \frac{4\pi G R \rho_w}{(2n+1)\gamma(\theta)}(1+z_n). \tag{5}$$

z_n represent the Love numbers for the elastic response of a surface-loaded Earth and $\gamma(\theta)$ is the normal gravity on the reference ellipsoid at co-latitude θ, respectively. $G \approx 6,67\ 10^{-11}$ m^3 kg^{-1} s^{-2} is the gravitational constant and R ($\approx$6378 km) is the mean radius of the Earth. ρ_W ($\approx$1000 kg m^{-3}) and ρ_E ($\approx$5517 kg m^{-3}) are mean water and Earth density, respectively.

The approximated values of $\gamma(\theta)$ are given by the theoretical formula adopted by the International Association of Geodesy (IAG) for Geodetic Reference System (Blakely 1995):

$$\gamma(\theta) = 9.7803267714 \frac{1 + 0.00193185138639\cos^2\theta}{\sqrt{1 - 0.00669437999013\cos^2\theta}}. \tag{6}$$

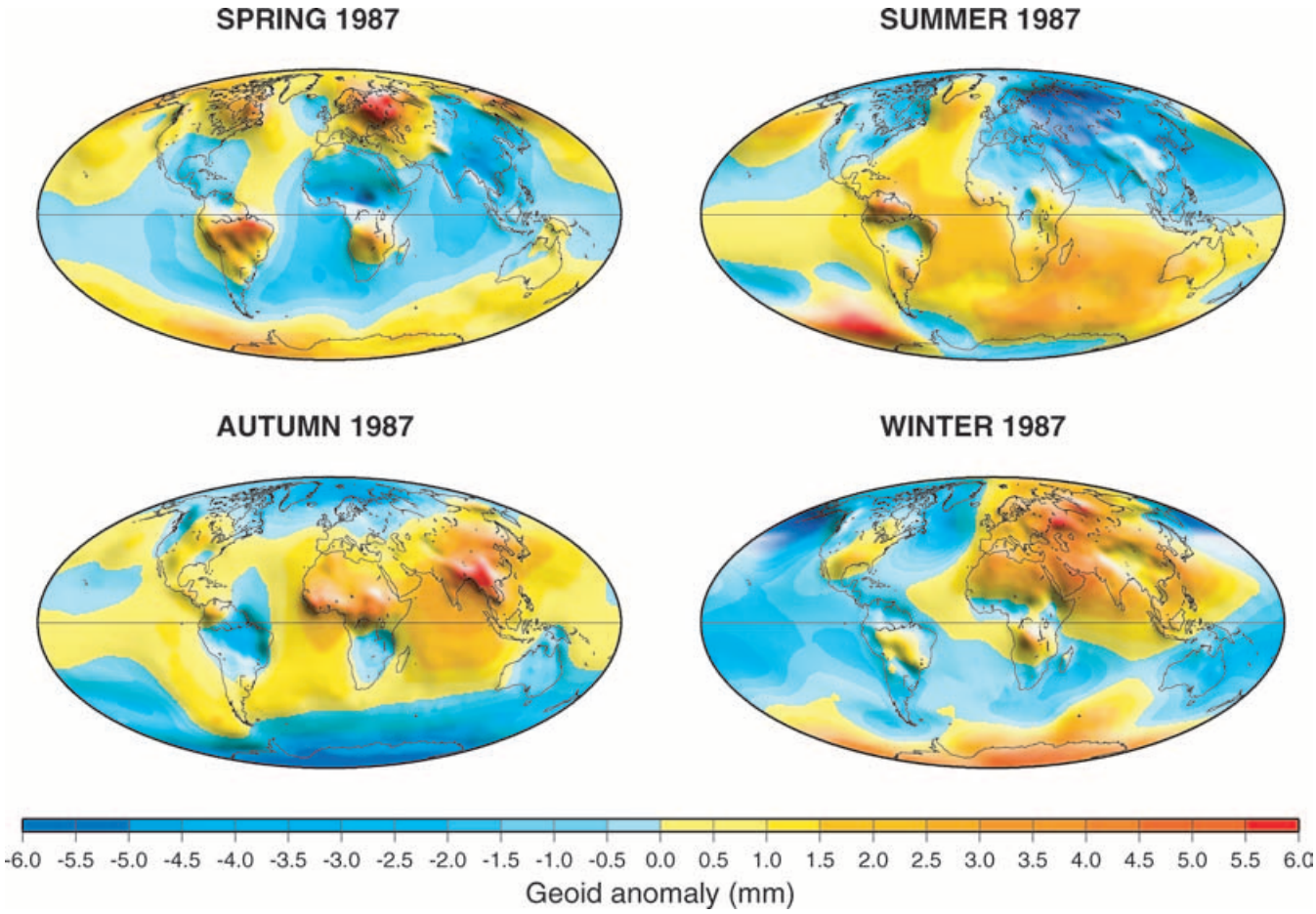

Figure 1. Seasonal variations of the synthetic geoids for 1987 computed using global model forecasts for the same period: atmosphere from ECMWF; oceans from POCM; total land water plus snow from LaD.

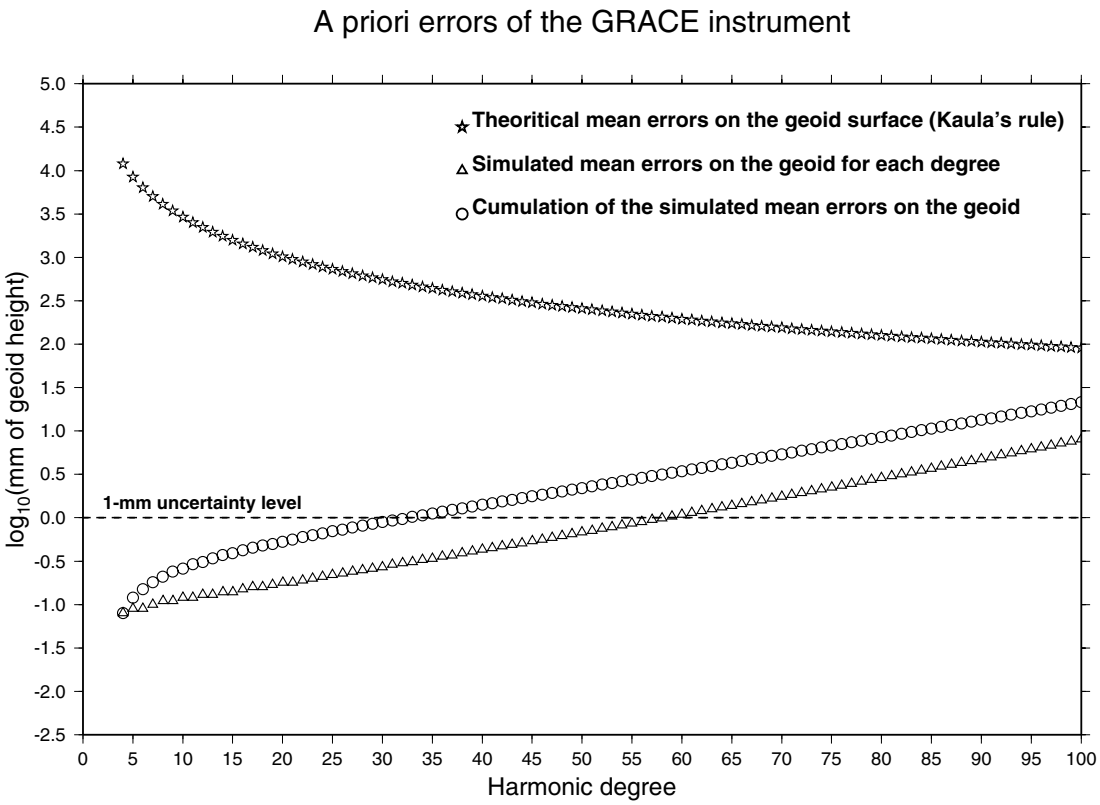

Figure 2. *A priori* error spectrum of the GRACE instrument deduced from modelling of perturbations of the GRACE orbit and analysis of co-variances (Pérosanz, private communication, 1995).

