Towards the Combination of Data Sets from Various Observation Techniques

M. Schmidt, F. Göttl, and R. Heinkelmann

Abstract

Nowadays, heterogeneous data sets are often combined within a parameter estimation process in order to benefit from their individual strengths and favorable features. Frequently, the different data sets are complementary with respect to their measurement principle, the accuracy, the spatial and temporal distribution and resolution, as well as their spectral characteristics.

This paper gives first a review on various combination strategies based on the Gauss-Markov model; special attention will be turned on the stochastic modeling of the input data, e.g. the influence of correlations between different sets of input data. Furthermore, the method of variance component estimation is presented to determine the relative weighting between the observation techniques.

If the input data sets are sensitive to different parts of the frequency spectrum the multi-scale representation might be applied which basically means the decomposition of a target function into a number of detail signals each related to a specific frequency band. A successive parameter estimation can be applied to determine the detail signals.

Keywords

Combination strategies • Gauss-Markov model • Operator-software impact • Variance component estimation • Multi-scale representation

1 Introduction

To achieve the goals of the Global Geodetic Observing System (GGOS) a combination of complementary sensor and observation systems has to be applied (Rummel et al. 2005). This combination should be performed within a parameter estimation process in order to benefit from their individual strengths and favorable features. Data sets could be complementary, e.g. with respect to their measurement principle, accuracies, spatial and temporal distribution and resolution or their spectral characteristics. Various combination strategies can be applied, e.g. the rigorous combination within the *Gauss-Markov model* or the model of *variance component estimation* (Koch 1999).

In reference frame estimation geometric space-geodetic techniques such as VLBI (Very Long Baseline Interferometry), GNSS (Global Navigation Satellite Systems) and SLR (Satellite Laser Ranging) are usually combined; special attention has to be turned on the datum definition and the consideration of constraints; see e.g. Seitz (2015). High-resolution geoid determination is performed by combining the modern gravimetric space-geodetic techniques, i.e. CHAMP (CHAllenging Minisatellite Payload), GRACE (Gravity Recovery and Climate Experiment) and GOCE (Gravity field and steady-state Ocean Circulation Explorer)

M. Schmidt (✉) • F. Göttl
Deutsches Geodätisches Forschungsinstitut (DGFI),
Alfons-Goppel-Str. 11, 80539 München, Germany
e-mail: schmidt@dgfi.badw.de

R. Heinkelmann
Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany

H. Kutterer et al. (eds.), *The 1st International Workshop on the Quality of Geodetic Observation and Monitoring Systems (QuGOMS'11)*, International Association of Geodesy Symposia 140,
DOI 10.1007/978-3-319-10828-5_6,

with terrestrial, airborne and shipborne measurements as well as data from altimetry missions. This can be achieved within a rigorous combination, mostly by considering *prior information* in form of a given gravity model (see e.g. van Loon and Kusche 2007) or by *least-squares collocation* (see e.g. Kühtreiber and Abd-Elmotaal 2007). Since GRACE provides accurate satellite data for the long wavelength part (larger 500 km) of the gravity field, GOCE however for the wavelength part between 100 and 500 km a *spectral combination* (see e.g. Gitlein et al. 2007) or a multi-scale representation (see e.g. Haagmans et al. 2002) can be applied for high-resolution geoid modeling.

Recently data combination procedures entered other fields of geodetic interest such as in hydrosphere or atmosphere modeling; see e.g. Dettmering et al. (2011). A crucial point is always the proper choice of the complete covariance matrix of the observations. A detailed discussion on this topic is, for instance, given by van Loon (2008).

The outline of this paper is the following. After introducing the general Gauss-Markov model in Sect. 2, we present in Sect. 3 altogether six models for combining data sets from various observations techniques. As an example we demonstrate in Sect. 4 how a full covariance matrix can be constructed for time series processed in different analysis centers (ACs) from the same original measurements. The different ACs are treated as here like different observation techniques. In Sect. 5 finally we discuss the multi-scale representation for providing a concept for a multi-scale (spectral) combination.

2 Gauss-Markov Model

The Gauss-Markov model is defined as

$$\mathbf{y} + \mathbf{e} = \mathbf{X}\boldsymbol{\beta} \quad \text{with} \quad D(\mathbf{y}) = \sigma^2\,\mathbf{P}^{-1} = \sigma^2\,\mathbf{Q}\,, \tag{1}$$

wherein $\mathbf{y}$ is the $n \times 1$ vector of the observations, $\mathbf{e}$ the related $n \times 1$ vector of the observation errors, $\mathbf{X}$ the $n \times u$ given coefficient (design) matrix, $\boldsymbol{\beta}$ the $u \times 1$ vector of the unknown parameters, σ^2 the unknown variance factor, $\mathbf{P}$ the $n \times n$ given positive definite weight matrix of the observations and $\mathbf{Q} = \mathbf{P}^{-1}$ the $n \times n$ given cofactor matrix, furthermore $n > u$ holds; see e.g. Koch (1999). In case of rank$\mathbf{X} = u$ we denote the model (Eq. (1)) as *Gauss-Markov model of full rank*. However, if rank$\mathbf{X} = r < u$ holds, the model (Eq. (1)) is called *Gauss-Markov model not of full rank* and the rank deficiency amounts $d = u - r$.

The *least squares method* yields the normal equation system

$$\mathbf{N}\boldsymbol{\beta} = \mathbf{b}\,, \tag{2}$$

wherein we introduced the normal equation matrix $\mathbf{N} = \mathbf{X}^T\mathbf{P}\,\mathbf{X}$ as well as the "right-hand side" vector $\mathbf{b} = \mathbf{X}^T\mathbf{P}\,\mathbf{y}$. In case of full rank, i.e. rank$\mathbf{N} = u$ we obtain the solution

$$\hat{\boldsymbol{\beta}} = \mathbf{N}^{-1}\,\mathbf{b}\,. \tag{3}$$

In case of a rank deficiency, i.e. rank$\mathbf{N} = r < u$ the solution reads

$$\overline{\boldsymbol{\beta}} = \mathbf{N}^{-}\,\mathbf{b}\,, \tag{4}$$

wherein $\mathbf{N}^{-}$ means a generalized inverse of $\mathbf{N}$, e.g. the pseudoinverse $\mathbf{N}^{+}$.

