Schmeer M., M. Schmidt, W. Bosch, F. Seitz: Separation of mass signals within GRACE monthly gravity field models by means of empirical orthogonal functions. Journal of Geodynamics, 59, 124-132, 2012.

DOI: 10.1016/j.jog.2012.03.001

Separation of mass signals within GRACE monthly gravity field models by means of Empirical Orthogonal Functions

Martin Schmeer[a], Michael Schmidt[b], Wolfgang Bosch[b], Florian Seitz[c]

[a]*German Research Centre for Geosciences, Helmholtz Centre Potsdam (GFZ), Germany*
[b]*German Geodetic Research Institut (DGFI), Germany*
[c]*Earth Oriented Space Science and Technology (ESPACE), Technische Universität München, Munich, Germany*

Abstract

Since 2002 the two GRACE satellites observe the time varying gravity signal mainly caused by the sum of mass variations within the Earth subsystems ocean, atmosphere, and continental hydrosphere. It is a challenging problem to separate the integral GRACE signal and to identify and quantify the mass variations of the individual subsystems. This work proves first by a closed loop simulation that such a decomposition is successful by means of empirical orthogonal functions (EOF) derived from geophysical models and a least-squares adjustment with a multivariate Gauss-Markov model with time coefficients parameterized. The geophysical models are used to synthesize GRACE observations which are subsequently separated leading to time coefficients coinciding with those of the predefined models. In a second step the separation is performed with real, unfiltered time series of five years of monthly GRACE gravity field models (with atmospheric and oceanic background models reconstructed) and a limited number of EOFs. The reconstructed time coefficients are in good agreement with the original ones and exhibit high correlations (0.70 for ocean, 0.91 for atmosphere

Email address: schmidt@dgfi.badw.de (Michael Schmidt)

and 0.93 for continental hydrosphere). Analysis of GRACE residuals and the correlation among the time coefficients substantiate a successful identification.
Keywords: GRACE, Principal Component Analysis, Empirical Orthogonal Functions, Geophysical Models, Multivariate Gauss-Markov Model, Signal Separation

1. Introduction

Since March 2002 the twin Gravity Recovery and Climate Experiment (GRACE) satellites observe the time varying gravity signal (Tapley et al., 2004) caused by the sum of all mass variations within the Earth system with unprecedented precision and resolution. These integral GRACE observations can be used for monitoring the variation of the global water masses. Within the global water cycle the mass variation of the continental hydrosphere is known to be the most uncertain component (see e.g. the 4th assessment report of IPCC (2007), Güntner et al. (2007)). Since standard GRACE products are reduced for atmospheric and oceanic effects by so called background models, the largest part of the remaining signal can be attributed to mass redistributions within the continental hydrosphere on sub-seasonal to seasonal time scales. However, the background models themselves are subject to errors with the risk to be interpreted as additional hydrological signal. Consequently, it is therefore much more challenging to identify and quantify the mass variations of all individual subsystems, continental hydrosphere, oceans, and atmosphere in one single step (Cazenave et al. (1999), Andersen and Hinderer (2005)). Besides atmosphere, ocean and continental hydrosphere, effects from other Earth system components and sub-components have to be considered for a complete and realistic Earth system modelling, such as GIA or influences

of earthquakes. However, in this paper we neglect these effects and focus on the three main Earth system components.

Of course, such a seperation is only possible by introducing prior information. In order to enable a separation of mass signals and to describe the mass variations of the individual components H, O, and A an appropriate functional model has to be selected. Additional information can be introduced, for example, if this model is as far as possible adapted to the input data. Moreover, the target functions should be represented by a series expansion with (1) a minimal number of terms and (2) a neglegible truncation error.

In the following we use empirical orthogonal functions (EOF) derived by principal component analysis (PCA) of geophysical models (Preisendorfer (1988)). Expanding the observations as a time series of these EOFs results in time-dependent expansion coefficients which are denoted as principal components (PC). The EOFs describe mutually orthogonal spatial patterns and the PC form mutually uncorrelated time series. These spatial patterns and time series occur in matched pairs, which are generally referred to as modes. Further it can be shown that the truncated EOF expansion including the first I of altogether N modes with $N > I$ means an optimal $I-$dimensional approximation of the signal under consideration in the sense that their mean squared difference is minimal (Jolliffe (2002), Monahan et al. (2009)). Truncated EOF expansions therefore provide a powerful tool for data compression or dimensionality reduction when it is desired to capture as much variance as possible in a lower-dimensional approximation.

The general strategy for separating an integral mass signal into detail signals is based on a two stage procedure: First, for each of the three components H, O, and A a set of EOFs are derived by PCA of geophysical models. Then a least

squares adjustment is applied to estimate the expansion coefficients for all three sets of base functions using the total mass signal of GRACE as observations in a single step. The main intention of this paper is to introduce the mathematical procedure and to illustrate the computational procedure for identifying mass signals in monthly GRACE gravity fields. The geophysical interpretation of the results is intentionally kept short in this contribution and will be highlighted in a follow-up article. The outline of this contribution is as follows: In Section 2 a presentation of the applied geophysical models is given including GRACE level 2 monthly gravity fields. Section 3 treats the mathematical modeling for the separation of individual mass signals from integral observations. Beginning with a simulation procedure in Section 4.1 we show within a closed loop application that a separation of an integral signal into individual components can be successful. With this knowledge real GRACE gravity observations replace in Section 4.2 the simulated data. It is shown that the estimations give reasonable results for the geophysical models. Conclusions will be given in Section 5.

2. Geophysical Data

In our approach we use global geophysical models for continental hydrosphere, oceans and atmosphere describing mass variations in their respective subsystem. For separating mass signals within the time variable gravity field of the Earth, GRACE monthly gravity fields are used. This input data shall be modeled as the sum of the main contributions of the system Earth, namely the continental hydrology, the oceans and the atmosphere. Furthermore these contributions are modeled as series expansions in terms of appropriate base functions derived from geophysical models briefly introduced in the following subsection.

