Journal of Geodesy (2002) 76: 259–268
DOI 10.1007/s00190-002-0245-x

Journal of Geodesy

Regularization of geopotential determination from satellite data by variance components

K.-R. Koch[1], **J. Kusche**[2]

[1] Institute for Theoretical Geodesy, University of Bonn, Nussallee 17, 53115 Bonn, Germany
e-mail: koch@theor.geod.uni-bonn.de; Tel.: +49-228-73-3395; Fax: +49-228-73-3029
[2] Delft Institute for Earth-Oriented Space Research (DEOS), Delft University of Technology, Thijsseweg 11, 2629 JA Delft, The Netherlands; e-mail: j.kusche@citg.tudelft.nl; Tel.: +31-1527-82562; Fax: +31-1527-83711

Received: 5 June 2001 / Accepted: 28 November 2001

Abstract. Different types of present or future satellite data have to be combined by applying appropriate weighting for the determination of the gravity field of the Earth, for instance GPS observations for CHAMP with satellite to satellite tracking for the coming mission GRACE as well as gradiometer measurements for GOCE. In addition, the estimate of the geopotential has to be smoothed or regularized because of the inversion problem. It is proposed to solve these two tasks by Bayesian inference on variance components. The estimates of the variance components are computed by a stochastic estimator of the traces of matrices connected with the inverse of the matrix of normal equations, thus leading to a new method for determining variance components for large linear systems. The posterior density function for the variance components, weighting factors and regularization parameters are given in order to compute the confidence intervals for these quantities. Test computations with simulated gradiometer observations for GOCE and satellite to satellite tracking for GRACE show the validity of the approach.

Keywords: Geopotential Determination – Regularization – Variance Components – Bayesian Statistics – Confidence Intervals

1 Introduction

The satellite mission CHAMP [GeoForschungsZentrum (GFZ) 2000] and the future mission GRACE [Jet Propulsion Laboratory (JPL) 1999] with its satellite-to-satellite tracking as well as GOCE [European Space Agency (ESA) 1999] with its gradiometer measurements will considerably improve the knowledge of the gravity field of the Earth. A large number of parameters of the geopotential, for instance the coefficients of an expansion into spherical harmonics, have to be determined. Unfortunately, the matrix of normal equations for the solution of the unknown gravity field parameters tends to be ill-conditioned. This fact is due to the inversion problem, i.e. the geopotential is determined by measurements outside the surface of the Earth; it already occurred with all determinations of the gravity field of the Earth by satellite observations. Another reason for ill-conditioned systems are data gaps such as the polar gaps in the GOCE mission.

Usually a Tikhonov-regularization (Tikhonov and Arsenin 1977, p. 103) has been applied to smooth or regularize the determination. Thus, a positive definite matrix times the regularization parameter is added to the matrix of normal equations to stabilize the solution. This regularization is also known as ridge regression (see e.g. Vinod and Ullah 1981, p. 169), and can be interpreted by Bayesian statistics as estimation with prior information (see e.g. O'Sullivan 1986; Koch 1990, p. 70; 2000, p. 119). The matrix to be added to the matrix of normal equations is the inverse of the covariance matrix of the parameters which is given by prior knowledge. In the Kaula regularization, for instance, the covariance matrix is a diagonal matrix with the variances of the harmonic coefficients given by Kaula's rule of thumb (Kaula, 1966, p. 98). This method has been frequently applied for gravity field determinations from satellite data (see e.g. Reigber and Ilk 1976; Marsh et al. 1988; Schwintzer et al. 1997). Measures for the contribution of the observations for regularized gravity field solutions have been discussed by Bouman (2000, p. 29).

Statistical properties of the regularization parameter or the ridge parameter for geopotential determinations have been investigated. For instance, Xu and Rummel (1994) derive the variance of the ridge parameter and discuss iterative solutions for estimating it. Xu and Rummel (1994) use the generalized ridge estimator to determine the regularization parameter, while Xu (1998) applies the truncated singular value decomposition together with F-statistics.

Correspondence to: J. Kusche

A further method which has been frequently applied to determine the regularization parameter is the generalized cross-validation by Wahba (1977), (see also Craven and Wahba 1979; Golub et al. 1979). It minimizes the weighted sum of squares of the residuals divided by a quantity containing the trace of the inverse of the matrix of normal equations. A large number of gravity field parameters have to be determined by the GRACE and GOCE missions, for which iterative methods will be applied to solve the large systems of normal equations (Kusche 2001). The computation of the inverse of the matrix of the normal equations is avoided so that the trace of the inverse cannot be directly obtained.

However, a stochastic estimator has been presented by Girard (1989) for computing the necessary trace in a cross-validation. The method has been refined by Hutchinson (1990) (see also Golub and Von Matt 1997), and applied to the regularization of the results of simulated gradiometer observations for GOCE by Kusche and Klees (2001). When estimating variance components in large systems of normal equations, the traces of matrices connected with the inverses of the normal equations also have to be computed. Stochastic estimators can therefore be applied, as shown in the following.

When combining heterogeneous data in one parameter estimation the proper weighting needs to be chosen. For estimating the gravity field of the Earth by satellite data this problem was solved by Lerch (1991). He required the differences of the parameters of subset solutions and the complete solution to be in agreement by adjusting the data weights iteratively. A different approach is given by estimating variance components, which goes back to Helmert (1924, p. 358). Rao (1973, p. 302) developed the MINQUE theory for estimating variance and covariance components, and Grafarend and D'Hone (1978) investigated estimation methods for analyzing geodetic data. In the following years many papers on the evaluation of geodetic data by variance components were published, quite a number are referenced by Crocetto et al. (2000) or Koch (1999a, p. 227).

If as mentioned above the Tikhonov regularization is interpreted as prior information in the Bayesian sense, the regularization parameter can be obtained as the ratio of two variance components, as proposed by Arsenin and Krianev (1992). For regularizing geopotential determinations we therefore apply Bayesian inference on variance components (see e.g. Koch 1987, Ou 1991). In order to compute the traces for the estimates of the variance components, we use the stochastic trace estimator of Hutchinson (1990). This leads to a new method of determining variance components which is still feasible for very large systems of normal equations for which the computation of the inverse is avoided. It can be shown that the posterior density function for the variance components results from an inverse gamma distribution (Ou and Koch 1994; Koch 2000, p. 145). Confidence intervals for the variance components can therefore be established, or hypotheses for the variance components can be tested. Finally, the posterior density function for the ratio of two variance components, which gives the weighting factor or regularization parameter, is derived. Thus, the statistical analysis of the regularization parameter can go beyond the one suggested, for instance, by Xu and Rummel (1994).

The present paper is organized as follows. Section 2 presents the linear model with unknown variance components for heterogeneous data and prior information on the parameters. The estimates of the variance components lead to the weighting of the data of different types and to the regularization parameter. Section 3 deals with the stochastic trace estimation and Sect. 4 presents the posterior density function for the variance components and the regularization parameter. Section 5 finally gives an example for regularizing results from simulated gradiometer measurements for GOCE and satellite-to-satellite tracking for GRACE.