In the two cases the covariance matrices of the estimations $\hat{\boldsymbol{\beta}}$ and $\overline{\boldsymbol{\beta}}$ are given as

$$D(\hat{\boldsymbol{\beta}}) = \sigma^2\,\mathbf{N}^{-1}\,, \tag{5}$$

$$D(\overline{\boldsymbol{\beta}}) = \sigma^2\,\mathbf{N}^{-}\,. \tag{6}$$

An estimation of the variance factor can be derived from the *maximum-likelihood method*; see e.g. Koch (1999).

3 Combination Models

Whereas the vector $\mathbf{y}$ introduced in the Gauss-Markov model (Eq. (1)) can be assumed to be the observation vector of a single technique, we discuss in the following several combination models (CMs) within the *multi-technique* case. To be more specific, we combine the observation vectors $\mathbf{y}_p$ with $p = 1, \ldots, P$ from altogether P techniques such as GNSS, VLBI or SLR. Note, that we identify a set of GNSS observation sites or GRACE ACs (see Sect. 4) also with different observation techniques or groups. In the CMs we distinguish between different stochastic approaches for the observation vectors and different kinds of unknown deterministic parameters.

CM 1: First we assume that for each technique p the vector $\boldsymbol{\beta}$ of the unknown parameters is the same; i.e. we reformulate the Gauss-Markov model (Eq. (1)) as

$$\mathbf{y}_p + \mathbf{e}_p = \mathbf{X}_p\,\boldsymbol{\beta} \quad \text{with} \quad C(\mathbf{y}_p, \mathbf{y}_q) = \sigma^2\,\mathbf{Q}_{p,q} \tag{7}$$

for $q, p = 1, \ldots, P$, wherein $\mathbf{y}_p$ is the $n_p \times 1$ vector of the observations and $\mathbf{e}_p$ the $n_p \times 1$ vector of the observation errors. Furthermore, $\mathbf{X}_p$ is the $n_p \times u$ given design matrix and $\mathbf{Q}_{p,q} = \mathbf{Q}_{q,p}^T$ the $n_p \times n_q$ given cofactor matrix between the observation vectors $\mathbf{y}_p$ and $\mathbf{y}_q$; the other quantities have been already defined in the context of Eq. (1). Next we rewrite CM 1 (Eq. (7)) and obtain the formulation

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \cdot \\ \mathbf{y}_P \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \cdot \\ \mathbf{e}_P \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \cdot \\ \mathbf{X}_P \end{bmatrix} \boldsymbol{\beta}$$

with

$$D(\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \cdot \\ \mathbf{y}_P \end{bmatrix}) = \sigma^2 \begin{bmatrix} \mathbf{Q}_{1,1} & \mathbf{Q}_{1,2} & \dots & \mathbf{Q}_{1,P} \\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} & \dots & \mathbf{Q}_{2,P} \\ & \dots\dots\dots & & \\ \mathbf{Q}_{P,1} & \mathbf{Q}_{P,2} & \dots & \mathbf{Q}_{P,P} \end{bmatrix} . \tag{8}$$

Note, we require for the total number $n = \sum_{p=1}^{P} n_p$ of observations that the inequality $n > u$ holds. The complete design matrix determines the total rank of the model. CM 1 (Eq. (8)) could be transferred into the Gauss-Markov model (Eq. (1)) by defining the $n \times 1$ vectors $\mathbf{y} = \left[\mathbf{y}_1^T, \mathbf{y}_2^T, \dots, \mathbf{y}_p^T\right]^T$ and $\mathbf{e} = \left[\mathbf{e}_1^T, \mathbf{e}_2^T, \dots, \mathbf{e}_p^T\right]^T$ as well as the corresponding coefficient matrix $\mathbf{X}$ and weight matrix $\mathbf{P}$. Thus, the estimation of the unknown parameter vector $\boldsymbol{\beta}$ is obtained from Eqs. (3) or (4), respectively.

CM 2: In our second approach we assume that the observation vectors of the different techniques are mutually uncorrelated, i.e. we set $\mathbf{Q}_{p,q} = \mathbf{0}$ for $p \neq q$ in CM 1 (Eq. (8)). With $\mathbf{P}_{p,p} = \mathbf{Q}_{p,p}^{-1}$ we obtain with $\mathbf{N}_p = \mathbf{X}_p^T \mathbf{P}_{p,p} \mathbf{X}_p$ and $\mathbf{b}_p = \mathbf{X}_p^T \mathbf{P}_{p,p} \mathbf{y}_p$ the normal equation system

$$\left(\sum_{p=1}^{P} \mathbf{N}_p\right) \boldsymbol{\beta} = \sum_{p=1}^{P} \mathbf{b}_p . \tag{9}$$

Due to the neglect of the cofactor matrices $\mathbf{Q}_{p,q}$ for $p \neq q$, i.e. the correlations between the observation vectors $\mathbf{y}_p$ and $\mathbf{y}_q$, the estimated variances of the estimated unknown parameter vector

$$\hat{\boldsymbol{\beta}} = \left(\sum_{p=1}^{P} \mathbf{N}_p\right)^{-1} \sum_{p=1}^{P} \mathbf{b}_p \tag{10}$$

are usually too optimistic. Note, in Eq. (10) we require $\mathrm{rank}(\sum_{p=1}^{P} \mathbf{N}_p) = u$.

CM 3: In the third approach we separate for each technique the $u \times 1$ vector $\boldsymbol{\beta} = [\boldsymbol{\beta}_p^T, \boldsymbol{\beta}_c^T]^T$ into a $u_p \times 1$ technique dependent subvector $\boldsymbol{\beta}_p$ of so-called *local* parameters and a technique independent $u_c \times 1$ subvector $\boldsymbol{\beta}_c$ of *common* or *global* parameters. Under this assumption the extended version of CM 1 (Eq. (7)) reads

$$\mathbf{y}_p + \mathbf{e}_p = \begin{bmatrix} \mathbf{X}_{p,p} & \mathbf{X}_{p,c} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_p \\ \boldsymbol{\beta}_c \end{bmatrix}$$
$$\text{with} \quad C(\mathbf{y}_p, \mathbf{y}_q) = \sigma^2 \mathbf{Q}_{p,q} , \tag{11}$$

wherein $\mathbf{X}_{p,p}$ and $\mathbf{X}_{p,c}$ are given $n_p \times u_p$ and $n_p \times u_c$ coefficient block matrices. Reformulating this model yields