2.1. Models for Continental Hydrology, Ocean and Atmosphere

Among a large variety of state-of-the-art regional and global geophysical models for each subsystem we processed and analyzed the global models WGHM (WaterGAP Global Hydrology Model) for continental hydrology (Döll et al., 2003) and OMCT (Ocean Model for Climate and Tides) for the ocean (Thomas, 2002) as well as the operational analysis data from ECMWF (European Centre for Medium Weather Forecasts) for atmosphere (Untch et al., 2006). From these models gridded data sets of surface mass variations were derived and introduced into our procedure described below.

ECMWF delivers atmospheric surface pressure fields on a reduced Gaussian grid of approximately 0.5° spacing and for a temporal resolution of 6 h. OMCT is forced by ECMWF which ensures a consistent representation of dynamics and mass transports in the two subsystems atmosphere and ocean. Ocean bottom pressure fields from OMCT are made available for a regular 1.875° grid in monthly mean values. ECMWF and OMCT are routinely used for the dealiasing of the GRACE observations (Flechtner, 2003). For the continental hydrology we derived monthly mean values of total water storage from WGHM on a grid of $1^\circ \times 1^\circ$. The latter model is forced by monthly grid point values of terrestrial surface climate and accounts for surface runoff, groundwater recharge and river discharge. For our approach – explained in the third section – we interpolate in a preparatory step the input data from the geophysical models for each subsystem onto a $1.8^\circ \times 1.8^\circ$ grid with a temporal spacing of one month. Thus, we generate time series of 60 monthly mean values of equivalent water heights beginning with January 2003.

2.2. GRACE Gravity Fields

For identifying mass signals within the time variable part of the Earth's gravity field a time series from 2003 to 2007 of GRACE monthly coefficients from Helmholtz Centre Potsdam, GFZ (GRACE-RL04) of spherical harmonics (Flechtner et al., 2010) up to degree and order 100 are transformed via a spherical harmonic synthesis into values of gravity potential for a regular $1.8° \times 1.8°$ grid at an initial GRACE altitude of 450 km. For reconstructing the total signal detectable from GRACE we added to the provided GRACE model (GSM) the background models for ocean and atmosphere (GAC) afore removed in the dealiasing procedure (Flechtner, 2003). A truncation of the GRACE fields is introduced as we do not rely on coefficients above degree and order 100 due to an increase of the signal-to-noise ratio. The time series of gravity fields contains 57 months within the same time span as for the geophysical models. The months January and June 2003 and January 2004 are excluded where no gravity field models are provided in GFZ-RL04 due to GRACE data gaps (Schmidt R. et al., 2008). More details on the simulated data sets from GRACE are given in Subsection 3.2.

3. Mathematical Modeling

3.1. Procedure

The flowchart shown in Fig. 1 illustrates the basic ideas of our approach. A large number of geophysical models can be used to calculate the contributions of the individual subsystems of the Earth system, e.g. hydrology, oceans and atmosphere, to the total gravitational field. Introducing a geophysical model M_y of a specific subsystem $Y \in \{H$ = continental hydrosphere, O = oceans, A = atmosphere, $\ldots\}$ provides an output signal $g_y(\boldsymbol{x})$ ($y \in \{h, o, a, \ldots\}$) in any point

146 $P(\boldsymbol{x})$. We identify the signal $g_y(\boldsymbol{x})$ with the so-called equivalent water heights
147 (EWH) which can be derived from the gravitational potential (Wahr et al., 1998).
148 We expand g_y in a series

Figure 1: Flowchart of the first part of the applied procedure; GRACE observations are used to estimate the unknown series coefficients

$$g_y(\boldsymbol{x}, t) = \sum_{i=1}^{I_y} c_{y,i}(t)\, \phi_{y,i}(\boldsymbol{x}) \tag{1}$$

149 in terms of base functions $\phi_{y,i}(\boldsymbol{x})$. Herein

$$\boldsymbol{x} = r\ [\cos\varphi\cos\lambda,\ \cos\varphi\sin\lambda,\ \sin\varphi]' \tag{2}$$

150 means the geocentric position vector of any arbitrary observation point $P = P(\boldsymbol{x})$
151 expressed by means of the spherical latitude φ, the spherical longitude λ, the radial
152 distance r and additionally the time t. Examples for modeling a spatio-temporal
153 signal are given by Schmidt M. et al. (2008a), Schmidt M. et al. (2008b) and Seitz
154 et al. (2008).
155 Inserting Eq. (1) into Newton's integral

$$V_y(\boldsymbol{x}_p, t) = G\,\rho_w \int_{\Omega_y} \frac{g_y(\boldsymbol{x}_q, t)}{l_{pq}}\, d\Omega \tag{3}$$

156 (cf. Heiskanen and Moritz (1967)) yields

$$\begin{aligned} V_y(\boldsymbol{x}_p, t) &= G\,\rho_w \sum_{i=1}^{I_y} c_{y,i}(t) \int_{\Omega_y} \frac{\phi_{y,i}(\boldsymbol{x}_q)}{l_{pq}}\, d\Omega \\ &= \sum_{i=1}^{I_y} c_{y,i}(t)\, a_{y,i}(\boldsymbol{x}_p)\ , \end{aligned} \tag{4}$$

with

$$a_{y,i}(\boldsymbol{x}_p) = G\,\rho_w \int_{\Omega_y} \frac{\phi_{y,i}(\boldsymbol{x}_q)}{l_{pq}}\, d\Omega\ . \tag{5}$$