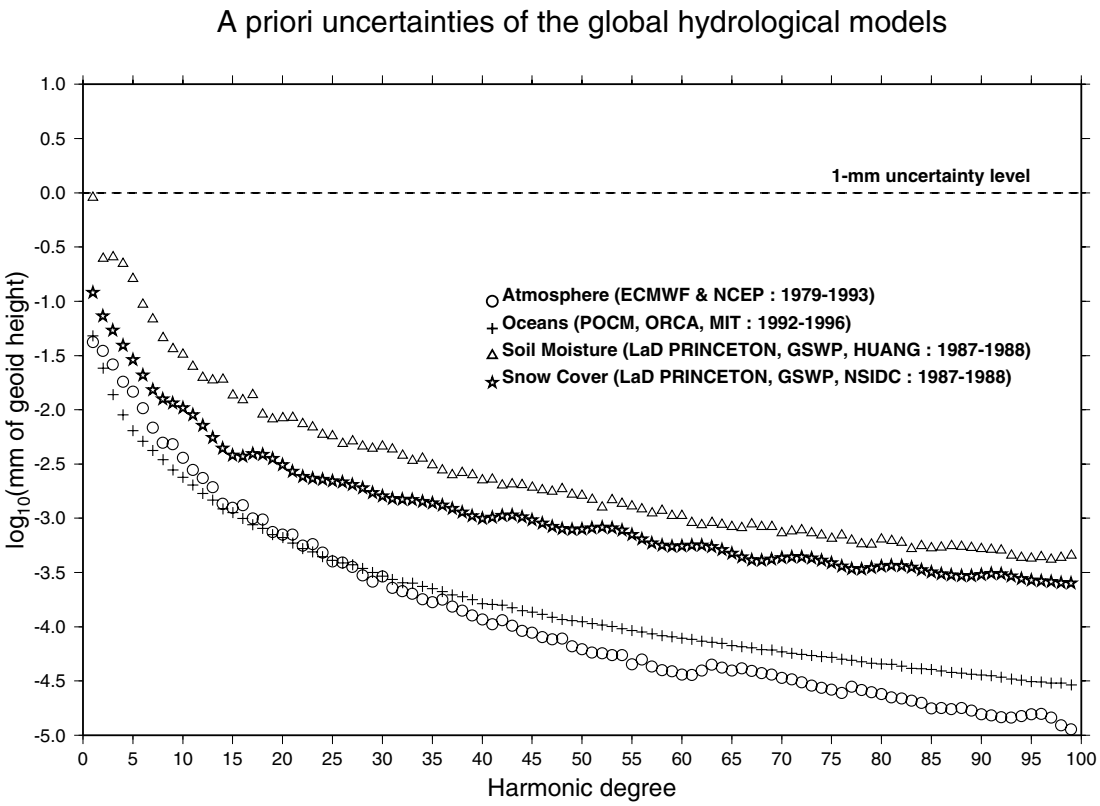

Figure 3. *A priori* uncertainties on the time variations of geoid height of the four hydrological contributions obtained by statistical comparisons of all the global model outputs available in our database: (i) atmosphere (data from ECMWF, NCEP for 1979–1993); (ii) ocean (bottom pressure from POCM, OPA, MIT for 1992–1996); (iii) soil moisture (GSWP, LaD and H96 for 1987–1988); (iv) snow cover (from GSWP, LaD and NSIDC for 1981–1987).

Another expression for the predicting filter was also proposed by Wahr *et al.* (1998):

$$W_n^0 = \frac{3\rho_w}{(2n+1)\rho_E}(1+z_n). \quad (7)$$

One can easily check that both expressions of W_n^0 (eqs 5 and 7) are numerically equivalent. Eq. (5) is used here to compute the harmonic coefficients of the synthetic geoids. The synthetic geoids themselves are derived from global climatic model outputs. The corresponding Stokes coefficients are defined as (Heiskanen & Moritz 1967):

$$\begin{Bmatrix} \delta C_{nm}(t) \\ \delta S_{nm}(t) \end{Bmatrix} = \frac{1}{(2n+1)MR^n} \times \iiint_V \left[\delta\rho(r', \theta', \lambda', t) r'^n P_{nm}(\cos\theta) \begin{Bmatrix} \cos \\ \sin \end{Bmatrix} (m\lambda') \right] dV, \quad (8)$$

where M ($\approx$5,97602 10^{24} kg) is the mass of the Earth and $\delta\rho$ (r', θ', λ'(, t') is the variable density inside the volume V of the Earth. Redistributions of mass $\delta\rho dV$ inside the volume V produces variation of the $\delta C_{nm}(t)$ and $\delta S_{nm}(t)$ coefficients. Because we consider here that the mass redistribution $\delta\rho$ (r', θ', λ') mostly occurs inside surface $\delta C_{nm}(t)$ and $\delta S_{nm}(t)$ coefficients, by $R^2\delta q\, dS$, where $\delta q(\theta'$, λ', t') is the surface load and S is the surface of the Earth. Taking into account the elastic deformation of the solid Earth under the variable load via the Love numbers z_n, the latter equation becomes (e.g. Chao *et al.* 1987; Chao 1994):

$$\begin{Bmatrix} \delta C_{nm}(t) \\ \delta S_{nm}(t) \end{Bmatrix} = \frac{(1+z_n)R^2}{(2n+1)M} \iint_S \delta q(\theta', \lambda') \begin{Bmatrix} \cos \\ \sin \end{Bmatrix} (m\lambda') P_{nm}(\cos\theta')\, dS. \quad (9)$$

3.1.1 *Atmospheric loading*

Atmospheric mass redistribution is classically deduced from atmospheric surface pressure data. The variable load $\delta q(\theta, \lambda, t)$ is related to the pressure variations $\delta p(\theta, \lambda, t)$ through:

$$\delta q(\theta, \lambda, t) = \frac{\delta p(\theta, \lambda, t)}{\rho_W \gamma(\theta)}, \quad (10)$$

where $\gamma(\theta)$ is the mean surface gravity at co-latitude θ computed using eq. (6).

3.1.2 *Ocean loading*

We will use the ocean bottom pressure data derived from ocean general circulation models (OGCM). The variable load δq is thus also given by eq. (10).

3.1.3 *Land water loading*

Land water considered here will include soil water, ground water and snow load. These three contributions will be expressed in terms

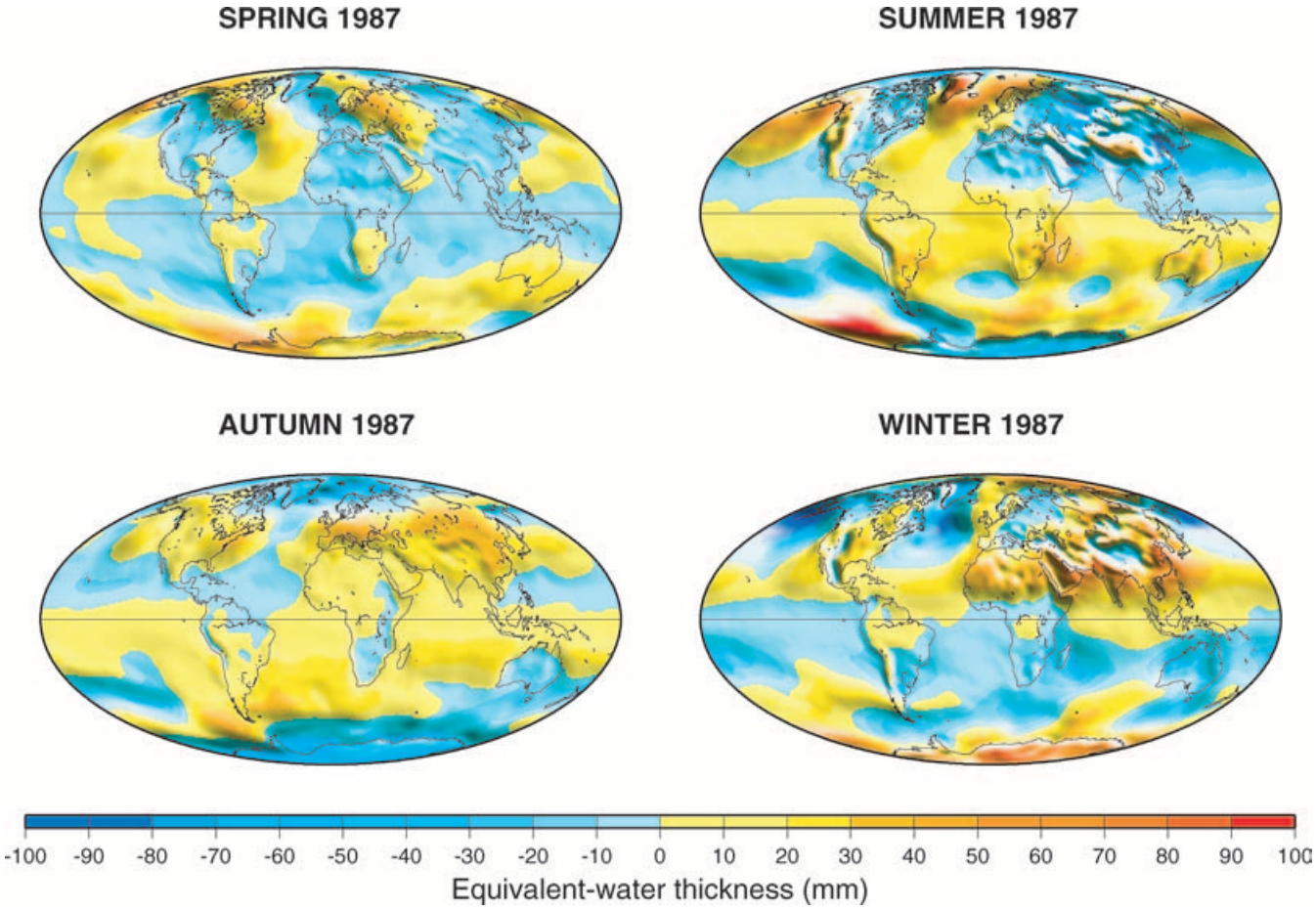

Figure 4. Example of recovery of seasonal variations of the global fluid mass for 1987 from a synthetic GRACE geoids of Fig. 1: for the atmosphere (a); the oceans (b); total continental waters including soil moisture and snow pack (c).