2 Variance components

The satellite data for determining the gravity field of the Earth will be analyzed in the linear model which in the formulation of Bayesian statistics is given by, (see e.g. Koch 2000, p. 85)

$$\mathbf{X}\boldsymbol{\beta} = E(\mathbf{y}|\boldsymbol{\beta}) \quad \text{with} \quad D(\mathbf{y}|\sigma^2) = \sigma^2\mathbf{P}^{-1} \tag{1}$$

where $\mathbf{X}$ denotes the $n \times u$ matrix of given coefficients, the so-called design matrix, which will be assumed of full column rank, although ill-conditioned normal equations are expected. $\boldsymbol{\beta}$ is the $u \times 1$ vector of unknown random parameters, σ^2 the unknown variance factor and $\mathbf{P}$ the known $n \times n$ positive definite weight matrix of the observations $\mathbf{y}$. In the following we also use the alternative to the linear model of Eq. (1) given by the observation equations

$$\mathbf{X}\boldsymbol{\beta} = \mathbf{y} + \mathbf{e} \quad \text{with} \quad E(\mathbf{e}|\boldsymbol{\beta}) = \mathbf{0} \quad \text{and}$$
$$D(\mathbf{e}|\boldsymbol{\beta}, \sigma^2) = D(\mathbf{y}|\sigma^2) = \sigma^2\mathbf{P}^{-1} \tag{2}$$

where $\mathbf{e}$ denotes the vector of errors of the observations.

When determining the gravity field of the Earth by satellite data we have to keep in mind that Eq. (1) or (2) contains a modeling error, also called aliasing. It is caused by the fact that when globally computing the gravity field the number of gravity field parameters has to be finite, although it should be infinite, and when locally determining the geopotential the effects of outer zones are neglected. An estimate of the variance factor larger than expected points to a modeling error (see e.g. Koch 1999, p. 180).

Different types of observations such as satellite-to-satellite tracking or gradiometer measurements enter into the determination of the gravity field, leading to the observation equations

$$\mathbf{X}_i\boldsymbol{\beta} = \mathbf{y}_i + \mathbf{e}_i \quad \text{with} \quad D(\mathbf{y}_i|\sigma_i^2) = \sigma_i^2\mathbf{P}_i^{-1} \quad \text{and}$$
$$i \in \{1, \ldots, o\} \tag{3}$$

if o different kinds of observations $\mathbf{y}_i$ are assumed. The smoothing or the regularization of the determination of the gravity field can be interpreted as being achieved by prior information, as already mentioned in Sect. 1. In agreement with the Bayesian approach, not only the covariance matrix of the unknown parameters is introduced as prior information, as happens for the Tikhonov regularization, but also the expected value of the unknown parameters, which gives the observation equation

$$\boldsymbol{\beta} = \boldsymbol{\mu} + \mathbf{e}_\mu \quad \text{with} \quad D(\boldsymbol{\mu}|\sigma_\mu^2) = \sigma_\mu^2 \mathbf{P}_\mu^{-1} \tag{4}$$

with the $u \times 1$ vector $\boldsymbol{\mu}$ of prior information on the unknown parameters, the variance factor σ_μ^2 and the $u \times u$ weight matrix $\mathbf{P}_\mu$ of the parameters. The vector $\boldsymbol{\mu}$ is especially helpful for gravity field determinations since the long-wave components of the geopotential are known with given variances and covariances so that by Eq. (4) realistic prior information can be introduced (see e.g. Blinken and Koch 2001).

The observations $\mathbf{y}_i$ in Eq. (3) and the prior information $\boldsymbol{\mu}$ in Eq. (4) can be assumed as being independent. The variance factors σ_i^2 in Eq. (3) are unknown random parameters since the weighting of different kinds of observations is generally unknown. The variance factor σ_μ^2 is also assumed as an unknown random parameter to account for a regularization parameter. Small variance factors ${\sigma_\mu}^2$ will lead to more rigorous regularizations, while larger variance factors give weaker regularizations. The variance factors σ_i^2 for $i \in \{1, \ldots, o\}$ and σ_μ^2 are considered as unknown variance components so that substituting Eqs. (3) and (4) in Eq. (2) gives the linear model with unknown variance components

$$\begin{vmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_o \\ \mathbf{I} \end{vmatrix} \boldsymbol{\beta} = \begin{vmatrix} \mathbf{y}_1 + \mathbf{e}_1 \\ \vdots \\ \mathbf{y}_o + \mathbf{e}_o \\ \boldsymbol{\mu} + \mathbf{e}_\mu \end{vmatrix} \quad \text{with}$$

$$D\left(\begin{vmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_o \\ \boldsymbol{\mu} \end{vmatrix}\right) = \sigma_1^2 \mathbf{V}_1 + \cdots + \sigma_o^2 \mathbf{V}_o + \sigma_\mu^2 \mathbf{V}_\mu \tag{5}$$

and

$$\mathbf{V}_1 = \begin{vmatrix} \mathbf{P}_1^{-1} & \cdots & \mathbf{0} & \mathbf{0} \\ & \vdots & & \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{vmatrix}, \ldots, \mathbf{V}_o = \begin{vmatrix} \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ & \vdots & & \\ \mathbf{0} & \cdots & \mathbf{P}_o^{-1} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{vmatrix},$$

$$\mathbf{V}_\mu = \begin{vmatrix} \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ & \vdots & & \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{P}_\mu^{-1} \end{vmatrix}$$

The estimates $\hat{\boldsymbol{\beta}}$ of the unknown parameters $\boldsymbol{\beta}$ follow from the normal equations

$$\begin{aligned} &\left(\frac{1}{\sigma_1^2}\mathbf{X}_1'\mathbf{P}_1\mathbf{X}_1 + \cdots + \frac{1}{\sigma_o^2}\mathbf{X}_o'\mathbf{P}_o\mathbf{X}_o + \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\right)\hat{\boldsymbol{\beta}} \\ &\quad = \frac{1}{\sigma_1^2}\mathbf{X}_1'\mathbf{P}_1\mathbf{y}_1 + \cdots + \frac{1}{\sigma_o^2}\mathbf{X}_o'\mathbf{P}_o\mathbf{y}_o + \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\boldsymbol{\mu} \end{aligned} \tag{6}$$

In the case that there is only one type of observations $\mathbf{y}_1$ together with the prior information $\boldsymbol{\mu}$, we obtain with $o = 1$ in Eq. (6) the normal equations

$$\left(\frac{1}{\sigma_1^2}\mathbf{X}_1'\mathbf{P}_1\mathbf{X}_1 + \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\right)\hat{\boldsymbol{\beta}} = \frac{1}{\sigma_1^2}\mathbf{X}_1'\mathbf{P}_1\mathbf{y}_1 + \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\boldsymbol{\mu}. \tag{7}$$

By introducing the regularization parameter λ with

$$\lambda = \frac{\sigma_1^2}{\sigma_\mu^2} \tag{8}$$

we obtain

$$(\mathbf{X}_1'\mathbf{P}_1\mathbf{X}_1 + \lambda\mathbf{P}_\mu)\hat{\boldsymbol{\beta}} = \mathbf{X}_1'\mathbf{P}_1\mathbf{y}_1 + \lambda\mathbf{P}_\mu\boldsymbol{\mu} \tag{9}$$

With $\boldsymbol{\mu} = \mathbf{0}$ the solution vector for the Tikhonov regularization, the ridge regression and the cross-validation is obtained. For two different types of observations such as gradiometer measurements and satellite-to-satellite tracking we find with $o = 2$ in Eq. (6) and with

$$\lambda = \frac{\sigma_1^2}{\sigma_\mu^2} \quad \text{and} \quad \omega = \frac{\sigma_1^2}{\sigma_2^2} \tag{10}$$

the normal equations

$$\begin{aligned} (\mathbf{X}_1'\mathbf{P}_1\mathbf{X}_1 + \omega\mathbf{X}_2'\mathbf{P}_2\mathbf{X}_2 + \lambda\mathbf{P}_\mu)\hat{\boldsymbol{\beta}} &= \mathbf{X}_1'\mathbf{P}_1\mathbf{y}_1 + \omega\mathbf{X}_2'\mathbf{P}_2\mathbf{y}_2 \\ &\quad + \lambda\mathbf{P}_\mu\boldsymbol{\mu}. \end{aligned} \tag{11}$$

The ratio σ_1^2/σ_μ^2 again gives the regularization parameter λ while the ratio ω expresses the relative weighting of the observations $\mathbf{y}_2$ with respect to $\mathbf{y}_1$.