Herein $\boldsymbol{x}_p$ is the position vector of any point $P(\boldsymbol{x}_p)$ on the surface or in the outer space of the Earth and $d\Omega$ is the variable of integration as a spherical element with an expansion of $1.8°\times\ 1.8°$. Furthermore, in the Eqs. (3) to (5) G and ρ_w denote the gravitational constant and the density of water, respectively. The quantity

$$l_{pq} = |\boldsymbol{x}_p - \boldsymbol{x}_q| = \sqrt{r_p^2 + R^2 - 2Rr_p\cos\psi_{pq}} \tag{6}$$

means the distance between the attracted point $P(\boldsymbol{x}_p) = P(\theta_p, \lambda_p, r_p)$ and the source point $P(\boldsymbol{x}_q) = P_q(\theta_q, \lambda_q, r_q = R)$ at the surface of the Earth. The spherical distance ψ_{pq} between the two points is computable from the relation

$$\cos\psi_{pq} = \cos\theta_p\cos\theta_q + \sin\theta_p\sin\theta_q\cos(\lambda_q - \lambda_p)\ . \tag{7}$$

Since V_y is the contribution of the subsystem Y related to the spatial domain Ω_y, the total gravitational potential V can be written as

$$\begin{aligned} V(\boldsymbol{x},t) &= V_h(\boldsymbol{x},t) + V_o(\boldsymbol{x},t) + V_a(\boldsymbol{x},t) + \ldots \\ &= \sum_{i_h=1}^{I_h} c_{h,i_h}(t)\, a_{h,i_h}(\boldsymbol{x}) + \\ &+ \sum_{i_o=1}^{I_o} c_{o,i_o}(t)\, a_{o,i_o}(\boldsymbol{x}) + \\ &+ \sum_{i_a=1}^{I_a} c_{a,i_a}(t)\, a_{a,i_a}(\boldsymbol{x}) + \ldots \end{aligned} \tag{8}$$

or

$$V(\boldsymbol{x},t) = \sum_{y=h,o,a,\ldots} V_y(\boldsymbol{x},t) = \sum_{y=h,o,a,\ldots}\ \sum_{i_y=1}^{I_y} c_{y,i_y}(t)\, a_{y,i_y}(\boldsymbol{x})\ . \tag{9}$$

Figure 2: Flowchart of the second part of the applied procedure; the estimated parts $\widehat{V}_y(\boldsymbol{x},t)$ of the gravitational potential related to the different subsystems Y are the results of the estimation procedure and provide improvements w.r.t. the initial geophysical models M_y introduced in the top row of Fig. 1.

In this equation the coefficients $c_{y,i_y}(t)$ are the time-dependent unknown parameters, which have to be determined from measurements of the gravitational potential $V(\boldsymbol{x},t)$ itself or from functionals $\mathcal{F}(V)(\boldsymbol{x},t)$ of $V(\boldsymbol{x},t)$ such as gravity anomalies. The estimations $\widehat{V}_y(\boldsymbol{x},t)$ of the gravitational potential parts $V_y(\boldsymbol{x},t)$ of the subsystems Y are symbolized in Fig. 2.

The principal component analysis (PCA) expands a spatio-temporal input signal into a series in terms of orthonormal spatial base functions denoted as empirical orthogonal functions (EOF). The corresponding series coefficients, i.e. the principal components (PC), represent the temporal evolution of the input signal. The EOFs are derived from an eigenvector and eigenvalue decomposition of the empirical covariance matrix of the input signal; the PCs are the result of the transformation of the signal into the space spanned by the EOFs. In other words PCA identifies the geographical patterns (EOFs) together with their temporal evolution (PCs). Since the EOFs are related to the eigenvalues, the input signal can be separated into dominant and non-dominant parts, also called modes (Preisendorfer, 1988). As mentioned before we identify the input signal $g_y(\boldsymbol{x},t)$ with the EWHs related to the subsystem Y. With $\phi_{y,i_y}(\boldsymbol{x}) =: EOF_{y,i_y}(\boldsymbol{x}) = EOF_{y,i_y}(\theta,\lambda)$ we rewrite Eq. (1) as

$$g_y(\theta,\lambda,t) = \sum_{i_y=1}^{I_y} c_{y,i_y}(t)\, \boldsymbol{EOF}_{y,i_y}(\theta,\lambda)\ . \tag{10}$$

Herein $i_y = 1,\ldots,I_y$ is the counting index for the modes related to subsystem

Y; the first mode ($i_y = 1$) comprises the most dominant structures of the EWHs from the subsystem Y. In our approach for each of the three subsystems, i.e. continental hydrosphere, oceans and atmosphere an independent PCA was applied to the *EWH* given within the chosen observation interval. Incidentally since the two compartments, oceans and continents are non-overlapping the complete set of base functions for these two Earth system components establish an orthogonal basis on the sphere. But this statement is not valid for introducing the third system component atmosphere.

3.2. Input Data

In the following we apply our method to two kinds of input data, namely,

1. simulated gravitational potential observations $V(\boldsymbol{x}_p, t_l)$ for GRACE calculated from geophysical models and
2. monthly solutions of the gravitational potential values from GRACE processing centers.

For that purpose we define grid points $P(\theta_p, \lambda_p, r_p) = P(\theta_j, \lambda_k, r_p)$ with latitude coordinates $\theta_j = \theta_0 + (j-1)\,\Delta\theta$, $j = 1, \ldots, J$ and longitude coordinates $\lambda_k = \lambda_0 + (k-1)\,\Delta\lambda$, $k = 1, \ldots, K$; $\Delta\theta$ and $\Delta\lambda$ are the discretization intervals. For the radial distance r_p we set $r_P = R + 450\ km$, i.e. the gravitational potential is calculated at a mean GRACE orbital height. For the time-dependency we chose analogously to the spatial part the temporal discretization $t_l = t_0 + (l-1)\,\Delta T$ with $l = 1, \ldots L$, wherein ΔT is set to one month.