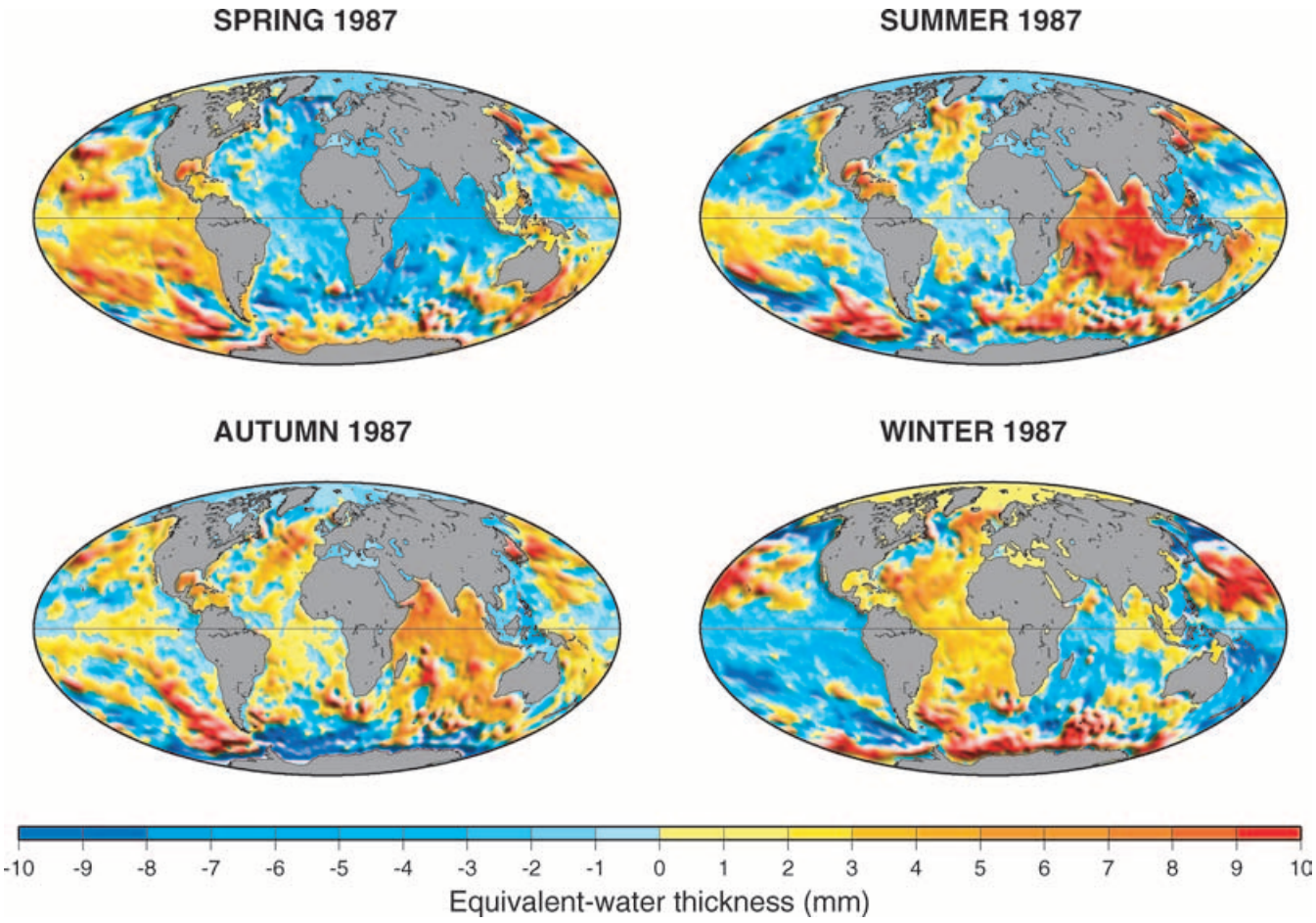

Figure 4. (*Continued*.)

of equivalent water height $\delta h(\theta, \lambda, t)$ such that corresponding load will be given by:

$$\delta q(\theta, \lambda, t) = \rho_{W,S} \delta h(\theta, \lambda, t), \tag{11}$$

where $\rho_{W,S}$ can be either liquid water or snow density.

3.2 The inverse problem: separation of several fluid contributions

The basic linear system to solve for the separate fluid contributions corresponds to eq. (3). According to this expression, there are K times more Stokes coefficients to adjust than observed ones. Therefore, the problem is highly underdetermined. For example, if we decide to separate four contributions with a maximum harmonic degree of $N = 100$, the total number of observations being $N^2 + 2N = 10\,200$, so there are ~41 600 coefficients to adjust. Because of non-uniqueness of this inverse problem, new *a priori* information need to be included as new constraints for recovering the coefficients of all water mass contributions.

In order to combine GRACE data and model prognostics, we propose to adapt the scheme previously developed by Tarantola (1987), which is based on the generalized least-square inversion for estimating the Stokes coefficients of K fluid contributions. In this formalism, the estimates are built as linear combinations of optimal fitted parameters:

$$\Gamma_k(t) = \Gamma_k^0(t) + C_k A \left(C_D + C_M + A C_k A^T\right)^{-1} \left[\Gamma^{\text{obs}}(t) - A\Gamma_k^0(t)\right], \tag{12}$$

where $\Gamma^{\text{obs}}(t)$ is the vector formed by the list of all the observed geoid coefficients and $\Gamma_k^0(t)$ represents its initial guess. C_D and C_M are the matrices of the *a priori* error covariance for the GRACE observations and the chosen model corresponding to the fluid contribution number k, respectively. C_k is the covariance matrix that describes the statistics of the water mass variations in the reservoir k, which are determined by the analysis of available models. The corresponding treatment of external constraints is described in Section 4.3.

As a result of the particular structure of the separating matrix A, eq. (12) can be split and then simplified into a system of two dependent equations:

$$\Gamma_k(t) = \Gamma_k^0(t) + C_k \xi, \tag{13}$$

where the vector ξ is the solution of the expression

$$\left(C_D + C_M + A C_k A^T\right) \xi = \Gamma^{\text{obs}}(t) - A\Gamma_k^0(t). \tag{14}$$

Moreover, the *a posteriori* covariance matrix is computed using:

$$C_k' = C_k - C_k A^T \left(A C_k A^T + C_D + C_M\right)^{-1} A C_k. \tag{15}$$

The *a posteriori* uncertainty associated with each fitted Stokes coefficient are given by the root-mean square of the diagonal elements of this latter matrix:

$$\sigma_k(t) = \sqrt{C_k'(\text{diag})}, \tag{16}$$

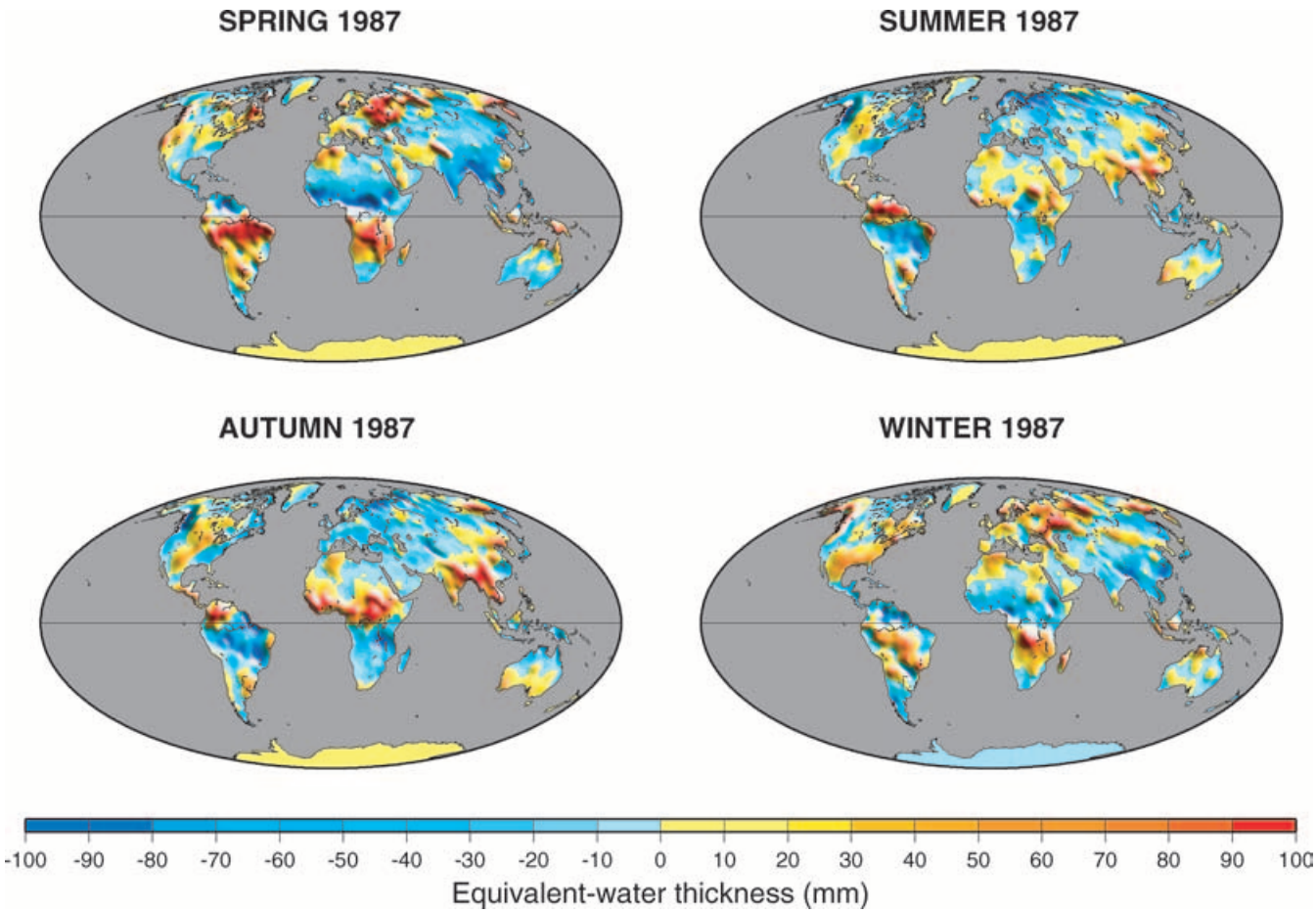

Figure 4. (*Continued*.)

where diag stands for individual diagonal elements of the C'_k matrix. Note that in the latter expressions, $A^T C_k A$ represents the sum of the model covariance matrices C_k.

In practice, the left-hand matrix in brackets to inverse in eq. (14) is always symmetric by construction and often definite positive. Here, the fast procedure of the Cholesky factorizations (Press *et al.* 1998) can be used to find ξ. However, as a result of even numerical instabilities of high-degree coefficients (i.e. short wavelengths of the geoid) estimated from the models and/or bad matrix conditioning, this configuration matrix is no longer definite positive and sometimes ill-conditioned. In these particular cases, one can use either schemes of pre-conditioning, singular value decomposition (SVD), or lower–upper (LU) factorization to ensure to find a stable solution ξ. This solution vector is replaced in eq. (13) to estimate the Stokes coefficients afterwards.

Then the corresponding surface density coefficients, expressed in terms of surface load variations, are computed using an inverse filtering:

$$\begin{Bmatrix} \delta C^h_{nm}(t) \\ \delta S^h_{nm}(t) \end{Bmatrix} = W_n^{-1} \begin{Bmatrix} \delta C_{nm}(t) \\ \delta S_{nm}(t) \end{Bmatrix}, \tag{17}$$

where W_n^{-1} is the predicting filter constructed as a modified version of the one from eq. (5) such as:

$$W_n^{-1} = \frac{V_n}{W_n^0}, \tag{18}$$

if V_n is a stabilizing function that tapers the amplitudes of the filter in the high degree for avoiding the development of short-wavelength instabilities in the prediction.

For example, this latter low-pass function can be defined as:

$$V_n = 1 \text{ if } n \le n_1, \quad V_n = 0 \text{ if } n \ge n_2 \quad \text{and}$$
$$V_n = \frac{1}{2}\left[1 + \cos\left(\pi \frac{n - n_1}{n_2 - n_1}\right)\right] \text{ otherwise.} \tag{19}$$

This attenuates slowly the values of the predicting operator at the cutting degree n_1.

The disadvantage of smoothing the spatial resolution is to remove short-wavelength features in the solution, as previously pointed out by Wahr *et al.* (1998), who used a disc function to recover mean soil water variations in continental basins.

4 DATA USED AS NEW INDEPENDENT STATISTICAL CONSTRAINTS

4.1 Global climatic data sets

As statistical constraints, as well as for generating synthetic GRACE geoids, we used several global data sets corresponding to the four contributions considered in this study: atmosphere, oceans, total soil water and snow cover.

4.1.1 Atmosphere

Global maps of surface pressure from 6-hr re-analysis of ground observations coupled with equations of general circulation are

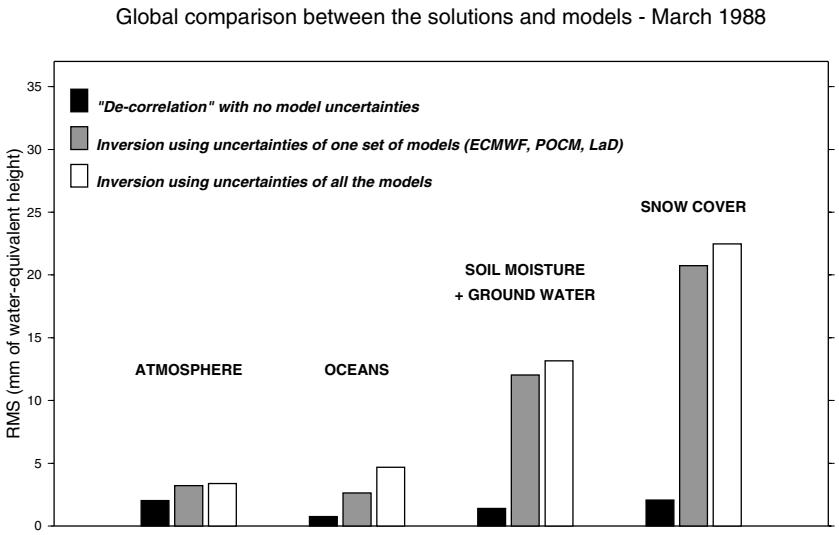

Figure 5. Histograms of global rms errors for the four water mass estimates. Errors of prediction are of a few mm rms of water thickness neglecting the *a priori* model uncertainties, whereas they reach tens of mm while these statistical information are input in the inversion.

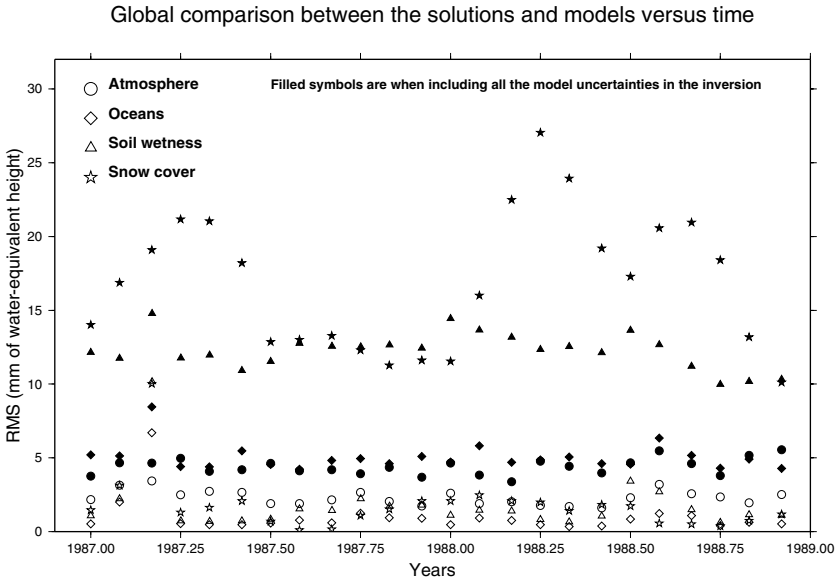

Figure 6. Rms differences between input model data and global solutions versus time (1987–1988), and for fluid mass variations of each contribution: atmosphere (circles), oceans (diamonds), soil water (triangles) and snow cover (stars). Filled symbols are for inversions including *a priori* uncertainties of all the models.

currently provided by National Centers for Environmental Predictions (NCEP) and European Centre for Medium Weather Forecast (ECMWF). Errors are around ∼1 mbar on the average for a monthly gridded value, but this uncertainty may be regionally and locally more important and correlated with the station distribution over the globe. The quality of the data is poor in some regions like Antarctica where observations are very sparse and sometimes non-existent. Monthly mean surface pressure fields are available from 1979 to 1996 for NCEP and from 1979 to 1993 for ECMWF.