The unknown variance components σ_i^2 for $i \in \{1, \ldots, o\}$ and σ_μ^2 are estimated iteratively. Either approximate values are introduced for the variance components such that their values are close to one and one iterates until at the point of convergence the estimated values are equal to one, or starting from approximate values the estimates of the variance components are iteratively computed until they converge. For the latter case we obtain the estimates, which may be derived by the maximum-likelihood method, by the best invariant quadratic unbiased estimation or by Helmert's method (see e.g. Förstner 1979; Crocetto et al. (2000); Koch 2000, p. 146)

$$\hat{\sigma}_i^2 = \hat{\mathbf{e}}_i'\mathbf{P}_i\hat{\mathbf{e}}_i/r_i\,, \quad i \in \{1, \ldots, o\} \tag{12}$$

and

$$\hat{\sigma}_\mu^2 = \hat{\mathbf{e}}_\mu'\mathbf{P}_\mu\hat{\mathbf{e}}_\mu/r_\mu \tag{13}$$

where $\hat{\mathbf{e}}_i$ and $\hat{\mathbf{e}}_\mu$ denote the vectors of residuals

$$
\begin{aligned}
\hat{\mathbf{e}}_i &= \mathbf{X}_i\hat{\boldsymbol{\beta}} - \mathbf{y}_i \\
\hat{\mathbf{e}}_\mu &= \hat{\boldsymbol{\beta}} - \boldsymbol{\mu}
\end{aligned}
\tag{14}
$$

and r_i and r_μ the partial redundancies, i.e. the contributions of the observations $\mathbf{y}_i$ and the prior information $\boldsymbol{\mu}$ to the overall redundancy $n + u - u = n$ of the model of Eq. (5). The partial redundancies are computed from

$$
r_i = \mathrm{tr}(\mathbf{W}(\sigma_i^2\mathbf{V}_i)) \quad \text{and} \quad r_\mu = \mathrm{tr}(\mathbf{W}(\sigma_\mu^2\mathbf{V}_\mu))
\tag{15}
$$

with the matrix $\mathbf{W}$, which is the matrix of the quadratic form of the residuals (Koch 1999, p. 230), as follows:

$$
\mathbf{W} = \begin{vmatrix} \frac{1}{\sigma_1^2}\mathbf{P}_1 \cdots & \mathbf{0} & \mathbf{0} \\ & \vdots & \\ \mathbf{0} & \cdots \frac{1}{\sigma_o^2}\mathbf{P}_o & \mathbf{0} \\ \mathbf{0} & \cdots \ \mathbf{0} & \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu \end{vmatrix}
- \begin{vmatrix} \frac{1}{\sigma_1^2}\mathbf{P}_1 \cdots & \mathbf{0} & \mathbf{0} \\ & \vdots & \\ \mathbf{0} & \cdots \frac{1}{\sigma_o^2}\mathbf{P}_o & \mathbf{0} \\ \mathbf{0} & \cdots \ \mathbf{0} & \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu \end{vmatrix}
\begin{vmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_o \\ \mathbf{I} \end{vmatrix} \mathbf{N}^{-1}
$$

$$
|\mathbf{X}_1', \ldots, \mathbf{X}_o', \mathbf{I}|
\begin{vmatrix} \frac{1}{\sigma_1^2}\mathbf{P}_1 \cdots & \mathbf{0} & \mathbf{0} \\ & \vdots & \\ \mathbf{0} & \cdots \frac{1}{\sigma_o^2}\mathbf{P}_o & \mathbf{0} \\ \mathbf{0} & \cdots \ \mathbf{0} & \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu \end{vmatrix}
= \begin{vmatrix} \frac{1}{\sigma_1^2}\mathbf{P}_1 \cdots & \mathbf{0} & \mathbf{0} \\ & \vdots & \\ \mathbf{0} & \cdots \frac{1}{\sigma_o^2}\mathbf{P}_o & \mathbf{0} \\ \mathbf{0} & \cdots \ \mathbf{0} & \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu \end{vmatrix}
$$

$$
- \begin{vmatrix}
\frac{1}{\sigma_1^2}\mathbf{P}_1\mathbf{X}_1\mathbf{N}^{-1}\mathbf{X}_1'\left(\frac{1}{\sigma_1^2}\mathbf{P}_1\right) \cdots \frac{1}{\sigma_1^2}\mathbf{P}_1\mathbf{X}_1\mathbf{N}^{-1}\mathbf{X}_0'\left(\frac{1}{\sigma_o^2}\mathbf{P}_o\right) & \frac{1}{\sigma_1^2}\mathbf{P}_1\mathbf{X}_1\mathbf{N}^{-1}\left(\frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\right) \\
\vdots & \\
\frac{1}{\sigma_o^2}\mathbf{P}_o\mathbf{X}_o\mathbf{N}^{-1}\mathbf{X}_1'\left(\frac{1}{\sigma_1^2}\mathbf{P}_1\right) \cdots \frac{1}{\sigma_o^2}\mathbf{P}_o\mathbf{X}_o\mathbf{N}^{-1}\mathbf{X}_o'\left(\frac{1}{\sigma_o^2}\mathbf{P}_o\right) & \frac{1}{\sigma_o^2}\mathbf{P}_o\mathbf{X}_o\mathbf{N}^{-1}\left(\frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\right) \\
\frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\mathbf{N}^{-1}\mathbf{X}_1'\left(\frac{1}{\sigma_1^2}\mathbf{P}_1\right) \cdots \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\mathbf{N}^{-1}\mathbf{X}_o'\left(\frac{1}{\sigma_o^2}\mathbf{P}_o\right) & \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\mathbf{N}^{-1}\left(\frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\right)
\end{vmatrix}
$$

and

$$
\mathbf{N} = \frac{1}{\sigma_1^2}\mathbf{X}_1'\mathbf{P}_1\mathbf{X}_1 + \cdots + \frac{1}{\sigma_o^2}\mathbf{X}_o'\mathbf{P}_o\mathbf{X}_o + \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu .
\tag{16}
$$