3.3. Parameter Estimation

The observation equation for the simulated gravitational potential $V(\theta_j, \lambda_k, r_p, t_l)$ is obtained by considering the residuals $e(\theta_j, \lambda_k, r_p, t_l)$ in Eq. (9) as

$$\begin{aligned} V(\theta_j, \lambda_k, r_p, t_l) + e(\theta_j, \lambda_k, r_p, t_l) \\ = \sum_{y=h,o,a,\ldots} \sum_{i_y=1}^{I_y} c_{y,i_y}(t_l)\, a_{y,i_y}(\theta_j, \lambda_k, r_p)\ . \end{aligned} \tag{11}$$

In this equation the coefficient sets $c_{y,i_y}(t_l)$ for $y \in \{h, o, a\}$ and $i_y = 1, \ldots, I_y$ are the unknown parameters of the hydrology, ocean and the atmosphere model parts. Since the simulated input signal of $N = J \cdot K$ observations given globally at L discrete times, we solve for altogether $u = I_h + I_o + I_a$ unknowns at each time moment t_l with $l = 1, \ldots, L$. On the right-hand side of the observation equation (11) a modification of the EOFs into the *transformed* base functions $a_{y,i_y}(\theta_j, \lambda_k, r_p)$ is performed according to Eq. (5). Whereas, as mentioned before, the EOFs establish an orthogonal basis, the transformed base functions $a_{y,i_y}(\theta_j, \lambda_k, r_p)$, however, do not fulfill this favorable property because of the integration. Next, we collect all *N* gravitational potential observations $V(\theta_j, \lambda_k, r_p, t_l) = V(\boldsymbol{x}_p, t_l)$ with $p = 1, \ldots, N$ for specific time moments t_l with $l \in \{1, \ldots, L\}$ in an $N \times 1$ observation vector $\boldsymbol{y}_l$ and define the $u \times 1$ vector

$$\begin{aligned} \boldsymbol{c}_l = [\, & c_{h,1}(t_l),\, \ldots,\, c_{h,I_h}(t_l),\, c_{o,1}(t_l),\, \ldots, \\ & \ldots,\, c_{o,I_o}(t_l),\, c_{a,1}(t_l),\, \ldots,\, c_{a,I_a}(t_l)\,]' \\ = \big[\, & \boldsymbol{c}'_{h,l},\, \boldsymbol{c}'_{o,l},\, \boldsymbol{c}'_{a,l}\,\big]'\ , \end{aligned} \tag{12}$$

wherein $\mathbf{c}_{h,l}$, $\mathbf{c}_{o,l}$ and $\mathbf{c}_{a,l}$ are $I_h \times 1$, $I_o \times 1$ and $I_a \times 1$ vectors of the unknown series coefficients $c_{h,i_h}(t_l)$, $c_{o,i_o}(t_l)$ and $c_{a,i_a}(t_l)$, respectively. Furthermore, we

introduce with $a_{y,i_y}(\theta_j, \lambda_k, r_p) = a_{y,i_y}(\boldsymbol{x}_p)$ the $N \times u$ matrix

$$\boldsymbol{A} = \begin{bmatrix} a_{h,1}(\boldsymbol{x}_1) & \dots & a_{o,1}(\boldsymbol{x}_1) & \dots & a_{a,1}(\boldsymbol{x}_1) & \dots \\ a_{h,1}(\boldsymbol{x}_2) & \dots & a_{o,1}(\boldsymbol{x}_2) & \dots & a_{a,1}(\boldsymbol{x}_2) & \dots \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots \\ a_{h,1}(\boldsymbol{x}_N) & \dots & a_{o,1}(\boldsymbol{x}_N) & \dots & a_{a,1}(\boldsymbol{x}_N) & \dots \end{bmatrix} \tag{13}$$

$$= [\, \boldsymbol{A}_h \quad \boldsymbol{A}_o \quad \boldsymbol{A}_a \,] .$$

The function $a_{o,1}(\boldsymbol{x})$ of the first mode for the subsystem ocean as one element of the $\boldsymbol{A}$-Matrix is depicted in Fig. 6. According to Eq. (13) the $N \times u$ matrix $\boldsymbol{A}$ consists of the $N \times I_h$ submatrix $\boldsymbol{A}_h$, the $N \times I_o$ submatrix $\boldsymbol{A}_o$ and the $N \times I_a$ matrix$\boldsymbol{A}_a$. Next we define the $N \times L$ observation matrix

$$\boldsymbol{Y} = [\, \boldsymbol{y}_1 \, \dots \, \mathbf{y}_L \,] \ , \tag{14}$$

wherein

$$\boldsymbol{y}_l = [\, y(\boldsymbol{x}_1, t_l),\, y(\boldsymbol{x}_2, t_l), \dots ,\, y(\boldsymbol{x}_N, t_l) \,]' \tag{15}$$

means the $N \times 1$ vector of the given N gravitational potential observations $V(\boldsymbol{x}_p, t_l)$ at time t_l as well as the $u \times L$ matrix

$$\boldsymbol{C} = [\, \boldsymbol{c}_1 \, \dots \, \boldsymbol{c}_L \,] \tag{16}$$

of the unknown parameters. Introducing the $N \times N$ given weight matrix $\boldsymbol{P}$ and the $L \times L$ unknown covariance matrix $\boldsymbol{\Sigma}$ we formulate the multivariate Gauss-Markov model (Koch, 1999):

$$\boldsymbol{A}\,\boldsymbol{C} = E(\boldsymbol{Y}) = \boldsymbol{Y} + \boldsymbol{E} \quad \text{with} \quad D(\text{vec}\boldsymbol{Y}) = \boldsymbol{\Sigma} \otimes \boldsymbol{P}^{-1} \tag{17}$$