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \cdot \\ \mathbf{y}_P \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \cdot \\ \mathbf{e}_P \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{1,1} & \dots & \mathbf{0} & \mathbf{X}_{1,c} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{X}_{2,c} \\ & \dots\dots\dots & & \\ \mathbf{0} & \dots & \mathbf{X}_{P,P} & \mathbf{X}_{P,c} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \cdot \\ \boldsymbol{\beta}_P \\ \boldsymbol{\beta}_c \end{bmatrix}$$

with

$$D(\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \cdot \\ \mathbf{y}_P \end{bmatrix}) = \sigma^2 \begin{bmatrix} \mathbf{Q}_{1,1} & \mathbf{Q}_{1,2} & \dots & \mathbf{Q}_{1,P} \\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} & \dots & \mathbf{Q}_{2,P} \\ & \dots\dots\dots & & \\ \mathbf{Q}_{P,1} & \mathbf{Q}_{P,2} & \dots & \mathbf{Q}_{P,P} \end{bmatrix} . \tag{12}$$

CM 4: In this case we introduce for each technique an unknown individual variance component σ_p^2 and an unknown covariance components $\sigma_{p,q}$ for the covariance matrices between two individual techniques. Thus, the generalized version of CM 3 (Eq. (11)) reads

$$\mathbf{y}_p + \mathbf{e}_p = \begin{bmatrix} \mathbf{X}_{p,p} & \mathbf{X}_{p,c} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_p \\ \boldsymbol{\beta}_c \end{bmatrix}$$
$$\text{with} \quad C(\mathbf{y}_p, \mathbf{y}_q) = \sigma_{p,q} \mathbf{Q}_{p,q} \tag{13}$$

with $\sigma_{p,q} = \sigma_{q,p}$ and $\sigma_{p,p} = \sigma_p^2$. Reformulating the model above yields the *linear model with unknown variance and covariance components*

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \cdot \\ \mathbf{y}_P \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \cdot \\ \mathbf{e}_P \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{1,1} & \dots & \mathbf{0} & \mathbf{X}_{1,c} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{X}_{2,c} \\ & \dots\dots\dots & & \\ \mathbf{0} & \dots & \mathbf{X}_{P,P} & \mathbf{X}_{P,c} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \cdot \\ \boldsymbol{\beta}_P \\ \boldsymbol{\beta}_c \end{bmatrix}$$

with

$$D(\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \cdot \\ \mathbf{y}_P \end{bmatrix}) = \begin{bmatrix} \sigma_1^2 \mathbf{Q}_{1,1} & \sigma_{1,2} \mathbf{Q}_{1,2} & \dots & \sigma_{1,P} \mathbf{Q}_{1,P} \\ \sigma_{1,2} \mathbf{Q}_{2,1} & \sigma_2^2 \mathbf{Q}_{2,2} & \dots & \sigma_{2,P} \mathbf{Q}_{2,P} \\ & \dots\dots\dots\dots & & \\ \sigma_{1,P} \mathbf{Q}_{P,1} & \sigma_{2,P} \mathbf{Q}_{P,2} & \dots & \sigma_P^2 \mathbf{Q}_{P,P} \end{bmatrix} ; \tag{14}$$

for the solution see, for instance, Koch (1999).

CM 5: In our next approach we again assume that the observation vectors of the different techniques are mutually uncorrelated, i.e. $\mathbf{Q}_{p,q} = \mathbf{0}$ for $p \neq q$ in CM 4 (Eq. (14)). Thus, we obtain from CM 4 (Eq. (13)) the *linear model with unknown variance components*

$$\mathbf{y}_p + \mathbf{e}_p = \begin{bmatrix} \mathbf{X}_{p,p} & \mathbf{X}_{p,c} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_p \\ \boldsymbol{\beta}_c \end{bmatrix}$$
$$\text{with} \quad C(\mathbf{y}_p, \mathbf{y}_q) = \delta_{p,q} \sigma_{p,q} \mathbf{Q}_{p,q} , \tag{15}$$

wherein the delta symbol $\delta_{p,q}$ is defined as $\delta_{p,q} = 1$ for $p = q$ and $\delta_{p,q} = 0$ for $p \neq q$. Note, that e.g. the computation of the International Terrestrial Reference Frame (ITRF) is based on CM 5 (Eq. (15)); see Seitz (2015).

As a specialization of CM 5 (Eq. (15)) we remove the local parameter vectors $\boldsymbol{\beta}_p$ and obtain with $\boldsymbol{\beta}_c = \boldsymbol{\beta}$ the linear model with unknown variance component components as

$$\mathbf{y}_p + \mathbf{e}_p = \mathbf{X}_p\,\boldsymbol{\beta} \quad \text{with } C(\mathbf{y}_p, \mathbf{y}_q) = \delta_{p,q}\,\sigma_{p,q}\,\mathbf{Q}_{p,q} \tag{16}$$

as an extension of CM 1 (Eq. (7)). Applying e.g. the least squares method to CM 5 (Eq. (16)) we obtain the normal equation system

$$\left(\sum_{p=1}^{P} \frac{1}{\sigma_p^2}\,\mathbf{N}_p\right)\boldsymbol{\beta} = \sum_{p=1}^{P} \frac{1}{\sigma_p^2}\,\mathbf{b}_p\ . \tag{17}$$

In case that the system is of full rank, the estimation of the unknown vector $\boldsymbol{\beta}$ reads

$$\hat{\boldsymbol{\beta}} = \left(\sum_{p=1}^{P} \frac{1}{\sigma_p^2}\,\mathbf{N}_p\right)^{-1}\left(\sum_{p=1}^{P} \frac{1}{\sigma_p^2}\,\mathbf{b}_p\right)\ . \tag{18}$$

The variance components σ_p^2 can be estimated from the residuals

$$\hat{\mathbf{e}}_p = \mathbf{X}_p\,\hat{\boldsymbol{\beta}} - \mathbf{y}_p \tag{19}$$

according to $\hat{\sigma}_p^2 = (\hat{\mathbf{e}}_p^T\,\mathbf{P}_{p,p}\,\hat{\mathbf{e}}_p)/r_p$ by means of the partial redundancy r_p (Koch 1999). Consequently, the estimation (18) has to be performed iteratively. An efficient calculation of the variance components is given by a fast Monte-Carlo implementation of the iterative maximum-likelihood variance component estimation as described by Koch and Kusche (2001).