The products $\mathbf{W}(\sigma_1^2\mathbf{V}_1), \ldots, \mathbf{W}(\sigma_o^2\mathbf{V}_o), \mathbf{W}(\sigma_\mu^2\mathbf{V}_\mu)$ are obtained by

$$
\mathbf{W}(\sigma_1^2\mathbf{V}_1) = \begin{vmatrix} \mathbf{I} & \cdots & \mathbf{0} & \mathbf{0} \\ & \vdots & & \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{vmatrix}
- \begin{vmatrix} \frac{1}{\sigma_1^2}\mathbf{P}_1\mathbf{X}_1\mathbf{N}^{-1}\mathbf{X}_1' & \cdots & \mathbf{0} & \mathbf{0} \\ & \vdots & & \\ \frac{1}{\sigma_o^2}\mathbf{P}_o\mathbf{X}_o\mathbf{N}^{-1}\mathbf{X}_1' & \cdots & \mathbf{0} & \mathbf{0} \\ \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\mathbf{N}^{-1}\mathbf{X}_1' & \cdots & \mathbf{0} & \mathbf{0} \end{vmatrix}
$$

$$
\mathbf{W}(\sigma_0^2\mathbf{V}_0) = \begin{vmatrix} \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ & \vdots & & \\ \mathbf{0} & \cdots & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{vmatrix}
- \begin{vmatrix} \mathbf{0} & \cdots & \frac{1}{\sigma_1^2}\mathbf{P}_1\mathbf{X}_1\mathbf{N}^{-1}\mathbf{X}_o' & \mathbf{0} \\ & \vdots & & \\ \mathbf{0} & \cdots & \frac{1}{\sigma_o^2}\mathbf{P}_o\mathbf{X}_o\mathbf{N}^{-1}\mathbf{X}_o' & \mathbf{0} \\ \mathbf{0} & \cdots & \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\mathbf{N}^{-1}\mathbf{X}_o' & \mathbf{0} \end{vmatrix}
$$

$$
\mathbf{W}(\sigma_\mu^2\mathbf{V}_\mu) = \begin{vmatrix} \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ & \vdots & & \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{I} \end{vmatrix}
- \begin{vmatrix} \mathbf{0} & \cdots & \mathbf{0} & \frac{1}{\sigma_1^2}\mathbf{P}_1\mathbf{X}_1\mathbf{N}^{-1} \\ & \vdots & & \\ \mathbf{0} & \cdots & \mathbf{0} & \frac{1}{\sigma_o^2}\mathbf{P}_o\mathbf{X}_o\mathbf{N}^{-1} \\ \mathbf{0} & \cdots & \mathbf{0} & \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\mathbf{N}^{-1} \end{vmatrix}
\tag{17}
$$

The redundancy numbers r_i and r_μ therefore follow from Eq. (15) by

$$
\begin{aligned}
r_i &= n_i - \mathrm{tr}\Big(\frac{1}{\sigma_i^2}\mathbf{X}_i'\mathbf{P}_i\mathbf{X}_i\mathbf{N}^{-1}\Big) , \quad i \in \{1, \ldots, o\} \\
r_\mu &= u - \mathrm{tr}\Big(\frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\mathbf{N}^{-1}\Big)
\end{aligned}
\tag{18}
$$

where n_i denotes the number of observations $\mathbf{y}_i$ with $\sum_{i=1}^{o} n_i = n$.

3 Stochastic trace estimation

For the computation of the partial redundancies $r_1, \ldots, r_o, r_\mu$ in Eq. (18) the inverse $\mathbf{N}^{-1}$ of the matrix of normal equations from Eq. (16) is required. However, it is not available for large systems of normal equations which are only solved and not inverted. A similar problem is encountered for the cross-validation of large-scale systems. As already mentioned in Sect. 1, it is solved by a stochastic trace estimation which will be also applied here.

We use the theorem by Hutchinson (1990)

$$
E(\mathbf{u}'\mathbf{B}\mathbf{u}) = \mathrm{tr}\mathbf{B}
\tag{19}
$$

where $\mathbf{B}$ denotes a symmetric $n \times n$ matrix and $\mathbf{u}$ an $n \times 1$ vector of n independent samples from a random variable U with $E(U) = 0$ and $V(U) = 1$. If U is a discrete random variable which takes with probability 1/2 the values -1 and $+1$, then $\mathbf{u}'\mathbf{B}\mathbf{u}$ is an unbiased estimator of $\mathrm{tr}\mathbf{B}$ with minimum variance.

In order to apply Eq. (19) symmetric matrices have to be present in Eq. (18). They are obtained by a singular value decomposition of the matrices $\mathbf{X}_i'\mathbf{P}\mathbf{X}_i$ of normal equations and of the weight matrix $\mathbf{P}_\mu$, or with less computational effort by a Cholesky factorization. For gravity field determinations the matrices $\mathbf{X}_i'\mathbf{P}\mathbf{X}_i$ tend to be ill-conditioned, so that the Cholesky factorization might not work properly. It is therefore preferable to apply the Cholesky factorization to the weight matrices $\mathbf{P}_i$ and $\mathbf{P}_\mu$ which will already be diagonal or have strong diagonal elements. With

$$\mathbf{P}_i = \mathbf{G}_i\mathbf{G}_i' \quad \text{and} \quad \mathbf{P}_\mu = \mathbf{G}_\mu\mathbf{G}_\mu' \tag{20}$$

where $\mathbf{G}_i$ and $\mathbf{G}_\mu$ denote regular lower triangular matrices, we obtain instead of Eq. (18)

$$r_i = n_i - \mathrm{tr}\left(\frac{1}{\sigma_i^2}\mathbf{G}_i'\mathbf{X}_i\mathbf{N}^{-1}\mathbf{X}_i'\mathbf{G}_i\right), \quad i \in \{1, \ldots, o\}$$

$$r_\mu = u - \mathrm{tr}\left(\frac{1}{\sigma_\mu^2}\mathbf{G}_\mu'\mathbf{N}^{-1}\mathbf{G}_\mu\right) \tag{21}$$

The Cholesky factorization may be approximately computed, for instance by

$$\mathbf{G}_i = \mathrm{diag}\left(\sqrt{p_{i11}}, \ldots, \sqrt{p_{in_in_i}}\right) \quad \text{with} \quad \mathbf{P}_i = (p_{ijk}) \tag{22}$$

because the results can be checked and improved by the relation following from Eq. (16) and (18)

$$\sum_{i=1}^{o} r_i + r_\mu = \sum_{i=1}^{o} n_i + u - \mathrm{tr}\left(\left(\frac{1}{\sigma_1^2}\mathbf{X}_1'\mathbf{P}_1\mathbf{X}_1 + \cdots + \frac{1}{\sigma_o^2}\mathbf{X}_o'\mathbf{P}_o\mathbf{X}_o + \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\right)\mathbf{N}^{-1}\right) = n \tag{23}$$

By inserting the symmetric matrices of Eq. (21) into Eq. (19) the products

$$\mathbf{u}'\mathbf{G}_i'\mathbf{X}_i\mathbf{N}^{-1}\mathbf{X}_i'\mathbf{G}_i\mathbf{u} \tag{24}$$

need to be determined. To avoid the computation of the inverse $\mathbf{N}^{-1}$, the unknown parameter vectors $\boldsymbol{\alpha}_i$ are defined

$$\boldsymbol{\alpha}_i = \mathbf{N}^{-1}\mathbf{X}_i'\mathbf{G}_i\mathbf{u} \tag{25}$$

and the linear equations

$$\mathbf{N}\boldsymbol{\alpha}_i = \mathbf{X}_i'\mathbf{G}_i\mathbf{u} \tag{26}$$

are solved for $\boldsymbol{\alpha}_i$ so that Eq. (24) follows with

$$\mathbf{u}'\mathbf{G}_i'\mathbf{X}_i\boldsymbol{\alpha}_i \tag{27}$$

Different vectors $\mathbf{u}$ of independent samples of U give different values for the estimator of the trace so that the trace is obtained by the mean. Golub and von Matt (1997) recommend just one sample vector $\mathbf{u}$ to compute the trace by Eq. (27) and their recommendation is followed here.

This stochastic trace estimator leads to a new method of computing estimates of variance components. This method can be applied even for large systems of normal equations for which the computation of the inverse of the matrix of normal equations is avoided.