Herein $E(\cdot)$ and $D(\cdot)$ mean the expectation and the covariance matrix of a matrix or vector, respectively, 'vec' is an operator which orders the columns of a matrix one below the other as a vector. The symbol '$\otimes$' denotes the Kronecker product between the covariance matrix and the weight matrix. The $N \times L$ matrix $\boldsymbol{E}$ contains the residuals $e(\boldsymbol{x}_p, t_l)$. Since the matrix $\boldsymbol{A}$ is of full column rank, the least squares adjustment yields the estimation

$$\widehat{\boldsymbol{C}} = (\boldsymbol{A}'\,\boldsymbol{P}\,\boldsymbol{A})^{-1}\,\mathbf{A}'\,\boldsymbol{P}\,\boldsymbol{Y} \tag{18}$$

of the unknown parameter matrix $\boldsymbol{C}$. The corresponding covariance matrix reads

$$D(\text{vec}\widehat{\boldsymbol{C}}) = \boldsymbol{\Sigma} \otimes (\boldsymbol{A}'\,\boldsymbol{P}\,\boldsymbol{A})^{-1}\ . \tag{19}$$

The matrix $\widehat{\boldsymbol{E}}$ of the residuals between the observations and the recovery of the gravitational potential is given as

$$\widehat{\boldsymbol{E}} = \boldsymbol{A}\,\widehat{\mathbf{C}} - \boldsymbol{Y}\ . \tag{20}$$

Finally, an estimation of the covariance matrix $\boldsymbol{\Sigma}$ is obtained from

$$\widehat{\boldsymbol{\Sigma}} = \frac{\widehat{\boldsymbol{E}}'\,\boldsymbol{P}\,\widehat{\boldsymbol{E}}}{N-u}\ . \tag{21}$$

This matrix allows for calculating the correlations between the observation vectors $\boldsymbol{y}_l$, i.e. it is a measure of the time-dependency.

4. Results

4.1. Simulated Gravitational Potential

In order to check our method we compose simulated gravitational potential observations $V(\boldsymbol{x}_p, t_l)$ on grid points $P(\theta_p, \lambda_p, r_p) = P(\theta_j, \lambda_k, r_p)$ as introduced

Figure 3: Eigenvalues of OMCT, the solid curve shows the total variance, i.e. the sum of the eigenvalues, the vertical bars are the amounts of the eigenvalues. Note, the eigenvalues are always non-negative. A percentage of 98% of the total variance is achieved by the first $I_o = 37$ eigenvalues.

in Section 2.2. The gravitational potential values are computed from the combination of the three geophysical models WGHM, OMCT and ECMWF introduced in Section 2.1. Their single *EWH* contributions are added to one mass equivalent layer for each time step t_l according to Eq. (9). In order to deal with more realistic input data we add white noise with a standard deviation of $\sigma = 0.003$ m^2/s^2; cf. Han et al. (2006). A spherical harmonic expansion until degree and order 100 requires a sampling with discretaization intervals $\Delta\theta = \Delta\lambda = 1.8°$ as introduced at the end of Section 3.2. With $\theta_0 = -88.2°$ and $\lambda_0 = 0°$ we obtain $N = 19,800$ observations at $L = 60$ discrete times $t = t_l$ for $l = 1, \ldots, L = 60$ were chosen.

For approximating each subsystem by a selected threshold of 98% of the total variance of the mass variations we obtain from Fig. 3 for the ocean model OMCT $I_{o,max} = 37 =: I_o$, for the hydrological model WGHM $I_{h,max} = 22 =: I_h$ and for the atmosphere model ECMWF $I_{a,max} = 35 =: I_a$. Figures 4 and 5 show the results of the PCA of the ocean model OMCT for the first and the second mode. In series expansions the base functions are usually assumed to be dimensionless. Thus, the measurement unit of the input data is transferred into the PCs. The top panels depict the base function $EOF_{o,i}(\boldsymbol{x})$ with $i = 1, 2$. Figure 6 shows the transformed function $a_{o,1}(\boldsymbol{x})$ for the subsystem ocean at GRACE orbital height, i.e. the radial distance $r_p = R + 450$km. Note that only mass variations in the oceans can be identified. Due to the integration according to Eq. (5) non-zero values appear all over the globe and are not restricted to the area of the oceans.

Figure 4: First mode of OMCT; the top panel shows $EOF_{o,1}(\boldsymbol{x})$, the bottom panel depicts the corresponding PCs $c_{o,1}(t_l)$ in [m].

Figure 5: Second mode of OMCT; the top panel shows $EOF_{o,2}(\boldsymbol{x})$, the bottom panel depicts the corresponding PCs $c_{o,2}(t_l)$ in [m].

The components $a_{o,1}(\boldsymbol{x})$ for $\boldsymbol{x} = \boldsymbol{x}_p$ with $p = 1, \ldots, N$ are introduced into the $N \times I_o$ submatrix $\boldsymbol{A}_o$ defined in Eq. (13).

Since the simulated input signal $V(\boldsymbol{x}_p, t_l)$ is given globally at $L = 60$ discrete times t_l, we solve with $u = I_h + I_o + I_a = 94$ for altogether $u \cdot L = 94 \cdot 60 = 5640$ unknowns for all 60 time epochs. For the parameter estimation according to the Eqs. (17) to (21) we choose with $\boldsymbol{P} = \boldsymbol{I}$ the $N \times N$ unit matrix as weight matrix $\boldsymbol{P}$. The estimations $\widehat{c}_{y,1}(t_l)$ and the PCs $c_{y,1}(t_l)$ of the PCA for $y \in \{h, o, a\}$ are shown in Fig. 7. To be more specific, the green circles in each panel display the original time-dependent PCs from the PCA for each subsystem. The red circles are the estimated coefficients $\widehat{c}_{y,1}(t_l)$ from the adjustment according to Eq. (18). Since the first mode comprises the part of the input signal with the highest variance, it reflects for the subsystems hydrology and atmosphere the most dominating annual signal. High correlations between the original coefficients and the estimations are obtained for the subsystems hydrology ($99.9\ \%$) and atmosphere ($99.8\ \%$); the largest differences are detected for the oceanic part but the correlation still remains high with $94.9\ \%$. As a showcase Fig. 8 depicts the correlations between the original coefficients – the PC – and the adjusted time series $\widehat{c}_{o,1}$ for the subsystem ocean (OMCT) and all $I_o = 37$ applied modes. A clear decrease

Figure 6: Function $a_{o,1}(\boldsymbol{x})$ of the first mode for the subsystem ocean.