CM 6: In this approach we assume that *prior information* for the expectation vector $E(\boldsymbol{\beta}) = \boldsymbol{\mu}_\beta$ and the covariance matrix $D(\boldsymbol{\beta}) = \mathbf{Q}_\beta$ of the unknown parameter vector $\boldsymbol{\beta}$ are available. To be more specific we introduce the additional Gauss-Markov model

$$\boldsymbol{\mu}_\beta + \mathbf{e}_\beta = \boldsymbol{\beta} \ \text{ with } \ D(\boldsymbol{\mu}_\beta) = \sigma_\beta^2\,\mathbf{Q}_\beta \tag{20}$$

wherein $\mathbf{e}_\beta$ is the $u \times 1$ vector of the errors for the prior information and σ_β^2 an unknown variance factor of the pseudo observation vector $\boldsymbol{\mu}_\beta$. By combining the P models (Eq. (16)) and the additional model (Eq. (20)) we obtain the *linear model with unknown variance components and prior information* as

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\\ .\\ \mathbf{y}_P\\ \boldsymbol{\mu}_\beta\end{bmatrix} + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\\ .\\ \mathbf{e}_P\\ \mathbf{e}_\beta\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1\\ \mathbf{X}_2\\ .\\ \mathbf{X}_P\\ \mathbf{I}\end{bmatrix}\boldsymbol{\beta} \quad \text{with}$$

$$D(\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\\ .\\ \mathbf{y}_P\\ \boldsymbol{\mu}_\beta\end{bmatrix}) = \begin{bmatrix}\sigma_1^2\mathbf{Q}_{1,1} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \sigma_2^2\mathbf{Q}_{2,2} & \dots & \mathbf{0} & \mathbf{0}\\ \dots & \dots & \dots & \dots & \dots\\ \mathbf{0} & \mathbf{0} & \dots & \sigma_P^2\mathbf{Q}_{P,P} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \sigma_\beta^2\mathbf{Q}_\beta\end{bmatrix}\ . \tag{21}$$

The extended normal equations for the unknown parameter vector $\boldsymbol{\beta}$ including the unknown variance components σ_p^2 and σ_β^2 are given as

$$\left(\sum_{p=1}^{P} \lambda_p\,\mathbf{N}_p + \mathbf{P}_\beta\right)\hat{\boldsymbol{\beta}} = \sum_{p=1}^{P} \lambda_p\,\mathbf{b}_p + \mathbf{P}_\beta\,\boldsymbol{\mu}_\beta \tag{22}$$

with $\lambda_p = \sigma_\beta^2/\sigma_p^2$. Since the matrix $\mathbf{P}_\beta = \mathbf{Q}_\beta^{-1}$ is positive definite and the P matrices $\mathbf{N}_p$ at least positive semidefinite, the normal equation matrix $(\sum_{p=1}^{P} \lambda_p\,\mathbf{N}_p + \mathbf{P}_\beta)$ is regular and Eq. (22) can be solved for

$$\hat{\boldsymbol{\beta}} = \left(\sum_{p=1}^{P} \lambda_p\mathbf{N}_p + \mathbf{P}_\beta\right)^{-1}\left(\sum_{p=1}^{P} \lambda_p\mathbf{b}_p + \mathbf{P}_\beta\boldsymbol{\mu}_\beta\right)\ , \tag{23}$$

see Koch (2000) or Koch and Kusche (2001).

4 Numerical Example

Assuming a non-rigid Earth body the motion of the Celestial Intermediate Pole (CIP) is calculable from the differential equation

$$p(t) + \frac{i}{\Omega_{CW}}\frac{dp(t)}{dt} = \chi(t)\ , \tag{24}$$

wherein Ω_{CW} means the Chandler frequency and $i = \sqrt{-1}$; for more details see e.g. Gross (2007). Equation (24) states that the complex-valued motion of the CIP, $p(t)$, is driven by the equatorial *excitation function* $\chi(t) = \chi_1(t) + i\,\chi_2(t)$. According to $\chi_j(t) = \chi_j^{mass}(t) + \chi_j^{motion}(t)$, the components $\chi_j(t)$ with $j = 1, 2$ can each be separated into a matter term $\chi_j^{mass}(t)$ and a motion term $\chi_j^{motion}(t)$. Whereas the latter is caused by moving masses the matter term is the consequence of mass changes within the Earth system and can be calculated by the degree two spherical harmonic coefficients $\Delta C_{2,1}(t)$ and $\Delta S_{2,1}(t)$ of the gravitational potential

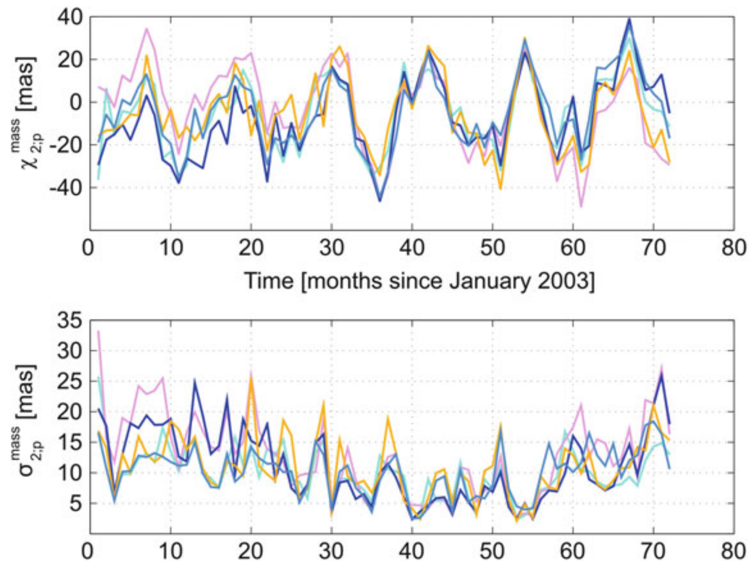

Fig. 1 *Top*: monthly time series of the excitation function $\chi^{mass}_{2;p}(t_k)$ calculated from the degree 2 spherical harmonic coefficients $\Delta S_{2,1;p}(t_k)$ from $P = 5$ GRACE ACs, namely CSR (*cyan*), GFZ (*dark blue*), JPL (*blue*), IGG (*orange*) and DEOS (*magenta*) at times t_k with $k = 1, \ldots 72$ within the time interval between January 2003 and December 2008; *bottom*: empirical standard deviations $\sigma^{mass}_{2;p}(t_k)$ of the time series; all data in mas

according to $\chi^{mass}_1(t) = \chi^{mass}_1(\Delta C_{2,1}(t))$ and $\chi^{mass}_2(t) = \chi^{mass}_2(\Delta S_{2,1}(t))$.