If according to Eq. (7) there is only one type of data available, the variance components σ_1^2 and σ_μ^2 have to be computed. We obtain from Eq. (18)

$$\begin{aligned} r_1 &= n - \mathrm{tr}\left(\left(\frac{1}{\sigma_1^2}\mathbf{X}_1'\mathbf{P}_1\mathbf{X}_1 + \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu - \frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\right)\mathbf{N}^{-1}\right) \\ &= n - u + \mathrm{tr}\left(\frac{1}{\sigma_\mu^2}\mathbf{P}_\mu\mathbf{N}^{-1}\right) \end{aligned} \tag{28}$$

and from Eq. (21)

$$r_1 = n - u + \mathrm{tr}\left(\frac{1}{\sigma_\mu^2}\mathbf{G}_\mu'\mathbf{N}^{-1}\mathbf{G}_\mu\right)$$

$$r_\mu = u - \mathrm{tr}\left(\frac{1}{\sigma_\mu^2}\mathbf{G}_\mu'\mathbf{N}^{-1}\mathbf{G}_\mu\right) \tag{29}$$

Only one trace needs be computed in this case and we have $r_1 + r_\mu = n$ in agreement with Eq. (23). In the case of one data set the regularization parameter λ is obtained with Eq. (28) by $\lambda = \sigma_1^2/\sigma_\mu^2$ and with Eqs. (12), (13), (14), (16) and (29) by

$$\lambda = \frac{\left(u - \mathrm{tr}\left(\frac{\sigma_1^2}{\sigma_\mu^2}\mathbf{G}_\mu'\bar{\mathbf{N}}^{-1}\mathbf{G}_\mu\right)\right)(\mathbf{X}_1\hat{\boldsymbol{\beta}} - \mathbf{y}_1)'\mathbf{P}_1(\mathbf{X}_1\hat{\boldsymbol{\beta}} - \mathbf{y}_1)}{\left(n - u + \mathrm{tr}\left(\frac{\sigma_1^2}{\sigma_\mu^2}\mathbf{G}_\mu'\bar{\mathbf{N}}^{-1}\mathbf{G}_\mu\right)\right)(\hat{\boldsymbol{\beta}} - \boldsymbol{\mu})'\mathbf{P}_\mu(\hat{\boldsymbol{\beta}} - \boldsymbol{\mu})} \tag{30}$$

with

$$\bar{\mathbf{N}} = \mathbf{X}_1'\mathbf{P}_1\mathbf{X}_1 + \frac{\sigma_1^2}{\sigma_\mu^2}\mathbf{P}_\mu \tag{31}$$

In comparison to the result of Eq. (30) for the regularization parameter λ from an estimation of variance components, the result for λ will now be given as determined by a cross-validation. It is obtained such that the function $V(\lambda)$ is minimized (Wahba 1977)

$$V(\lambda) = \frac{\frac{1}{n}(\mathbf{X}_1\hat{\boldsymbol{\beta}} - \mathbf{y}_1)'\mathbf{P}_1(\mathbf{X}_1\hat{\boldsymbol{\beta}} - \mathbf{y}_1)}{\left[1 - \frac{1}{n}\mathrm{tr}(\mathbf{G}_1'\mathbf{X}_1\bar{\mathbf{N}}^{-1}\mathbf{X}_1'\mathbf{G}_1)\right]^2} \tag{32}$$

With Eq. (28) we obtain

$$V(\lambda) = \frac{n(\mathbf{X}_1\hat{\boldsymbol{\beta}} - \mathbf{y}_1)'\mathbf{P}_1(\mathbf{X}_1\hat{\boldsymbol{\beta}} - \mathbf{y}_1)}{\left[n - u + \mathrm{tr}\left(\frac{\sigma_1^2}{\sigma_\mu^2}\mathbf{G}_\mu'\bar{\mathbf{N}}^{-1}\mathbf{G}_\mu\right)\right]^2} \tag{33}$$

and finally with Eqs. (12), (14) and (29)

$$V(\lambda) = \frac{n}{r_1}\hat{\sigma}_1^2 \tag{34}$$

When computing the regularization parameter λ from Eq. (30), the estimates $\hat{\sigma}_1^2$ and $\hat{\sigma}_\mu^2$ of the variance components are separately iterated by means of the ratios involving the quadratic form $(\mathbf{X}_1\hat{\boldsymbol{\beta}} - \mathbf{y}_1)'\mathbf{P}_1(\mathbf{X}_1\hat{\boldsymbol{\beta}} - \mathbf{y}_1)$ of the residuals of the data and the quadratic form $(\hat{\boldsymbol{\beta}} - \boldsymbol{\mu})'\mathbf{P}_\mu(\hat{\boldsymbol{\beta}} - \boldsymbol{\mu})$ of the residuals of the prior information. In a cross-validation the regularization parameter λ follows according to Eqs. (32)–(34) by minimizing $V(\lambda)$, which contains (besides λ) the quadratic form $(\mathbf{X}_1\hat{\boldsymbol{\beta}} - \mathbf{y}_1)'\mathbf{P}_1(\mathbf{X}_1\hat{\boldsymbol{\beta}} - \mathbf{y}_1)$ of the residuals of the data.

4 Distributions for the variance components and their ratios

In the case of non-informative priors the variance components σ_i^2 are independently distributed like the inverted gamma distribution with the posterior density function (Ou and Koch 1994; Koch 2000, p. 147)

$$p(\sigma_i^2|\mathbf{y}_i) = \left(\frac{1}{2}\hat{\mathbf{e}}_i'\mathbf{P}_i\hat{\mathbf{e}}_i\right)^{\frac{r_i}{2}}\Gamma\left(\frac{r_i}{2}\right)^{-1}\left(\frac{1}{\sigma_i^2}\right)^{\frac{r_i}{2}+1} \times \exp\left(-\frac{1}{2\sigma_i^2}\hat{\mathbf{e}}_i'\mathbf{P}_i\hat{\mathbf{e}}_i\right) \tag{35}$$

In order to obtain the posterior density function for the regularization parameter λ in Eq. (8) or for the parameter ω of the relative weighting in Eq. (10), the density function for the ratio v

$$v = \sigma_i^2/\sigma_j^2 \tag{36}$$

of the variance components σ_i^2 and σ_j^2 is required.

Since σ_i^2 and σ_j^2 are independently distributed, their joint posterior density function follows from Eq. (35) by

$$p(\sigma_i^2, \sigma_j^2|\mathbf{y}_i, \mathbf{y}_j) \propto \left(\frac{1}{\sigma_i^2}\right)^{\frac{r_i}{2}+1}\exp\left(-\frac{1}{2\sigma_i^2}\hat{\mathbf{e}}_i'\mathbf{P}_i\hat{\mathbf{e}}_i\right) \left(\frac{1}{\sigma_j^2}\right)^{\frac{r_j}{2}+1}\exp\left(-\frac{1}{2\sigma_j^2}\hat{\mathbf{e}}_j'\mathbf{P}_j\hat{\mathbf{e}}_j\right) \tag{37}$$

where the constants are not being considered and $\propto$ denotes proportionality. The transformation $\sigma_i^2 = v\sigma_j^2$ of variables according to Eq. (36) with $\partial\sigma_i^2/\partial v = \sigma_j^2$ leads to

$$p(v, \sigma_j^2|\mathbf{y}_i, \mathbf{y}_j) \propto \left(\frac{1}{v}\right)^{\frac{r_i}{2}+1}\left(\frac{1}{\sigma_j^2}\right)^{\frac{r_i+r_j}{2}+1} \times \exp\left(-\frac{1}{2\sigma_j^2}\left(\frac{1}{v}\hat{\mathbf{e}}_i'\mathbf{P}_i\hat{\mathbf{e}}_i + \hat{\mathbf{e}}_j'\mathbf{P}_j\hat{\mathbf{e}}_j\right)\right) \tag{38}$$