4.1.2 Oceans

We used ocean bottom pressure data provided by several OGCM: Parallel Ocean Circulation Model_4C (POCM_4C), Océan PArallélisé (OPA) and Massachussets Institute of Technology (MIT), that are deduced from analysis of global oceanic circulation combined with surface and subsurface observations.

(i) POCM_4C: the model is described in Semtner & Chervin (1992), Stammer *et al.* (1996) and Tokmakian (1996). The model is run from 1979 to 1998 forced by daily ECMWF re-analysis of heat, freshwater flux (evaporation minus precipitation and river run-off) and wind stress. The model resolution is approximately 1/4° latitude and longitude, permitting resolution of the western boundary currents. As of 1994, the forcing is prescribed by ECMWF operational forecasts. A correction to heat and fresh water flux is applied within 20° of the equator to conserve global net annual heat and freshwater fluxes in the model. POCM_4C is forced by ECMWF

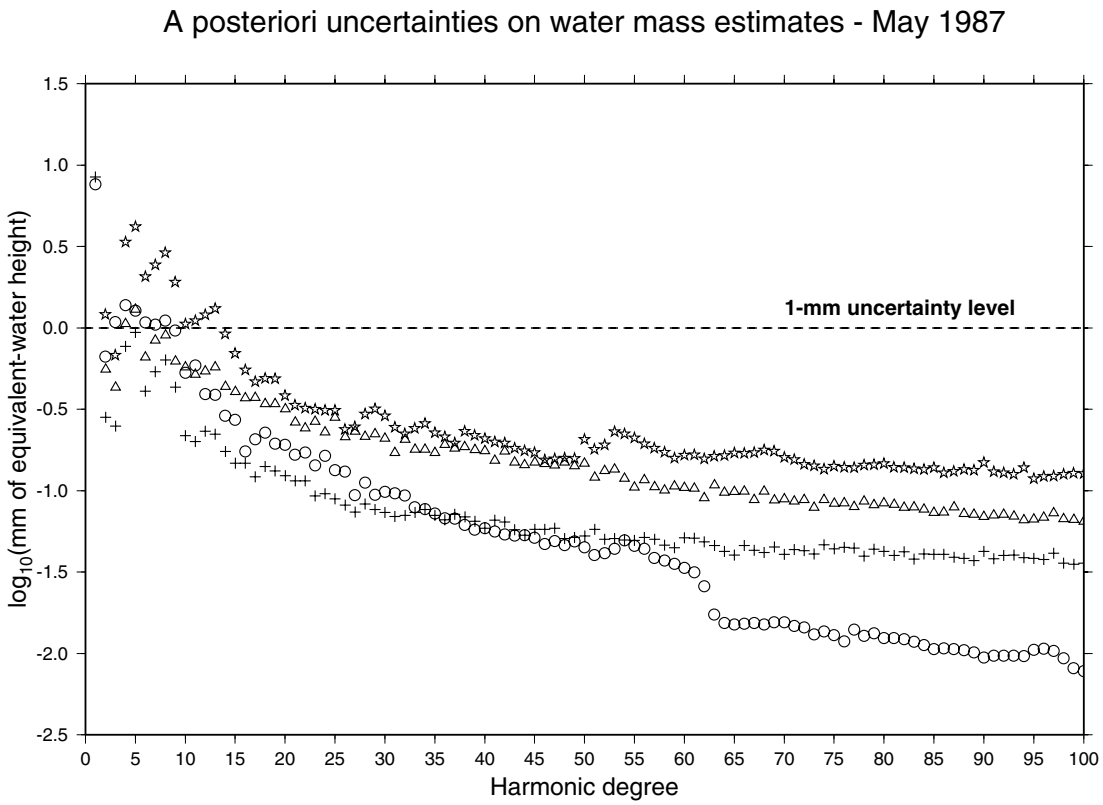

Figure 7. *A posteriori* uncertainties (a) and final errors (b) versus degree for each estimated water mass component of 1987 May: atmosphere (circles), oceans (crosses), soil moisture (triangles) and snow cover (star). Errors are computed as the differences between model and predicted coefficients of water heights. Uncertainties for continental water storage are greater than ∼1 mm of water thickness for low-degree coefficients but they decrease rapidly. Combining different global models for estimating the four water reservoirs leads to errors >10 mm on each water mass coefficient.

re-analysed wind from 1979 to 1993 and by ECMWF operational winds for 1994–1997.

(ii) OPA from Laboratoire d'Océanographie DYnamique et de Climatologie (LODYC): this is a global ocean model solved in Z coordinates on a modified world grid whose poles are situated on land to avoid singularities in the calculation (Madec & Imbard 1996; Madec *et al.* 1999). The model is forced by the European Remote Sensing (ERS) satellite-derived surface winds and ECMWF heat and freshwater fluxes. The modelled period is 1992–2000 inclusive. The correction to bottom pressure as a result of the Boussinesq approximation is applied (Greatbatch 1994).

(iii) MIT: the model is described in Marshall *et al.* (1997a,b, 1999, 2001). Model resolution is 1°. The model is forced with real NCEP parameters (daily heat and fresh water fluxes; twice daily wind stress) and relaxed to monthly Levitus mean temperature and salinity fields. The model was run from 1985 to 1995.

4.1.3 Land water mass

Land water mass data used in this study are based on outputs of global hydrological models. The data consist of water stored in the root-zone (i.e. SM), water in shallow aquifers (underground water) and snow. We used data from three models:

(i) The Huang *et al.* (1996) model, further denoted as H96. This model provides only soil water for the upper 2-m-thick surface layer. Neither snow nor ground waters are calculated by the model. The time span of data covers 1979–1998. Grid size is 1.875° in longitude and 1.9° in latitude. This is an old generation model in that it uses Thornthwaite's empirical equation for estimating evapotranspiration and monthly meteorological forcing. Besides, the soil wetness data present (unexplained) temporal discontinuities, occurring at the beginning of the year, although not systematically. We considered this model nevertheless because it covers a long time span.

(ii) The Land Dynamics (LaD) model developed by Milly & Shmakin (2002). LaD provides global (exclusive of Antarctica and Greenland) monthly 1° × 1° gridded time series of snow depth, SM and underground water for 1981–1998.

(iii) Global Soil Wetness Project (GSWP) developed by Douville *et al.* (1999) in the context of the GSWP project and based on the Interactions between Soil Biosphere and Atmosphere (ISBA) land surface scheme (Noilhan & Planton 1989). This model provides monthly 1° × 1° gridded time-series of snow depth and SM for two years (1987–1988). As for LaD, GSWP also excludes Greenland and Antarctica.

Main characteristics of all these global models are summarized in Table 1.

4.2 Construction of the synthetic monthly geoids

The steps for generating synthetic monthly geoids are:

(i) linear interpolation of the model values onto a global 1° grid if necessary;

(ii) conversion of the gridded model values into surface load variations in terms of equivalent-water heights;

(iii) expansion of the gridded data from step (ii) into spherical harmonics, up to degree and order 100;