Several processing or analysis centers (ACs) of the GRACE K-band measurements, e.g. GFZ (GeoForschungsZentrum Potsdam), CSR (Center of Space Research, University of Texas, Austin, USA), JPL (Jet Propulsion Laboratory, Pasadena, USA), IGG (Institute for Geodesy and Geoinformation, University of Bonn) and DEOS (Delft Institute of Earth Observation and Space Systems), provide time-variable gravitational potential models. Indicating the different ACs by the index $p = 1, \ldots, P$ with P the total number of ACs we decompose $\chi^{mass}_j(t) =: \chi^{mass}_{j;p}(t)$ according to

$$\chi^{mass}_{j;p}(t) = \chi^{mass}_j(t) + \Delta\chi^{mass}_{j;p}(t) \tag{25}$$

into an *AC independent* term $\chi^{mass}_j(t)$ and an *AC dependent* correction term $\Delta\chi^{mass}_{j;p}(t)$. The first term stays for the fact that all gravitational potential models are derived from the same GRACE input data. However, different parameterizations, software packages, background models, standards and procedures (e.g. outlier detection), etc. cause AC dependent influences which can be summarized as so-called operator-software impact (OSI) parameter; see e.g. Kutterer et al. (2009) or Fang (2007). Thus, the term $\Delta\chi^{mass}_{j;p}(t)$ means the *OSI deviation* of the processed value $\chi^{mass}_{j;p}(t)$ from the "true" value $\chi^{mass}_j(t)$. Figure 1 shows exemplarily 5 monthly time series $\chi^{mass}_{2;p}(t_k)$ with $p = 1, \ldots, P = 5$ at discrete times t_k for a time span of 6 years between January 2003 and December 2008, wherein the index $k = 1, \ldots, K = 72$ indicates the 72 months starting with January 2003. As can be seen from the bottom panel of Fig. 1 the empirical monthly standard deviations $\sigma^{mass}_{2;p}(t_k)$ calculated via the relation

$$\sigma^{mass}_{j;p}(t_k) = \frac{1}{\sqrt{P-1}} \left(\sum_{\substack{q=1\\q\neq p}}^{P} \left(\chi^{mass}_{j;q}(t_k) - \chi^{mass}_{j;p}(t_k) \right)^2 \right)^{1/2}$$

are for $j = 2$ in the range between 10 and 15 mas. Usually the time series $\chi^{mass}_{j;p}(t_k)$ are considered as uncorrelated signals, although they are derived from the same GRACE raw measurements. As a consequence of this the accuracies of estimations from a combination of these time series are usually too optimistic.

For simplification we assume in the following that no systematic offsets or trends exist between the time series shown in Fig. 1 and that the OSI deviations $\Delta\chi^{mass}_{j;p}(t_k)$ could be interpreted as random variables with expectation values $E(\Delta\chi^{mass}_{j;p}(t_k)) = 0$. By introducing the $K \times 1$ observation vectors $\mathbf{y}_p = \mathbf{y} + \Delta\mathbf{y}_p = (\chi^{mass}_{j;p}(t_k))$ we reformulate CM 1 (Eq. (7)) as

$$\mathbf{y}_p + \mathbf{e}_p = \mathbf{I}_K \boldsymbol{\beta} \quad \text{with} \quad C(\mathbf{y}_p, \mathbf{y}_q) = \sigma^2 \mathbf{Q}_{p,q} \tag{26}$$

for $q, p = 1, \ldots, P$ with the $K \times K$ unit matrix $\mathbf{X}_p = \mathbf{I}_K$, the $K \times 1$ vector $\boldsymbol{\beta} = (\chi^{mass}_j(t_k))$ of the unknown excitation values $\chi^{mass}_j(t_k)$.

In the *traditional approach*, i.e. neglecting the cofactor matrices $\mathbf{Q}_{p,q}$ for $p \neq q$, the Gauss-Markov model (Eq. (26)) is reduced to CM 2 with the normal Eqs. (9). To be more specific, we define

$$\mathbf{Q}_{p,q} = \delta_{p,q}\, \mathbf{Q}_{y,y} = \delta_{p,q} \left(\frac{1}{P} \sum_{p=1}^{P} \mathbf{Q}_{y_p,y_p} \right) \tag{27}$$

with $\mathbf{Q}_{y_p,y_p} = \operatorname{diag}\left[(\sigma^{mass}_{j;p}(t_1))^2, \ldots, (\sigma^{mass}_{j;p}(t_K))^2 \right]$.

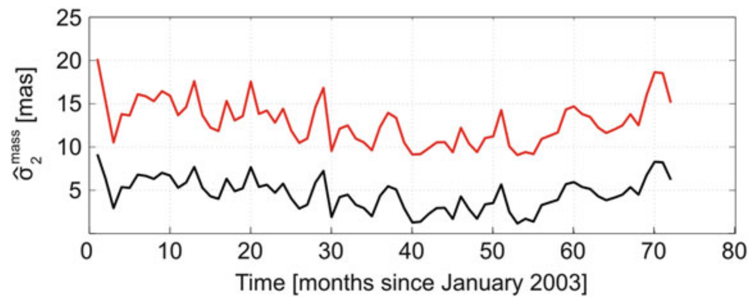

Fig. 2 Estimated standard deviations $\hat{\sigma}_2^{mass;trad}(t_k)$ and $\hat{\sigma}_2^{mass;osi}(t_k)$ from the traditional approach (*black curve*) and from the alternative approach (*red curve*)

Since the normal equation matrix is regular, the estimation $\hat{\boldsymbol{\beta}} =: \hat{\boldsymbol{\beta}}_{trad} = (\hat{\chi}_j^{mass;trad}(t_k))$ is obtained from Eq. (10).