In order to obtain the marginal density function for v, the variance component σ_j^2 is integrated out

$$p(v, \sigma_j^2|\mathbf{y}_i, \mathbf{y}_j) \propto \left(\frac{1}{v}\right)^{\frac{r_i}{2}+1}\left(\frac{1}{2v}\hat{\mathbf{e}}_i'\mathbf{P}_i\hat{\mathbf{e}}_i + \frac{1}{2}\hat{\mathbf{e}}_j'\mathbf{P}_j\hat{\mathbf{e}}_j\right)^{-\frac{r_i+r_j}{2}} \times \Gamma\left(\frac{r_i+r_j}{2}\right)\left(\frac{1}{2v}\hat{\mathbf{e}}_i'\mathbf{P}_i\hat{\mathbf{e}}_i + \frac{1}{2}\hat{\mathbf{e}}_j'\mathbf{P}_j\hat{\mathbf{e}}_j\right)^{\frac{r_i+r_j}{2}}\Gamma\left(\frac{r_i+r_j}{2}\right)^{-1} \times \int_0^\infty \left(\frac{1}{\sigma_j^2}\right)^{\frac{r_i+r_j}{2}+1}\exp\left(-\frac{1}{2\sigma_j^2}\left(\frac{1}{v}\hat{\mathbf{e}}_i'\mathbf{P}_i\hat{\mathbf{e}}_i + \hat{\mathbf{e}}_j'\mathbf{P}_j\hat{\mathbf{e}}_j\right)\right)d\sigma_j^2 \tag{39}$$

As a comparison with Eq. (35) shows, integrating over σ_j^2 means integrating a special inverted gamma distribution so that we finally obtain the posterior density function for the ratio v of variance components

$$p(v|\mathbf{y}_i, \mathbf{y}_j) \propto \left(\frac{1}{v}\right)^{\frac{r_i}{2}+1}\left(\frac{1}{2v}\hat{\mathbf{e}}_i\mathbf{P}_i\hat{\mathbf{e}}_i + \frac{1}{2}\hat{\mathbf{e}}_j'\mathbf{P}_j\hat{\mathbf{e}}_j\right)^{-\frac{r_i+r_j}{2}} \tag{40}$$

An analytical integration of this density function could not be achieved, so that the normalization constant cannot be given. However, as explained in the following, confidence intervals for the ratio v of variance components or hypotheses tests for v can be derived from Eq. (40) by numerical methods.

In Sect. 5 confidence intervals will be computed for the variance components σ_1^2 and σ_2^2 of two types of data and for the variance component σ_μ^2 of the prior information. In addition, confidence intervals are computed for the regularization parameter λ from Eqs. (8) and (36) by

$$\lambda = v = \frac{\sigma_1^2}{\sigma_\mu^2} \tag{41}$$

and for the parameter ω of the relative weighting from Eq. (10) and Eq. (36) by

$$\omega = v = \frac{\sigma_1^2}{\sigma_2^2} \tag{42}$$

The examples give large quadratic forms $\hat{\mathbf{e}}_i'\mathbf{P}_i\hat{\mathbf{e}}_i$ of the residuals and very large partial redundancies r_i. In such cases the posterior density functions of Eqs. (35) and (40) are not well suited for numerical computations due to floating-point underflows or overflows. Random values for the standard gamma distribution have therefore been generated by the log-logistic method recommended for large parameters (Dagpunar, p. 110), which in our case are due to the large partial redundancies. The generated random values have been transformed to random numbers of the gamma distribution and then to those of the inverted gamma distribution so that random values for the variance components σ_i^2 follow with the density function of Eq. (35). By independently generating random values for σ_i^2 and σ_j^2, random values for the ratio v from Eq. (36) or for λ from Eq. (41) and for ω from Eq. (42) with the density function of Eq. (40) are then obtained. Discrete density values or probabilities for σ_i^2 and v follow by counting the generated random values within small intervals; see, for instance, Koch (2000, p. 208).

The $(1-\alpha)$ confidence region, which has the property of being an HPD (highest posterior density) region, then results by adding the probabilities on both sides of the graph of the distribution until the probability α is reached.

This method of establishing the confidence intervals has been checked with smaller values for the quadratic forms of the residuals and the partial redundancies for the variance components σ_i^2 by numerically integrating Eq. (35) and by the tables of Ou (1991), and for the regularization parameter λ and the parameter ω for the relative weighting by numerically integrating Eq. (40).

5 Example

Variance component estimation for geopotential regularization and data weighting as well as the performance of the stochastic trace estimation have been investigated using simulated data. Our examples deal with a regional gravity field recovery from satellite data of different origins.

The first data set was generated following the baseline of the GOCE mission (ESA 1999): 52.000 gradiometer measurements collected along the orbit of a satellite at 250 km altitude during a period of 30 days. For simplification, we used only the V_{zz} Earth-pointing component of the four components which are measured on-board. These observations have been simulated at a sampling rate of 0.2 Hz and corrupted by a colored noise. The colored noise is generated by a power spectral density model with a flat part of 3 $\mathrm{mE}/\sqrt{\mathrm{Hz}}$ between 0.005 and 0.1 Hz and a $1/f$ behaviour between 3.7×10^{-4} and 0.005 Hz (ESA 1999). The regional recovery area was chosen to be from 57 to 132° longitude and -24 to 43° latitude and the anomalous potential was modeled by 5025 $1 \times 1°$ mean gravity anomalies. The vector containing these gravity anomalies is the vector $\boldsymbol{\beta}$ of unknown parameters in our example. Orbit and gradiometry data were derived from the EGM96 gravity model complete to degree and order 300, up to which the harmonic coefficients might be determined by GOCE. The effect of a reference gravity model, represented by the OSU91 model complete to degree and order 36, has been reduced from the data as a preprocessing step. Thus the estimated gravity anomalies must be considered as residuals refering to the reference field. The methology is presented in Kusche and Ilk (2000). The pseudo-true gravity anomalies are shown in Fig. 1.

Our second data set follows the GRACE mission baseline (JPL 1999): two satellites orbiting the Earth at 400-km altitude equipped with a high-precision distance-measuring system. Intersatellite range rates have been simulated at 0.2 Hz sampling rate and corrupted by Gaussian white noise of variance $(1\,\mu\mathrm{m/s})^2$, again for a period of 30 days. The methology for the regional gravity recovery is described in Ilk et al. (1995). In a preprocessing step 10 280 coefficients of the expansion of the variations of the intersatellite distances into a series have been computed as pseudo-observations

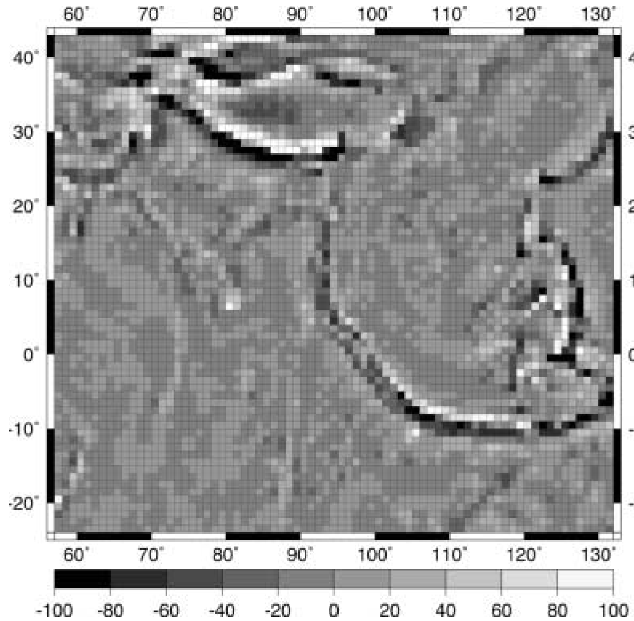

Fig. 1. Pseudo-true gravity anomalies (mGal)

(Ilk et al. 1995) and the effect of the reference model has been removed.