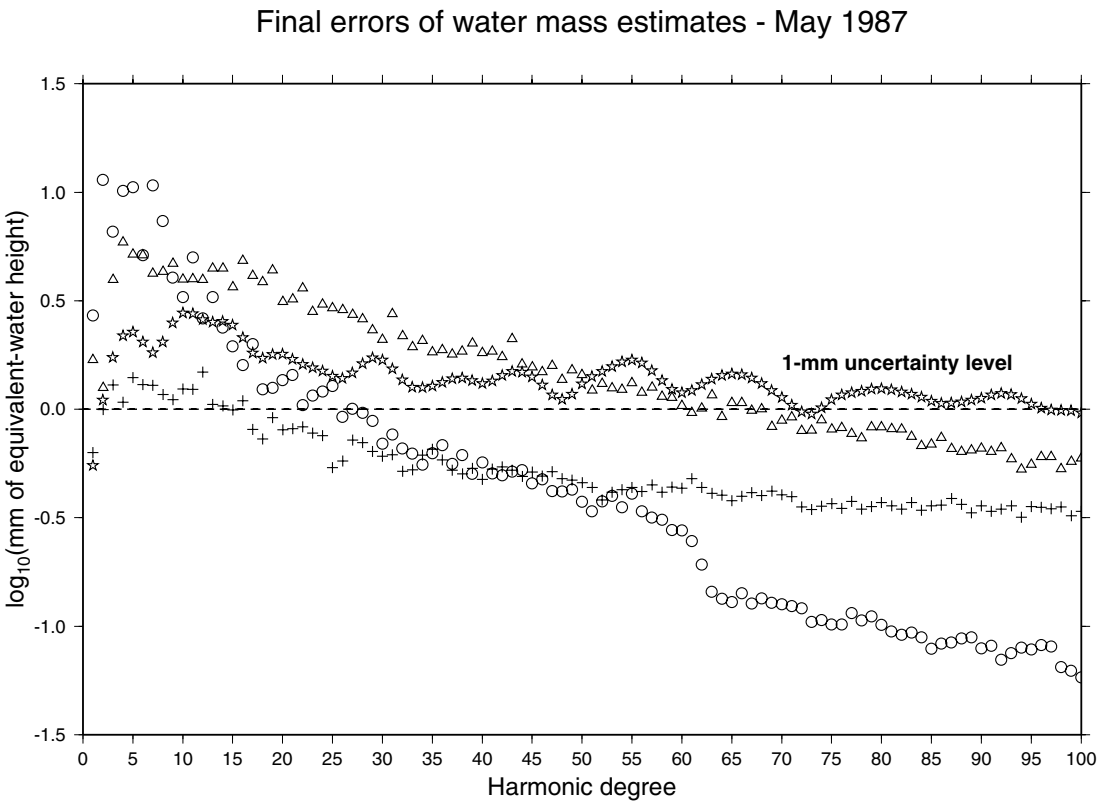

Figure 7. (*Continued*.)

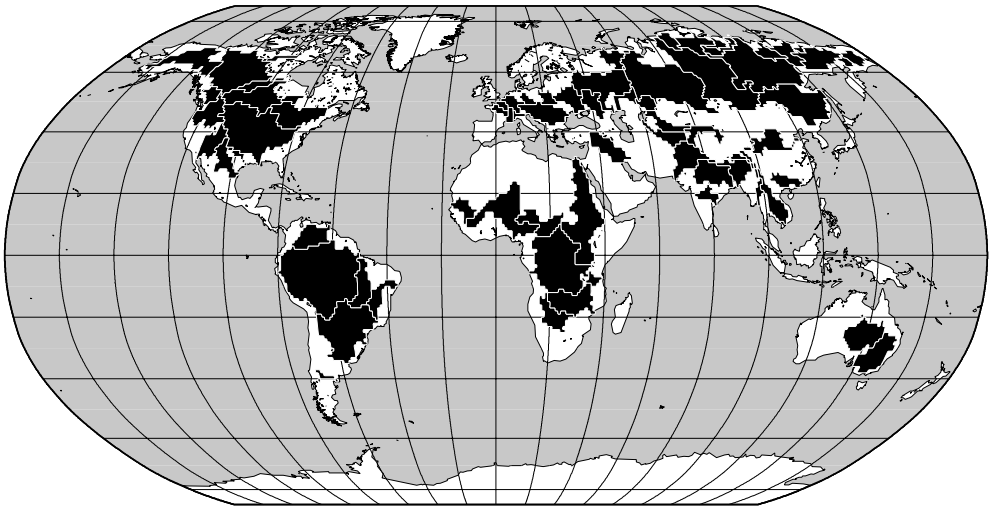

Figure 8. Geographical locations of the 71 main drainage basins in the world. These are mainly located in the Northern Hemisphere. General geographical information about each of them are provided in Table 2.

(iv) filtering the corresponding geoid coefficients using eq. (4);

(v) estimate of the time-average of geoid coefficients over the time span of each climate model data and removing it to each monthly geoid coefficient;

(vi) computation of the sum of the coefficients of four contributions (atmosphere, oceans, soil water and snow) for a given month, over a given time span.

The starting monthly synthetic geoids were constructed using the following climatic fields:

(i) Atmosphere: surface pressure fields from ECMWF.

(ii) Ocean bottom pressure from POCM_4C.

(iii) Continental waters (soil and ground waters plus snow) from LaD.

We consider below two complete years 1987–1988, which is the common time span of the above models. Fig. 1 shows for the four seasons of 1987 the synthetic geoids constructed with these data. In these maps, the dominant signals arise from atmosphere and continental hydrology, which is about 10 times larger than the oceanic signals.

Table 2. General information on the 71 drainage basins that are considered in the present study. They are listed by increasing size order.

Name	Area (millions of km^2)	Mean latitude (degree)	Name	Area (millions of km^2)	Mean latitude (degree)
Amazon	6,035	−7	Senegal	0,3830	15
Congo	3,587	−3	Don	0,3790	50
Mississippi	3,091	41	Tarim	0,3653	39
Ob	2,862	57	Syr Darya	0,3507	43
Paraná-Uruguay	2,831	−24	Xi	0,3487	24
Nile	2,788	13	Khatanga	0,3189	71
Yenissei	2,472	59	Indigirka	0,3087	67
Lena	2,221	62	Godavari	0,2779	19
Niger	1,952	14	Pechora	0,2741	65
Amur	1,753	50	Salween	0,2341	26
MacKenzie	1,651	61	Churchill	0,2151	56
Volga	1,298	56	Tugaj	0,2120	50
Zambeze	1,235	−15	Yana	0,1973	67
Lake Eyre	1,162	−26	Fraser	0,1938	52
Saint-Lawrence	1,022	46	Rhin	0,1900	49
Nelson	1,012	52	Ural	0,1875	51
Murray	0,9893	−32	Anadyr	0,1710	65
Ganges	0,9490	26	Vistula	0,1645	51
Indus	0,8854	31	Olenek	0,1626	69
Orinoco	0,8315	6	Orange	0,1626	69
Chari	0,8185	11	Rufigi	0,1510	−8
Araguaia	0,7855	−10	Paranaiba	0,1487	72
Yukon	0,7665	64	Pyasina	0,1487	72
Danube	0,7360	46	Albany	0,1189	51
Okavango	0,7143	−19	Elbe	0,1136	51
Mekong	0,6930	19	Loire	0,0939	47
Euphrates	0,6843	35	Brazos	0,0873	33
Huanghe	0,6787	37	Anabar	0,0830	71
Columbia	0,6548	46	Rhône	0,0828	45
Kolyma	0,5826	65	Negro	0,0804	−39
Colorado-Arizona	0,5758	37	Back	0,0726	65
Brahmapurta	0,5742	28	Seine	0,0643	49
Rio Grande	0,5429	31	Penzina	0,0630	64
São Francisco	0,5325	−13	Alazeja	0,0449	69
Dniepr	0,4634	51	Po	0,0407	46
Amu Darya	0,4285	39			

4.3 Covariance matrices construction

We assume that the surface fluid fields correspond to stationary processes (i.e. their statistical properties, such as mean and standard deviation, are stationary with time). We define as $D_k(\Delta t)$ the matrix formed by the list of all the Stokes coefficients up to the maximum degree $N = 100$, related to the global fluid model number k, for a period of analysis of Δt months. By construction, this matrix is such that each row corresponds to a particular month and each column to a given coefficient of degree n and order m. An estimation of the covariance matrix C_k related to the kth contribution for this period of time is then:

$$C_k = \left[D_k(\Delta t) - \bar{D}_k\right]^T \left[D_k(\Delta t) - \bar{D}_k\right], \tag{20}$$

where $\bar{D}_k$ is the time-mean values of the model coefficients computed during Δt. Several tests have shown that Δt should be ∼2–3 months around the chosen month t of GRACE observations to be inverted. Greater values of Δt (i.e. >3 months) lead to numerical smoothing and so provide less precise geoid solutions.

In order to take into account the spatial correlation between couples of distinct harmonic coefficients of degrees n and u, and orders m and v, the elements of the covariance matrix C_k are merely multiplied by the weighting function:

$$\eta(n, m, u, v) = (2|n - u| + 1)^{-1}(2|m - v| + 1)^{-1}. \tag{21}$$

Estimates of *a priori* of GRACE instrumental errors versus harmonic degree (Fig. 2) related to the covariance matrix C_D were provided by simulations of errors of orbit parameters and covariance analysis (e.g. Pérosanz 1995). *A priori* model uncertainties of the matrix C_M are estimated from statistical comparisons of the different climatic models of each contribution. *A priori* errors on each model harmonic coefficients (i.e. diagonal elements of C_M) were computed as the temporal variances of these coefficients, so that important variances indicate that global models, such as for soil waters, remain very different. Model covariance matrices representing fluid mass contributions were evaluated using eq. (20) for nominal periods of 2–3 months of analysis and centred at the considered month. For this purpose, we used all the available climate model data sets described in Table 1 and we considered their own complete non-coincident time spans in the computation of the variances. Fig. 3 shows for each contribution the *a priori* uncertainty spectra computed as explained above. According to the models used here, we note that the largest errors arise from continental hydrology. We also notice that beyond harmonic degree 30, the atmosphere represents the smallest source of error.