Figure 7: Estimations $\widehat{c}_{h,1}(t_l)$, $\widehat{c}_{o,1}(t_l)$ and $\widehat{c}_{a,1}(t_l)$ (red circles) of the first mode from simulation procedure in [m]. The correlations w.r.t. the PCs $c_{h,1}(t_l)$, $c_{o,1}(t_l)$ and $c_{a,1}(t_l)$ of the PCAs (green circles) are higher than 0.94.

Figure 8: Correlations (in $\%$) between the estimated coefficients $\widehat{c}_{o,i_o}(t_l)$ and the original PCs $c_{o,i_o}(t_l)$ for all $I_o = 37$ modes for subsystem ocean within the simulation process.

of the correlations for higher modes appears as well as the amounts of the corresponding eigenvalues are diminishing significantly (c.f. Fig 3). Comparable structures in the correlation plots (not shown here) for hydrology and atmosphere are detectable. However, the correlations are at a higher level for these both subsystems.

Further the $u \times u$ correlation matrix $\boldsymbol{R}$ of the unknown parameter vectors $\widehat{\boldsymbol{c}}_l$ can be derived from the inverse of the normal equation matrix $\boldsymbol{A}' \boldsymbol{P} \boldsymbol{A} = \boldsymbol{A}' \boldsymbol{A}$ (cf. Eq. (18)). As can be seen from Fig. 9 there exist at least partly correlative structures between the subsystems. Higher (anti-)correlations between the subsystems ocean and atmosphere might be caused by the fact that the ocean model OMCT is driven by atmospheric data from ECMWF (c.f. Section 2.1). Smaller correlations within one subsystem are accounted for the transformation of the EOFs into the functions $a_{y,i_y}(\boldsymbol{x})$ according to Eq. (5).

4.2. Real GRACE Solutions

Since the results of the closed-loop simulation presented in the previous subsection are very promising we now apply the developed procedure to real GRACE gravity solutions from GFZ. We choose the monthly solutions of EIGEN-GRACE RL04 for a total time span between February 2003 and December 2007 including

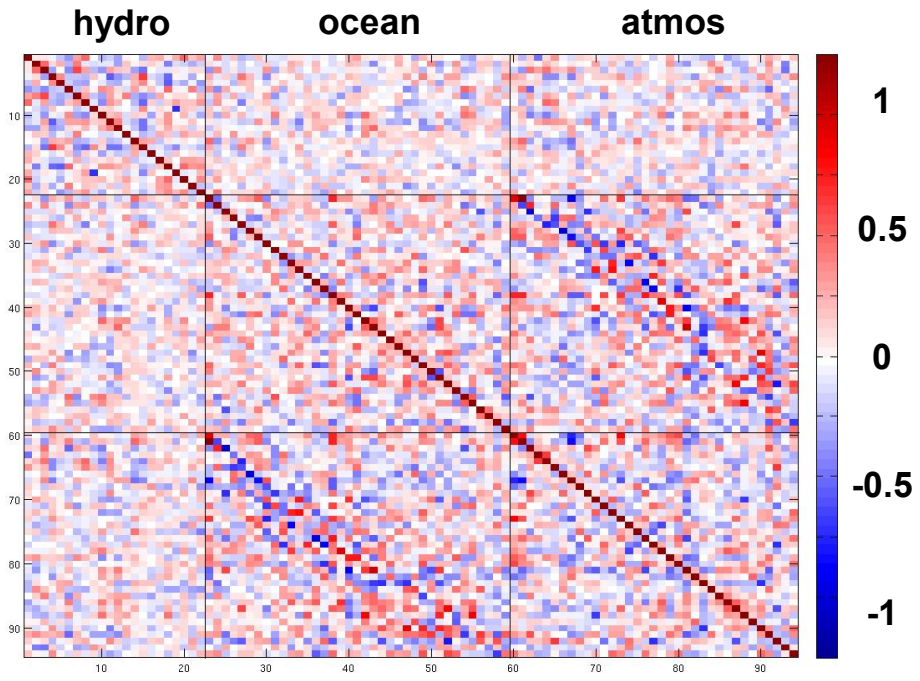

Figure 9: Correlation matrix $\boldsymbol{R}$ of the unknowns; the top block at the left is related to the hydrology part, the center part to the oceans an the bottom right block to atmosphere.

a few data gaps in 2003 and 2004. The model is evaluated up to degree and order 100 whereby no filtering was applied in a preprocessing step. For retrieving the entire mass signal detected from GRACE we use the sum of the RL04 GSM and the GAC models (c.f. Section 2.2). The spherical harmonic synthesis is performed at the orbital height $r = R + 450$ km in terms of geopotentials [m^2/s^2] as mentioned in Subsection 3.2.

Within a mass signal separation of GRACE data we again require an approximation rate of 98% of the total geophysical models consisting of the hydrological, the oceanic and the atmospheric part. For that purpose we apply again the chosen values $I_h = 22$, $I_o = 37$ and $I_a = 35$ from Section 4.1 for the upper limits in the observation equation (11). Thus, we solve for altogether $u = 94$ unknowns for each time step t_l with $l = 1, \ldots, L = 57$.