Our first numerical experiment concerns only the first data set and therefore the application of variance component estimation for the regularization of the gradiometric geopotential determination. The normal equations follow from Eq. (9) where $\mathbf{X}_1$ is the gradiometric design matrix, $\mathbf{y}_1$ the vector of gradiometer data and $\mathbf{P}_1$ the weight matrix according to the coloured noise model. $\mathbf{P}_1$ is block-diagonal since we assume that all correlations damp out between two consecutive orbital revolutions. The matrix $\mathbf{P}_1$ needs to be multiplied by the factor 1×10^{-24} which accounts for the conversion of mE units to $1/\mathrm{sec}^2$. For an easier comparison of the estimate of the regularization parameter λ from variance components with the result of the cross-validation we set $\mathbf{P}_\mu = \mathbf{I}$. Since the vector $\boldsymbol{\beta}$ of unknown parameters contains the gravity anomalies refering to a reference field up to degree and order 36, the vector $\boldsymbol{\mu}$ of prior information is set equal to zero.

Table 1 presents the results of our regularization method by means of estimating variance components and the regularized gravity field is shown in Fig. 2. It already closely resembles the pseudo-true field shown in Fig. 1. In addition to the estimates the 95% confidence intervals for the regularization parameter λ, the variance component σ_1^2 for the gradiometry and the variance component σ_μ^2 for the prior information are also shown in Table 1. Judging from the confidence intervals these

Table 1. Results for satellite gradiometry (example 1)

	Estimate	95% confidence interval
λ	0.713×10^{-13}	$0.662 \times 10^{-13} < \lambda < 0.752 \times 10^{-13}$
σ_1^2	0.118×10^{-19}	$0.117 \times 10^{-19} < \sigma_1^2 < 0.120 \times 10^{-19}$
σ_μ^2	0.166×10^{-6}	$0.156 \times 10^{-6} < \sigma_\mu^2 < 0.176 \times 10^{-6}$

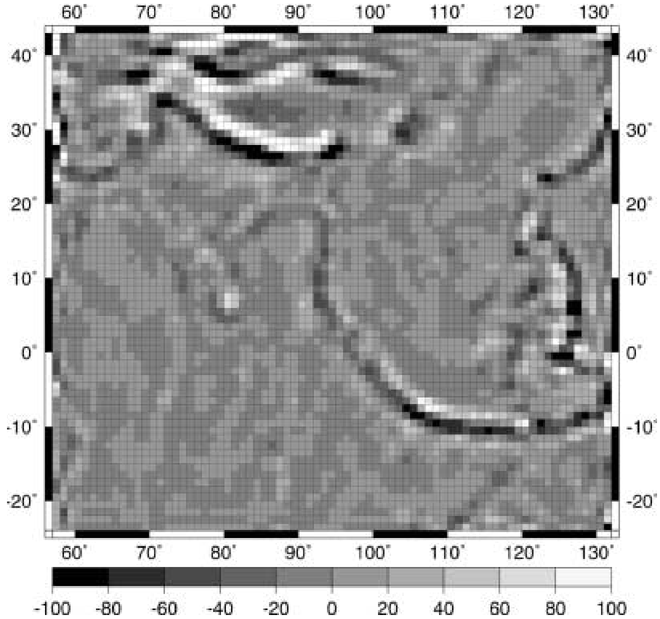

Fig. 2. Solution from satellite gradiometry (example 1)

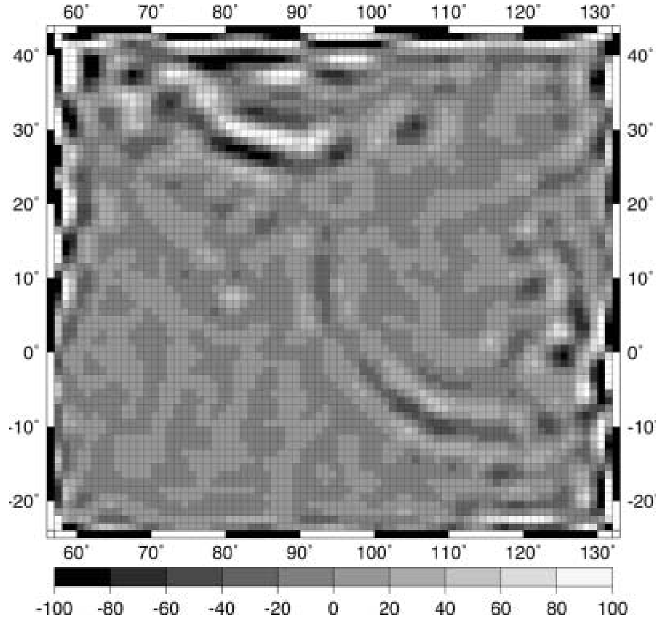

Fig. 3. Solution from satellite-to-satellite tracking (example 2)

quantities are quite well determined. Convergence is already obtained already after five iterations. The approximate values of the variance components for the beginning of the iterations were chosen such that the regularization parameter λ was large enough to avoid ill-conditioned normal equations. At the end of the iterations the regularization parameter $\lambda = 0.71 \times 10^{-13}$ in Table 1 gives a good solution. The corresponding cross-validation regularization parameter $\lambda = 1.37 \times 10^{-13}$ from Eq. (32) is slightly worse than the parameter from variance-component estimation, since the RMS difference between the pseudo-true gravity field and the regularized field is a bit larger than the one for the regularization parameter of Table 1. Without the modeling errors mentioned in connection with Eq. (2) we should obtain the estimate of σ_1^2 to be 1×10^{-24}, the conversion factor of the units as explained above. In practice the estimate of $\hat{\sigma}_1^2 = 0.118 \times 10^{-19}$ has been computed, which is larger by a factor of 11 800 and indicates severe model errors such as edge effects and unmodeled parts of the signal. They are present in our simulation study and will be also present in the case of real data.

In a second experiment only the second data set was used to form the normal equations of Eq. (9). Here $\mathbf{X}_1$ is the design matrix for the vector $\mathbf{y}_1$ of satellite-to-satellite tracking data and $\mathbf{P}_1$ a diagonal weight matrix as explained above. Table 2 presents the results of our variance component estimation and the regularized solution is shown in Fig. 3. From the principle characteristics of the GRACE mission it is clear that we cannot expect the same spatial resolution as from a gradiometric mission, but GRACE will provide the long-wavelength part of the gravity field with superior quality. This explains the fact that the estimate of σ_1^2 without the modeling errors should be 1, but in practice $\hat{\sigma}_1^2 = 0.164 \times 10^3$ was obtained, which is larger by a factor of 164. As compared to the factor of 11 800 for the first example, this points to smaller modeling errors. One reason could be the long-wavelength solution for the second example which is better determined than for the first example. Another reason could be the discretization errors due to the $1 \times 1°$ grid size for the gravity anomalies which effect the solution for the first example with high spatial resolution more strongly than the solution for the second example with low spatial resolution. Looking at Fig. 3, it becomes obvious that the solution for $\boldsymbol{\beta}$ is quite satisfactory; our method therefore gives a reasonable regularization parameter.