5 RESULTS

In this section, we present the results of the monthly inversion according to Section 3.2, over 24 months (1987–1988) of the synthetic geoids constructed in Section 4.2. Figs 4(a)–(c) present the four seasons of the inversion solutions for the atmosphere, ocean and total continental waters (soil plus ground waters and snow). All the maps are expressed in equivalent-water height for purpose of comparison. Note that for the ocean bottom and atmospheric pressures, ∼10 mm of water height is equivalent to 1 hPa. Comparison with the input data (not shown) reveals very good agreement. To illustrate this, Fig. 5 presents for a single month (1987 May), the histograms of the differences between the input model data and solutions for three cases: (a) no model uncertainty are input; (b) only one model for each fluid contribution is used to compute the model errors (ECMWF for atmosphere; POCM for oceans; LaD for continental waters: SM + ground water and snow depth); (c) all the available models are used to estimate *a priori* model errors (note that the set of solutions presented in Fig. 4 corresponds to this latter case). As expected, the ideal (but unrealistic) case (a) provides the smallest differences. Case (c) is likely the most physically plausible result from the inversion. We note that the largest uncertainty arises from the continental water reservoirs.

Differences between global monthly solutions and corresponding model prognostics versus time (1987–1988) are presented on Fig. 6. It shows that: (i) they are more severe while including model uncertainties than neglecting this type of error; (ii) once again, the largest *a posteriori* errors are for soil waters and snow cover.

Monthly residuals were also computed as the difference between the input model coefficients and the sum of the estimates for each degree and order. They represent the errors of de-correlation by combining satellite gravity data and model forecasts.

Residuals should be close to zero if perfect separation of different geoid contributions was successful. However, this is not the case in practice, first because of errors resulting from numerical approximations, as well as instability problems with the condition of the co-variances matrices used in the inversion. This causes errors intrinsic to the method that are revealed by analysing *a posteriori* uncertainties versus harmonic degree. These uncertainties remain important for degrees up to 10–15 and then decrease to less than 1-mm water height. Figs 7(a)–(b)present, respectively, *a posteriori* error spectra and the difference between input data and solutions for each harmonic degree. Using different sets of climate models for computing *a priori* co-variances generates a significant long-wavelength bias as well and increases the errors related to the geoid and hence water mass estimates.

6 APPLICATION FOR RECOVERING SOIL WATER VARIATIONS IN DRAINAGE BASINS OF THE WORLD

In this section, we explore the ability of GRACE in monitoring regional variations of soil water in 71 drainage basins as listed in Table 2 and shown in Fig. 8. The time-series of soil water (sum of SM and underground water) in each basin were merely extracted from the global monthly gridded solutions.

Monthly mean soil water solutions were compared to the corresponding mean values interpolated from the LaD model used as input. Histograms of rms differences are plotted versus basin area in Figs 9(a)–(b). There is no clear relationship between the errors for soil water recovery and the size of the basin. A similar conclusion is drawn with respect to the mean latitude of the drainage basin. This

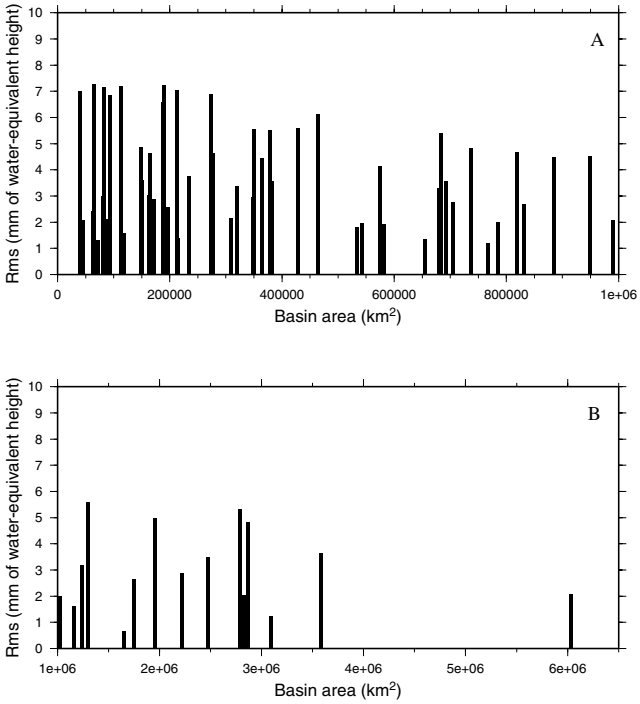

Figure 9. Rms differences between model and computed soil moisture time-series versus basin area of less than 1×10^6 km^2 (a) and greater than 1×10^6 km^2 (b).

implies that: (i) our reconstruction of soil water is geographically homogeneous because we used a global harmonic decomposition on the whole terrestrial sphere, and (ii) each large drainage basin keeps mainly its own characteristics, which do not depend upon mean latitude.

Nevertheless, drainage basins in arid regions, like Euphrates and Syr Darya basins, are generally associated with small errors (a few mm in water height only). Differences typically range from <1 to 16 mm rms in water height, suggesting that soil water solution variations remain quite close to the forecasts made by the LaD model.

Profiles of soil water solutions in the largest drainage basins of South America such as the Amazon basin are plotted on Fig. 10 where the seasonal cycle and interannual variations are clearly visible. Error bars of associated *a posteriori* uncertainty vary with time, probably a result related to the high variability of soil water signals in each basin.

7 CONCLUSION

In this paper, we successfully applied a new least-square method for solving a very unconstrained gravity inverse problem of decorrelation of fluid contributions (i.e. atmosphere, oceans, continental waters) by introducing independent information from different climate models. Instrumental errors and other *a priori* uncertainties on GRACE data and the physical model were included in this least-square approach.

Final errors on fitted harmonic coefficients of mass of each fluid component are of a few mm of equivalent-water height when no *a priori* uncertainty is included in the inversion. Besides, they are more severe (reaching tens of mm) if the starting errors of all the available global models are combined in the process, because the prognostics of the models remain too different.