In the *alternative approach* we consider the decomposition (25) and define for CM 1 (Eq. (7)) the covariance matrix $\mathbf{Q}_{p,q}$ as

$$\mathbf{Q}_{p,q} = \mathbf{Q}_{y,y} + \delta_{p,q}\,\mathbf{Q}_{\Delta y_p,\Delta y_p}\ , \tag{28}$$

wherein the matrix $\mathbf{Q}_{y,y}$ was already introduced by Eq. (27). According to Kutterer et al. (2009) we define the OSI deviation $\mathbf{Q}_{\Delta y_p,\Delta y_p}$ of the cofactor matrix as

$$\mathbf{Q}_{\Delta y_p,\Delta y_p} = \alpha^2\,\mathbf{Q}_{y_p,y_p}\ . \tag{29}$$

The calculation of the OSI parameter α is explained in detail by Fang (2007) and Heinkelmann et al. (2011) and is not repeated here. With a given value α the estimation $\hat{\boldsymbol{\beta}} =: \hat{\boldsymbol{\beta}}_{osi} = (\hat{\chi}_j^{mass;osi}(t_k))$ can be computed from the normal equation system of the Gauss-Markov model (Eq. (26)) considering Eqs. (28) and (29).

Due to the chosen cofactor matrices for the traditional and the alternative approach according to Eqs. (26)–(29) the estimated unknowns $\hat{\chi}_j^{mass;trad}(t_k)$ and $\hat{\chi}_j^{mass;osi}(t_k)$ (not shown here) are identical. However, as expected the estimated standard deviations are quite different as can be seen from Fig. 2. Whereas the black curve in Fig. 2, representing the standard deviations $\hat{\sigma}_2^{mass;trad}(t_k)$ of the estimated parameters according to the traditional approach, simulates unrealistic high accuracies, the range of the estimated standard deviations $\hat{\sigma}_2^{mass;osi}(t_k)$ of the alternative approach agrees well with the empirical accuracies shown in the bottom panel of Fig. 1. Thus, the results of the OSI approach seem to be more realistic.

The presented approach is just one way to consider the OSI deviations. Several other strategies are outlined by Kutterer et al. (2009), Fang (2007) or Heinkelmann et al. (2011) in detail. These procedures include variance component estimations according to CM 4 (Eq. (13)) and CM 5 (Eq. (15)).

5 Multi-scale Combination

One promising modern tool for the representation and the combination of input data from different observation techniques is the *multi-scale representation (MSR)*. The MSR means viewing on a signal under different resolutions (microscope effect). In other words the MSR provides approximations of the signal under different resolution levels. These approximations are representable as series expansions in so-called scaling functions; furthermore the differences between the approximations of two adjacent levels are called detail signals, also representable by series expansions, but this time in so-called wavelet functions.

In case of using scaling and wavelet functions as base functions for modeling the target function, e.g. the gravity field, two ways can be used for combining data from different observation techniques, namely (1) the combination of data sets from different observation techniques on the highest resolution level or (2) the estimation of the target function on different resolution levels due to the distribution and the sensitivity of the data from different techniques.

The MSR is explained in detail in many publications, e.g., by Schmidt (2007, 2012). Here we give a brief summary for the two-dimensional (2-D) case.

In the 2-D multi-scale approach we model the target function F as $F(\mathbf{x}) \approx F_J(\mathbf{x}) = F_J(x,y)$ depending on the 2-D position vector $\mathbf{x} = [x,y]^T$ on the highest resolution level J as

$$F_J(x,y) = \sum_{k_1=0}^{K_J-1}\sum_{k_2=0}^{K_J-1} d_{J;k_1,k_2}\,\phi_{J;k_1}(x)\,\phi_{J;k_2}(y)\ , \tag{30}$$

where $\phi_{J;k_1}(x)$ and $\phi_{J;k_2}(y)$ are two 1-D base functions depending on the coordinates x and y. In wavelet theory these functions are called level-J scaling functions and generate a 2-D tensor product MSR; see e.g. Schmidt (2001). As two examples the Daubechies scaling functions generate an orthogonal MSR, the endpoint-interpolating B-spline functions, however, provide a semi-orthogonal MSR; for the latter see e.g. Schmidt (2007). With the number K_J of scaling functions $\phi_{J;k}(z)$ with $z \in \{x, y\}$ and $k \in \{k_1, k_2\}$ we define the $(K_J \cdot K_J) \times 1$ vector $\mathbf{d}_J = \text{vec } \mathbf{D}_J$, wherein 'vec' means the vec-operator (Koch 1999). The $K_J \times K_J$ scaling coefficient matrix $\mathbf{D}_J$ is defined as

$$\mathbf{D}_J = \begin{bmatrix} d_{J;0,0} & d_{J;0,1} & \dots & d_{J;0,K_J-1} \\ d_{J;1,0} & d_{J;1,1} & \dots & d_{J;1,K_J-1} \\ \dots & \dots & \dots & \dots \\ d_{J;K_J-1,0} & d_{J;K_J-1,1} & \dots & d_{J;K_J-1,K_J-1} \end{bmatrix} . \quad (31)$$

With the $K_J \times 1$ level-J scaling vector $\boldsymbol{\phi}_J(z) = (\phi_{J;k}(z))$ and considering the computation rules for the Kronecker product '$\otimes$' (see e.g. Koch 1999) Eq. (30) can be rewritten as

$$\begin{aligned} F_J(x, y) &= (\boldsymbol{\phi}_J^T(y) \otimes \boldsymbol{\phi}_J^T(x)) \text{ vec } \mathbf{D}_J \\ &= \boldsymbol{\phi}_J^T(x)\, \mathbf{D}_J\, \boldsymbol{\phi}_J(y) . \end{aligned} \quad (32)$$