Results from the parameter estimation are depicted in Fig. 10 for the first mode. The green circles in each subplot display the original time-dependent coefficients, i.e. the PCs from the PCAs of each subsystem. The red circles are the estimated coefficients $\widehat{c}_{y,i_y}(t_l)$ from the parameter estimation (18). High correlations to the original coefficients exist for the subsystems hydrology (93.4 %) and atmosphere (91.0 %). However, the poor performance for the oceanic subsystem – the correlations are below 70 % – is still unsettled. An indication might be linked to the afore detected high (anti-)correlation to the subsystem atmosphere within the correlation matrix of the unknowns (cf. Fig. 9). The dominating annual signals for the subsystems hydrology and atmosphere are clearly detectable. However, the deviations between the original and the adjusted coefficients are rather high for the hydrological signal $\widehat{c}_{h,1}$. The well known phase shift for the annual signal of the GRACE solution w.r.t. WGHM is clearly visible and amounts up to one

Figure 10: Time-dependent coefficients: original PCs (green circles) and the estimations $\widehat{c}_{y,1}(t_l)$ (red circles) for each subsystem of the first mode in [m].

Figure 11: Time-dependent coefficients for the hydrological subsystem: original PCs (green circles) and estimations $\widehat{c}_{h,6}(t_l)$ (red circles) of the sixth mode in [m].

month (Seitz et al., 2008). Furthermore the magnitude of the estimated coefficients $\widehat{c}_{h,1}(t_l)$ is around the half of the magnitude of the original coefficients, i.e. the PCs $c_{h,1}(t_l)$.

In general, our procedure based on PCAs is capable for detecting aperiodic and transient signals. The sixth mode of the hydrological subsystem (Fig. 11), for example, shows a clear interannual variation over the entire period in the GRACE signal. A very similar progression can be seen in the same mode of the hydrological model WGHM which involves a correlation coefficient of 0.73 between both curves. Fig. 12 depicts for the three subsytems the correlations between the estimated and the original coefficients for all applied modes. A decrease of this consistency for higher mode numbers is evident whereupon the contribution to the total variance of the mass signal becomes increasingly insignificant.

Figure 13 shows the covariance matrix of the observations $\widehat{\mathbf{\Sigma}}$ according to Eq. (21). Characteristic samples in terms of a checkerboard pattern are detectable. This might be ascribed to a residual annual signal and suggest temporal deficiencies of the applied geophysical models. In addition a possible application is the reconstruction of the original signal of each subsytem *H*, *O*, and *A* from the esti-

Figure 12: Correlations (in $\%$) between the estimated coefficients $\widehat{c}_{h,i_h}(t_l)$, $\widehat{c}_{o,i_o}(t_l)$ and $\widehat{c}_{a,i_a}(t_l)$ and the original PCs $c_{h,i_h}(t_l)$, $c_{o,i_o}(t_l)$ and $c_{a,i_a}(t_l)$ for all applied modes of the three subsystems within a mass signal separation of GRACE data.

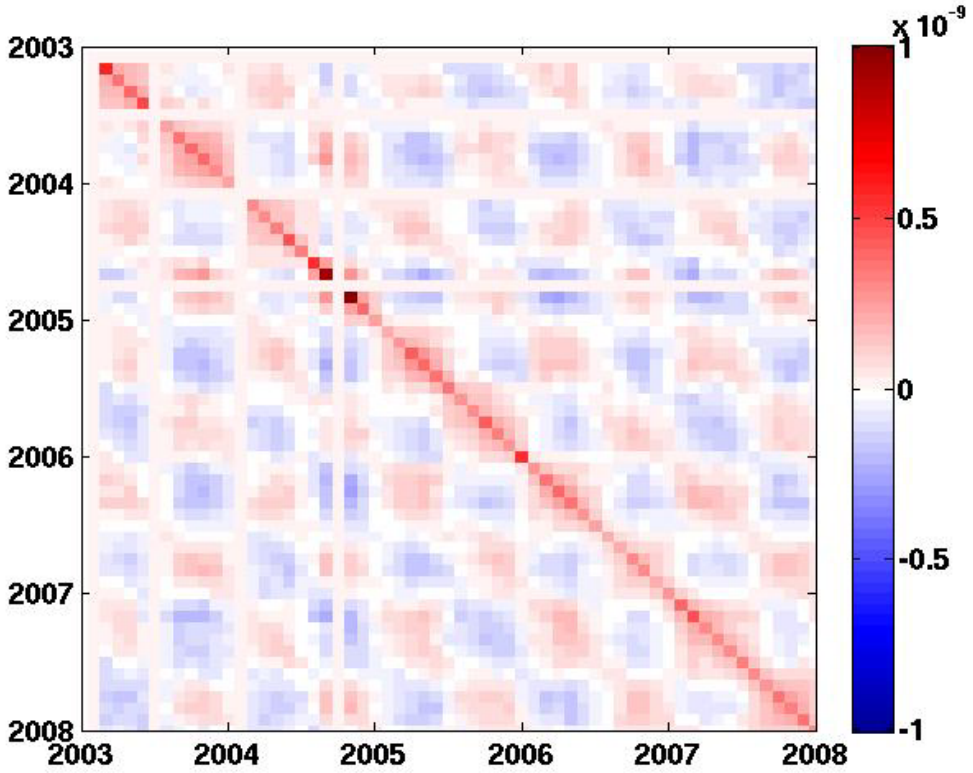

Figure 13: Covariance matrix $\widehat{\mathbf{\Sigma}}$ of the observations

mated temporal coefficients $\widehat{c}_{h,i_h}(t_l)$, $\widehat{c}_{o,i_o}(t_l)$ and $\widehat{c}_{a,i_a}(t_l)$ and the corresponding EOFs from the PCAs. However, a recalculation of the geophysical models based on all modes indicates high deviations relative to the original time series of the geophysical models. An interpretation is yet out of the scope of this contribution and will be presented in a follow-up paper.

5. Conclusion

In this paper we presented an approach for the separation and identification of mass signals from different components of the Earth system with a focus on the methodology. In a first step we applied PCAs to given geophysical models in order to establish the systems of base functions. It was shown by a simulation that a separation of an integral signal into its individual components of the Earth system could be successful.