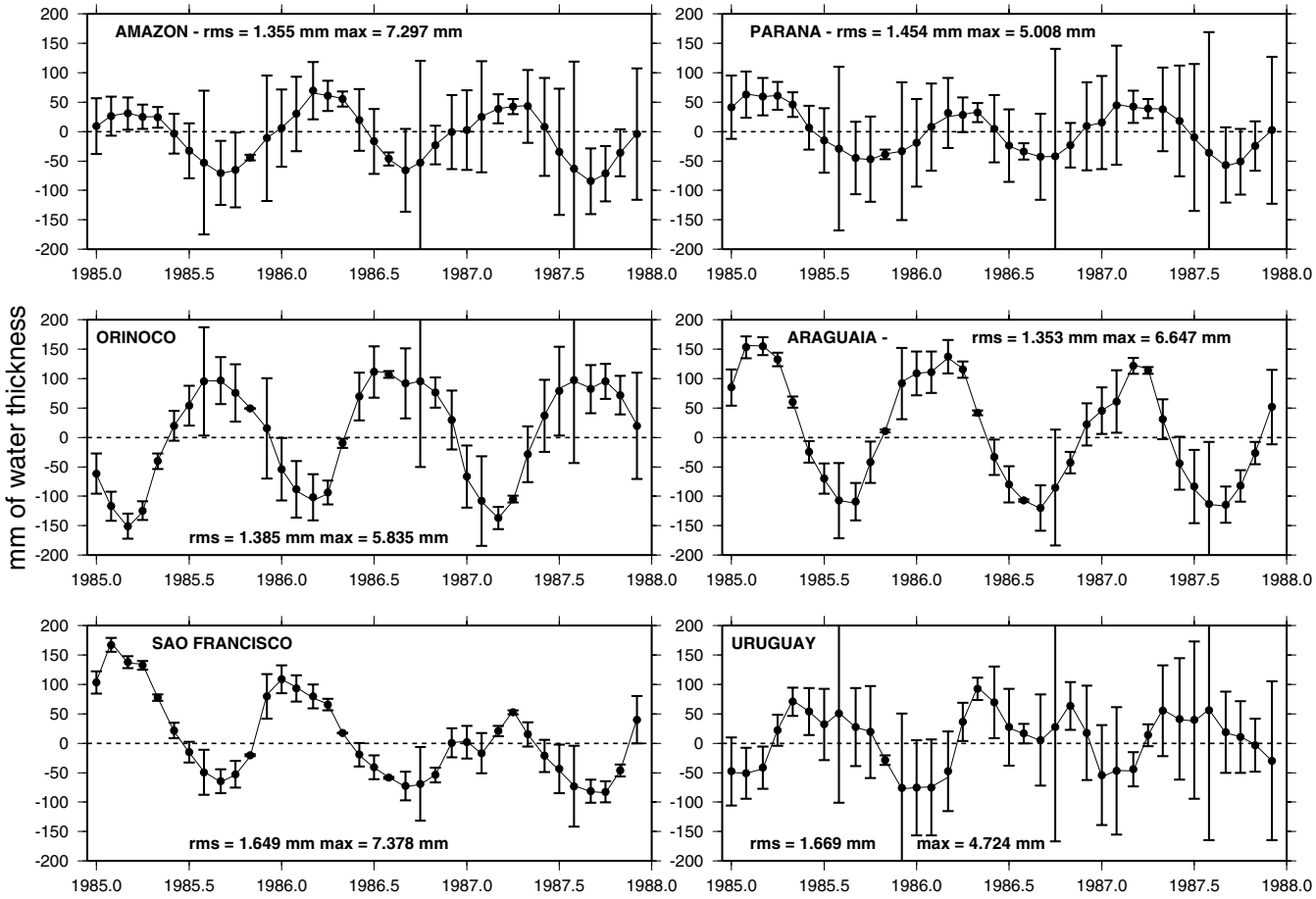

Figure 10. Soil moisture time-series (1985–1987) for six river basins in South America (Amazon, Paraná, Orinoco, Araguaia, São Francisco and Uruguay) computed by the least-square inversion (bold lines) and error bars. For comparison, superimposed (solid lines) is the mean soil moisture value obtained by averaging data from three global hydrological models: H96, LaD and GSWP.

Technically, our inverse approach for separating fluid contributions would be improved by:

(i) Introducing more realistic spatio-temporal correlations between the fluid mass harmonic coefficients through the separating matrix A (because there are obvious exchanges of mass and thus correlations between the water reservoirs).

(ii) Including more independent information like satellite/*in situ* data in the co-variance matrices such as extracting the ground water variations from the total water storage on the continents. In oceanography, GRACE data combined with satellite altimetry data will allow the extraction of ocean mass changes related to the eustatic from the altimetry-observed contributions providing information on the large-scale water balance.

(iii) Extending the time spans of the available global data sets we have, which presently remain an important limiting factor of our method, especially to the recent period of measurements made by GRACE.

ACKNOWLEDGMENTS

The authors are very grateful to B. J. Chao for his constructive review of the paper. The authors thank M.-C. Gennero for having kept the GEOFLUIDS database at our disposal, as well as V. Toumazou who initially developed the very first version of the code of inversion in Matlab and provided us fruitful advices on the methodology.

REFERENCES

Blakely, R.J., 1995. *Potential Theory in Gravity & Magnetic Applications,* Cambridge University Press, Cambridge, p. 441.

Chao, B.F., 1994. *The Geoid and Earth Rotation, Geophysical Interpretation of the Geoid,* eds Vanicek, P. & Christou, N., CRC Press, Boca Raton.

Chao, B.F., O'Connor, W.P., Chang, A.T.C., Hall, D.K. & Foster, J.L., 1987. Snowload effect on the Earth's rotation and gravitational fielg, 1979–1985, *J. geophys. Res.,* **92,** 9415–9422.

Dickey, J. & National Research Council Commission (NRC), 1997. *Satellite Gravity & The Geosphere,* National Academy Press, Washington DC, p. 112.

Douville, H., Basile, E., Caille, P., Giard, D., Noilhan, J., Peirone, L. & Taillefer, F., 1999. Global soil wetness project : forecast and assimilation experiments performed at Météo-France, *J. Meteorlogical Soc. of Japan,* **77**(1B), 305–316.

European Space Agency (ESA), 2000. *From Eötvös to Milligal*, Final Report, No. 13392/98/NL/GD.

Greatbatch, R., 1994. A note on the representation of steric sea level in models that conserve volume rather than mass, *J. geophys. Res.,* **99,** 12 767–12 771.

Heiskanen, W.H. & Moritz, H., 1967. *Physical Geodesy,* WH Freeman & Co, San Francisco.

Huang, J., van den Dool, H. & Georgakakos, K.P., 1996. Analysis of model-calculated soil moisture over the United Sates (1931–1993) and application to long range temperature forecasts, *J. Climate,* **9,** 1350–1362 (H96).

Jekeli, C., 1981. *Alternative method to smooth the Earth's gravity field,* Rep. 327, Dep. Of Geod. Sci. & Surv., Ohio State University, Columbus.

Madec, G. & Imbard, M., 1996. A global mesh to overcome the North pole singularity, *Clim. Dyn.,* **12,** 381–388.

Madec, G., Delecluse, P., Imbard, M. & Lévy, C., 1999. *OPA 8.1 Ocean General Circulation Reference Model,* Tech. Rep. 10, Institut Pierre Simon Laplace, Paris.

Marshall, J., Adcroft, A., Hill, C., Perelman, L. & Heisey, C., 1997a. A finite-volume, incompressible Navier-Stokes model for studies of the ocean on parallel computers, *J. geophys. Res.,* **102(C3)**, 5753–5766.

Marshall, J., Hill, C., Perelman, L. & Adcroft, A., 1997b. Hydrostatic, quasi-hydrostatic and non-hydrostatic ocean modeling, *J. geophys. Res.,* **1032(C3)**, 5733–5752.

Marshall, J., Jamous, D. & Nilsson, J., 1999. Reconciling 'thermodynamic' and 'dynamic' methods of computation of water-mass transformation rates, *Deep Sea Res. I,* **46,** 545–572.

Marshall, J., Nilsson, J. & Jamous, D., 2001. Entry, flux and exit of potential vorticity in ocean circulation, *J. Phys. Oceanogr.,* **31(3)**, 777–789.

Milly, P.C.D. & Shmakin, A.B., 2002. Global modeling of land water and energy balances: 1. The Land Dynamics (LaD) model, *J. Hydrometeorology,* **3,** 283–299.

Noilhan, J. & Planton, S., 1989. A simple parametrization of the land surface processes for meteorological models, *Monthly Weather Review,* **117,** 536–549.

Pérosanz, F., 1995. Utilisation des measures GPS pour la restitution dynamique precise d'orbites et l'amélioration des modèles globaux de champs de gravité terrestre. Application au satellite Topex/Poséïdon et à la simulation des performances des futures missions géodésiques, *PhD thesis*, Manuscript No. 2087, Université Paul Sabatier, Toulouse.

Press W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P., 1998. *Numerical Recipes in C,* Cambridge University Press, Cambridge, p. 994.

Rodell, M. & Famiglietti, J.S., 1999. Detectability of variations in continental water storage from satellite observations of the time dependent gravity field, *Water Resources Res.,* **35(9)**, 2715–2723.

Rodell, M. & Famiglietti, J.S., 2001. An analysis of terrestrial water storage variations in Illinois with implications for the Gravity Recovery and Climate Experiment (GRACE), *Water Resources Res.,* **37(5),** 1327–1339.

Rodell, M. & Famiglietti, J.S., 2002. The potential for satellite-based monitoring of groundwater storage changes using GRACE: the High Plains Aquifer, *Central US Journal of Hydrology,* **263,** 245–256.

Semtner, A.J. & Chervin, R.M., 1992. Ocean general circulation from a global eddy-resolving model, *J. geophys. Res.,* **97,** 5493–5550.

Stammer, D., Tokmakian, R., Semtner, A. & Wunsch, C., 1996. How well does a 1/4° global circulation model simulate large-scale oceanic observations?, *J. geophys. Res.,* **101,** 25 799–25 811.

Swenson, S. & Wahr, J., 2002. Methods for inferring regional surface-mass anomalies from Gravity Recovery and Climate Experiment (GRACE) measurements of time variable gravity, *J. geophys. Res.,* **107(B9)**, ETG 1–13.

Tarantola, A., 1987. *Inverse Problem Theory,* Elsevier, Amsterdam, p. 613.

Tokmakian, R., 1996. Comparisons of time series from two global models with tide-gauge data, *Geophys. Res. Lett.,* **23,** 3759–3762.

Velicogna, I., Wahr, J. & Van den Dool, H., 2001. Can surface pressure be used to remove atmospheric contributions from GRACE data with sufficient accuracy to recover hydrological signals?, *J. geophys. Res.,* **106(B8)**, 16 415–16 434.

Wahr, J., Molenaar, M. & Bryan, F., 1998. Time variability of the Earth's gravity field: hydrological and oceanic effects and their possible detection using GRACE, *J. geophys. Res.,* **103(B12)**, 30 205–20 229.