With $j' = J - i_0$ and $0 < i_0 \leq J$ the MSR is given as

$$F_J(x, y) = F_{j'}(x, y) + \sum_{i=1}^{i_0} \sum_{\lambda=1}^{3} G_{J-i}^{\lambda}(x, y) , \quad (33)$$

wherein the low-pass filtered signal $F_{j'}(x, y)$ and the band-pass filtered detail signals $G_{J-i}^{\lambda}(x, y)$ are computable via the relations

$$\begin{aligned} F_{j'}(x, y) &= \boldsymbol{\phi}_{j'}^T(x)\, \mathbf{D}_{j'}\, \boldsymbol{\phi}_{j'}(y) , \\ G_{J-i}^1(x, y) &= \boldsymbol{\phi}_{J-i}^T(x)\, \mathbf{C}_{J-i}^1\, \boldsymbol{\psi}_{J-i}(y) , \\ G_{J-i}^2(x, y) &= \boldsymbol{\psi}_{J-i}^T(x)\, \mathbf{C}_{J-i}^2\, \boldsymbol{\phi}_{J-i}(y) , \\ G_{J-i}^3(x, y) &= \boldsymbol{\psi}_{J-i}^T(x)\, \mathbf{C}_{J-i}^3\, \boldsymbol{\psi}_{J-i}(y) \end{aligned} \quad (34)$$

by means of the $L_j \times 1$ level-j wavelet vectors $\boldsymbol{\psi}_j(z) = (\psi_{j,l}(z))$ with $L_j = K_{j+1} - K_j$.

The level-j and level-$(j-1)$ scaling functions $\phi_{j;k}(z)$ with $k = 0, \dots, K_j - 1$ and $\phi_{j-1;n}(z)$ with $n = 0, \dots, K_{j-1} - 1$ as well as the level-$(j-1)$ wavelet functions $\psi_{j-1;l}(z)$ with $l = 0, \dots, L_{j-1} - 1$ are related to each other by means of the *two-scale relations*

$$\phi_{j-1;n}(z) = \sum_{k=2n-(K_j-1)}^{2n} p_{j;k}\, \phi_{j;2n-k}(z) , \quad (35)$$

$$\psi_{j-1;l}(z) = \sum_{k=2l-(K_j-1)}^{2l} q_{j;k}\, \phi_{j;2l-k}(z) \quad (36)$$

with given coefficients $p_{j;k}$ and $q_{j;k}$; for more details see e.g. (Schmidt 2012). The two Eqs. (35) and (36) can be reformulated as matrix equations

$$\begin{aligned} \boldsymbol{\phi}_{j-1}^T(x) &= \boldsymbol{\phi}_j^T(x)\, \mathbf{P}_j , \\ \boldsymbol{\psi}_{j-1}^T(x) &= \boldsymbol{\phi}_j^T(x)\, \mathbf{Q}_j \end{aligned} \quad (37)$$

with the $K_j \times K_{j-1}$ matrix $\mathbf{P}_j = (p_{j;k})$ and the $K_j \times L_{j-1}$ matrix $\mathbf{Q}_j = (q_{j;k})$. In Eqs. (34) we introduced the $K_{J-i} \times L_{J-i}$ matrix $\mathbf{C}_{J-i}^1$, the $L_{J-i} \times K_{J-i}$ matrix $\mathbf{C}_{J-i}^2$ and the $L_{J-i} \times L_{J-i}$ matrix $\mathbf{C}_{J-i}^3$. The corresponding 2-D downsampling equations read

$$\begin{bmatrix} \mathbf{D}_{j-1} & \mathbf{C}_{j-1}^1 \\ \mathbf{C}_{j-1}^2 & \mathbf{C}_{j-1}^3 \end{bmatrix} = \begin{bmatrix} \overline{\mathbf{P}}_j \\ \overline{\mathbf{Q}}_j \end{bmatrix} \mathbf{D}_j \begin{bmatrix} \overline{\mathbf{P}}_j^T & \overline{\mathbf{Q}}_j^T \end{bmatrix} \quad (38)$$

for $j = j' + 1, \dots, J$ and mean the *pyramid algorithm*. The $K_{j-1} \times K_j$ matrix $\overline{\mathbf{P}}_j$ and the $L_{j-1} \times K_j$ matrix $\overline{\mathbf{Q}}$ are defined via the relation

$$\begin{bmatrix} \overline{\mathbf{P}}_j \\ \overline{\mathbf{Q}}_j \end{bmatrix} = \begin{bmatrix} \mathbf{P}_j & \mathbf{Q}_j \end{bmatrix}^{-1} . \quad (39)$$

Viewing on the rigorous combination explained before from the point of the MSR we (1) have to choose an appropriate number for the highest resolution level J defined in Eq. (30), (2) we perform the parameter estimation by means of an appropriate CM as defined in Sect. 3 and calculate the series coefficients collected in the matrix $\mathbf{D}_J$ (Eq. (31)) and (3) we calculate all detail signals by applying the pyramid algorithm according to Eqs. (34) and (38). Since all calculations within the pyramid algorithm are based on linear equation systems the covariance matrices of all sets of coefficients and the detail signals can be computed by means of the law of error propagation.

As an alternative to the procedure described before the spectral behavior of the different observation techniques can be used directly in the estimation procedure. Since, for instance, GOCE data cover a higher frequency part than GRACE data, the MSR comes directly into play. The procedure is visualized by the flowchart in Fig. 3. To be more specific, the high-resolution observation vector $\mathbf{y}_1$ with cofactor matrix $\mathbf{Q}_{1,1}$ of the first observation technique determines the coefficient matrix $\mathbf{D}_J$ (Eq. (31)) via a CM. The detail signals $G_{J-1} \in \{G_{J-1}^1, G_{J-1}^2, G_{J-1}^3\}$ to $G_{j_2+1} \in \{G_{j_2+1}^1, G_{j_2+1}^2, G_{j_2+1}^3\}$ of the highest levels $j_2 + 1, j_2 + 2, \dots, J - 1$ are calculated via Eqs. (34) and (38). Applying in the next step again the pyramid algorithm the scaling coefficient matrix $\mathbf{D}_{j_2}$ of level j_2 is predicted as

$$\mathbf{D}_{j_2} = \overline{\mathbf{P}}_{j_2+1}\, \mathbf{D}_{j_2+1}\, \overline{\mathbf{P}}_{j_2+1}^T \quad (40)$$

with covariance matrix

$$\begin{aligned} D(\text{vec}\mathbf{D}_{j_2}) &= (\overline{\mathbf{P}}_{j_2+1} \otimes \overline{\mathbf{P}}_{j_2+1})\, D(\text{vec}\mathbf{D}_{j_2+1}) \\ &\quad \cdot (\overline{\mathbf{P}}_{j_2+1}^T \otimes \overline{\mathbf{P}}_{j_2+1}^T) \end{aligned} \quad (41)$$