Based on the findings of the simulation procedure we estimated the series coefficients of the same system of base functions from monthly GRACE solutions of the gravitational field. Correlations between the original and the estimated coefficients are mostly high, but discrepancies between the geophysical models and the mass signal observed from GRACE are obvious, especially in the case of continental hydrology. The residuals of the parameter estimation indicate deficiencies of the applied geophysical models with respect to their temporal evolution, but on the other hand they are able to describe physical processes well with respect to their spatial structures. This result suggests that GRACE can be especially useful for an improvement of the temporal behavior (i.e. of phases and magnitudes) of the spatial structures.

6. Acknowledgment

The authors thank the NASA/DLR GRACE project for the GRACE data products provided by GFZ. We are grateful to A. Güntner and M. Thomas from GFZ processing WGHM and OMCT data respectively. ECMWF data was distributed by GFZ's ISDC data base made accessible through the priority programme "mass transport and mass distribution in the system Earth" SPP 1257 of the German Re-

search Foundation (DFG). We thank two anonymous reviewers and the associated editor for their helpful comments to enhance the original manuscript.

References

References

Andersen O.B. and J. Hinderer, 2005: Global inter-annual gravity changes from GRACE: Early results. Geophys. Res. Lett., 32, L01402, doi:10.1029/2004GL020948.

Cazenave A., F. Mercier, F. Bouille, J.M. Lemoine, and J.F. Cretaux, 1999: Global-scale interactions between the solid Earth and its fluid envelopes at the seasonal time scale. Earth planet. Sci. Lett, 171, 549559.

Döll P., F. Kaspar, and B. Lehner, 2003: A global hydrological model for deriving water availability indicators: model tuning and validation. J. Hydrology, 270, 105-134.

Flechtner F., 2003: AOD1B product description document. GRACE Project Documentation, JPL 327-750, Rev. 1.0, JPL Pasadena, CA.

Flechtner F., Ch. Dahle, K.H. Neumayer, R. König, and Ch. Förste, 2010: The Release 04 CHAMP and GRACE EIGEN Gravity Models. In: Flechtner et al. (eds.): System Earth via Geodetic-Geophysical Space Techniques. Adv. Technologies in Earth Sciences, Springer Berlin, 41-58, doi 10.1007/978-3-642-10228-8-4.

Güntner A., J. Stuck, S. Werth, P. Döll, K. Verzano, and B. Merz, 2007: A global

403 analysis of temporal and spatial variations in continental water storage. Water
404 Resour. Res., 43, W05416, doi:10.1029/2006WR005247.

405 Han S., C.K. Shum, and C. Jekeli, 2006: Precise estimation of in situ geopotential
406 difference from GRACE lowlow satellite-to-satellite tracking and accelerome-
407 try data. J. Geophys. Res. 111:B04411. doi:10.1029/2005JB003719.

408 Heiskanen W. A. and H. Moritz, 1967: Physical Geodesy. Freeman and Co., San
409 Francisco.

410 IPCC, 2007: Climate Change 2007: The Physical Science Basis. Contribution of
411 Working Group I to the Fourth Assessment Report of the Intergovernmental
412 Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M.
413 Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.)]. Cambridge University
414 Press, Cambridge, United Kingdom and New York, NY, USA.

415 Jolliffe I.T., 2002: Principal Components Analysis. Springer Series in Statistics,
416 Springer, 2nd Edition, New York.

417 Koch K.R., 1999: Parameter estimation and hypothesis testing in linear models.
418 Springer, Berlin, Heidelberg, New York.

419 Monahan A.H., J.C. Fyfe, M.H.P. Ambaum, D.B. Stephenson, and G.R. North,
420 2009: Empirical Orthogonal Functions: The Medium is the Message. J. Cli-
421 mate, 22, 6501-6514, doi: 10.1175/2009JCLI3062.1

422 Peltier W.R., 2004: Global Glacial Isostasy and the Surface of the Ice-Age Earth:
423 ICE-5G (VM2) Model and GRACE, Annu. Rev. Earth Planet. Sci. 32, 111-149.

Preisendorfer R., 1988: Principal Component Analysis in Meteorology and Oceanography. Elsevier, Amsterdam.

Schmidt M., J. Kusche, S.-C. Han, C.K. Shum, and M.O. Karslioglu, 2008a: Multi-resolution representation of the gravity field from satellite data based on wavelet expansions with time-dependent coefficients. Proceedings of the 1st International Symposium of the International Gravity Field Service, 28.8.-1.9.2008, Journal of Mapping (Harita Dergisi), special issue: 18, 477-482

Schmidt, M., F. Seitz, C.K. Shum, 2008b: Regional 4-D hydrological mass variations from GRACE, atmospheric flux convergence and river gauge data. Journal of Geophysical Research, 113, doi:10.1029/2008JB005575

Schmidt R., S. Petrovic, A. Güntner, F. Barthelmes, J. Wünsch, and J. Kusche, 2008: Periodic components of water storage changes from GRACE and global hydrology models, J. Geophys. Res. 113.

Seitz F., M. Schmidt, and C.K. Shum, 2008: Signals of extreme weather conditions in Central Europe in GRACE 4-D hydrological mass variations, Earth and Planetary Science Letters, Vol. 268.

Tapley B.D., S. Bettadpur, J. Ries, P. Thompson, and M. Watkins, 2004: GRACE measurements of mass variability in the Earth system. Science 305, 503505.

Thomas M., 2002: Ozeanisch induzierte Erdrotationsschwankungen - Ergebnisse eines Simultanmodells für Zirkulation und Gezeiten im Weltozean, PhD Thesis, University of Hamburg.

445 Untch A., M. Miller, R. Buizza, and P. Janssen, 2006: Towards a global meso-
446 scale model: the high resolution system TL799L91 and TL399L62 EPS.
447 ECMWF newsletter (108).

448 Wahr J., M. Molenaar, and F. Bryan, 1998: Time-variability of the Earth's grav-
449 ity field: hydrological and oceanic effects and their possible detection using
450 GRACE. J. Geophys. Res., 103, 30205-30230.