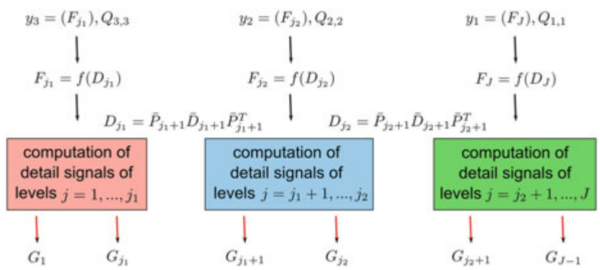

Fig. 3 Data combination within a MSR. The input data of different observation techniques are introduced into the evaluation model at different resolution levels. The coefficient matrix $\mathbf{D}_{j_2+1}$, estimated at the level $j_2 + 1$ is transferred to the lower level j_2 by means of the pyramid algorithm according to Eq. (40)

and improved by introducing the observation vector $\mathbf{y}_2$ with cofactor matrix $\mathbf{Q}_{2,2}$ of the second observation technique providing data of medium resolution. Continuing this process until the lowest level the MSR of the gravity field according to Eq. (30) is obtained.

Conclusion

In this paper we presented selected combination strategies for parameter estimation from input data of different geodetic observation techniques. Starting with the Gauss-Markov model we outline parameter estimation within different combination models including the estimation of unknown variance and covariance components. If furthermore prior information is available it can be used as additional information, e.g. by introducing pseudo observations from a given gravity field model. As discussed in various papers the variance components of the real observations and the prior information express the relative weighting and may be interpreted as regularization parameters; see e.g. Koch and Kusche (2001).

As an alternative approach to the rigorous combination we discussed the MSR for combining input data from observation techniques which are sensitive to different parts of the frequency spectrum. A further advantage of the MSR based on wavelet expansion is that usually a lot of elements of the matrices $\mathbf{C}_j^\lambda$ introduced in Eq. (34) are close to zero. Thus, as already standard in digital image processing efficient data compression techniques can be applied. This way, just the significant information of the signal is stored. Finally, it is worth to mention that all methods mentioned before can be formulated in the framework of Bayesian Inference (Koch 2000).

We discussed briefly the problem of the OSI deviation. This example demonstrates the importance of an realistic stochastic model. In our opinion it cannot be the main goal of data evaluation to achieve the smallest values for the estimated standard deviations. Instead of this we aim at *realistic accuracies*!

Finally we want to mention that future data combination goes far beyond the combination of geodetic techniques, since especially data from remote sensing missions have to be connected with geodetic data. In such a procedure, e.g. point measurements from GNSS have to be combined with superficial measurements from InSAR. These kinds of combination mean a further challenge for geodesy in the very near future.

References

Dettmering D, Schmidt M, Heinkelmann R, Seitz M (2011) Combination of different satellite observation data for regional ionosphere modeling. J Geod 85(12):989–998. doi:10.1007/s00190-010-0423-1
Fang X (2007) Statistische Analyse von Ausgleichungsergebnissen auf der Basis teilweise identischer Messwerte. Diploma thesis, Geodetic Institute, Leibniz University of Hannover
Gitlein O, Denker H, Müller J (2007) Local geoid computation by the spectral combination method. In: Jekeli C, Bastos L, Fernandes J (eds) Gravity, geoid and space missions. IAG symposia, vol 129. Springer, Berlin, pp 179–184
Gross R (2007) Earth rotation variations - long period. In: Schubert G (ed) Treatise on geophysics, vol 3 (Geodesy), Sect. 3.11. Elsevier, Amsterdam
Haagmans R, Prijatna K, Dahl-Omang O (2002) An alternative concept for validation of GOCE gradiometry results based on regional gravity. In: Proceedings of GG2002 meeting in Thessaloniki, pp 281–286
Heinkelmann R et al (2011) VLBI-derived troposphere parameter during CONT08. J Geod. doi:10.1007/s00190-011-0459-x
Koch KR (1999) Parameter estimation and hypothesis testing in linear models. Springer, Berlin
Koch KR (2000) Einführung in die Bayes-Statistik. Springer, Berlin
Koch KR, Kusche J (2001) Regularization of geopotential determination from satellite data by variance components. J Geod 76:259–268
Kühtreiber N, Abd-Elmotaal HA (2007) Ideal combination of deflection components and gravity anomalies for precise geoid computation. In: Tregonin P, Rizos C (eds) Dynamic planet, monitoring and understanding a dynamic planet with geodetic and oceanographic tools. IAG symposia, vol 130. Springer, Berlin, pp 259–265
Kutterer H, Krügel M, Tesmer V (2009). Towards an improved assessment of the quality of terrestrial reference frames. In: Drewes H (ed) Geodetic reference frames. IAG symposia, 134. Springer, New York, pp 67–72
Rummel R, Rothacher M, Beutler G (2005) Global geodetic observing system (GGOS): science rationale. J Geodyn 40:357–362
Schmidt M (2001) Grundprinzipien der Wavelet-Analyse und Anwendungen in der Geodäsie. Post doctoral thesis, Shaker, Aachen
Schmidt M (2007) Wavelet modeling in support of IRI. J Adv Sp Res 39:932–940
Schmidt M (2012) Towards a multi-scale representation of multi-dimensional signals. In: Sneeuw N et al (eds) VII Hotine-Marussi symposium on mathematical geodesy. IAG symposia, vol 137. Springer, Berlin, pp 119–127
Seitz M (2015) Combination of different space geodetic data types in order to compute terrestrial reference frames and time series of geophysical parameters. In: Kutterer H, Seitz F, Schmidt M (eds) Proceedings of the 1st international workshop on the quality of geodetic observation and monitoring system (QuGoMS), Munich,

2011. International association of geodesy symposia. Springer, Heidelberg, this volume
van Loon J (2008) Functional and stochastic modelling of satellite gravity data. In: Publication on geodesy, vol 67. Netherlands Geodetic Commission, Delft University of Technology
van Loon J, Kusche J (2007) Towards an optimal combination of satellite data and prior information. In: Tregonin P, Rizos C (eds) Dynamic planet, monitoring and understanding a dynamic planet with geodetic and oceanographic tools. IAG symposia, vol 130. Springer, Berlin, pp 345–353