Finally, in a third experiment both data sets are combined in the normal Eqs. (11) with λ and ω being determined from the variance components according to Eq. (10). Approximate values for λ and ω are obtained from the variance components estimated in the first two examples. The outcome of this experiment can be seen in Table 3 and the respective solution is plotted in Fig. 4.

Table 2. Results for satellite-to-satellite tracking (example 2)

	Estimate	95% confidence interval
λ	0.196×10^9	$0.173 \times 10^9 < \lambda < 0.214 \times 10^9$
σ_1^2	0.164×10^3	$0.158 \times 10^3 < \sigma_1^2 < 0.168 \times 10^3$
σ_μ^2	0.834×10^{-6}	$0.762 \times 10^{-6} < \sigma_\mu^2 < 0.923 \times 10^{-6}$

Table 3. Results for the combined solution (example 3)

	Estimate	95% confidence interval
λ	0.695×10^{-13}	$0.652 \times 10^{-13} < \lambda < 0.736 \times 10^{-13}$
ω	0.585×10^{-22}	$0.568 \times 10^{-22} < \omega < 0.604 \times 10^{-22}$
σ_1^2	0.122×10^{-19}	$0.121 \times 10^{-19} < \sigma_1^2 < 0.124 \times 10^{-19}$
σ_2^2	0.209×10^3	$0.203 \times 10^3 < \sigma_2^2 < 0.215 \times 10^3$
σ_μ^2	0.176×10^{-6}	$0.165 \times 10^{-6} < \sigma_\mu^2 < 0.186 \times 10^{-6}$

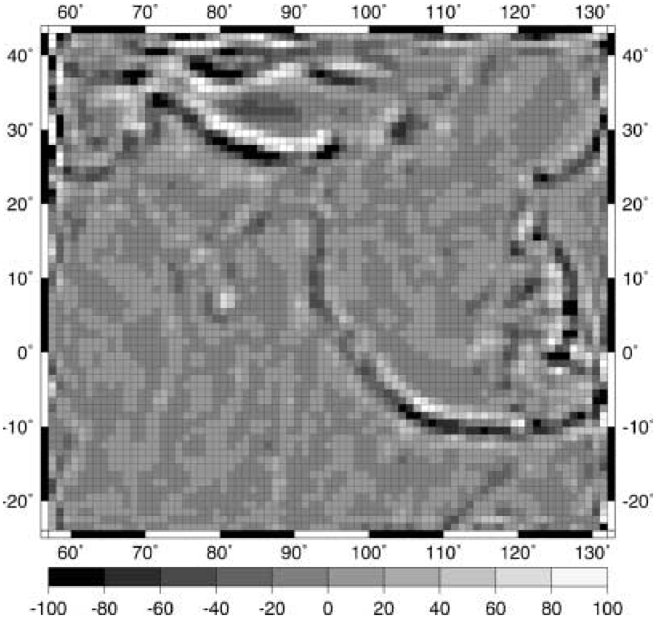

Fig. 4. Combined solution (example 3)

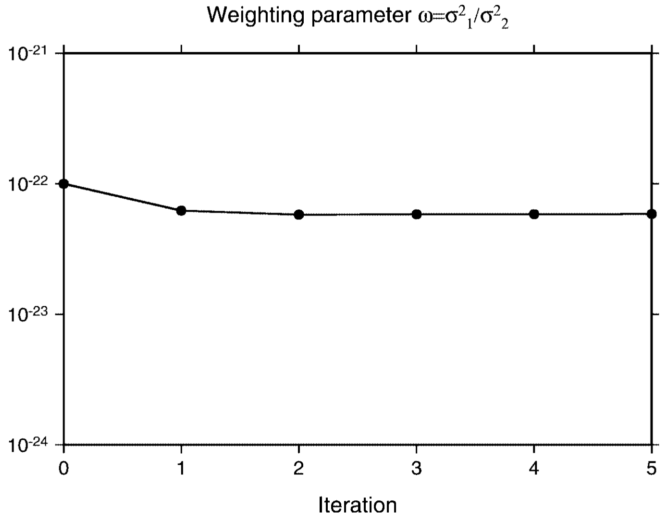

Fig. 6. Change of the weighting parameter ω during the iteration process (example 3)

Again, we needed only five iterations. Figure 5 shows the change of the regularization parameter during the iterations and Fig. 6 the change of the weighting parameter. The stochastic trace estimator and the approximate Cholesky factorization of the block-diagonal weight matrix of gradiometer measurements according to Eq. (22) work quite well because for the $n = 52000 + 10280 = 62280$ observations the correction from Eq. (23) of the partial redundancies amounts in the last iteration to 423.3, or to a relative correction of 0.0068. For the combined gravity field solution the gradiometry data obviously plays a dominating role, which is not surprising in this type of regional-scale simulation. However, adding the satellite-to-satellite tracking data gives a slightly smaller regularization parameter than using only gradiometry data. From a numerical point of view, this means that we improve the stability of the normal equations by combining the two data sets.

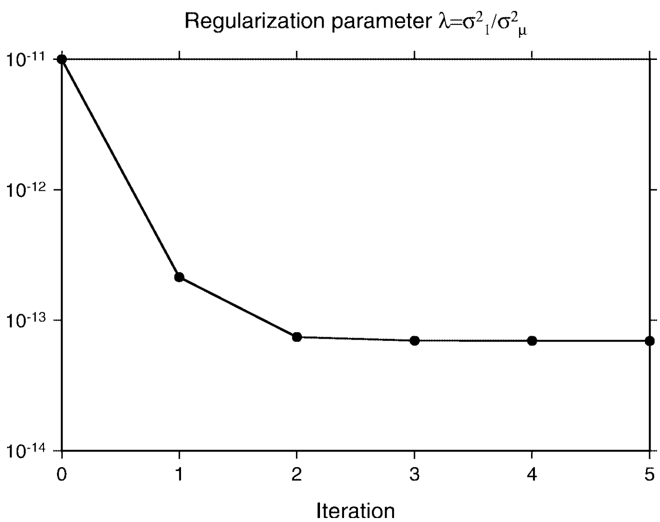

Fig. 5. Change of the regularization parameter λ during the iteration process (example 3)

The variance component σ_μ^2 of the prior information is estimated for the first and third example with $\hat{\sigma}_\mu^2 = 0.17 \times 10^{-6} = (0.41 \times 10^{-3})^2$, thus about $(41\ \mathrm{mGal})^2$. Because $\mathbf{P}_\mu = \mathbf{I}$, as mentioned above, $\hat{\sigma}_\mu^2$ estimates the variance with which the prior information is introduced for the regularization. The estimate $\hat{\sigma}_\mu^2 = (41\ \mathrm{mGal})^2$ is rather small in comparison to the variance of $(45\ \mathrm{mGal})^2$ of the pseudo-true gravity anomalies of Fig. 1. For the pure satellite-to-satellite tracking the estimate $\hat{\sigma}_\mu^2$ is about five times larger. Looking at Fig. 3, we observe that this is not due to a general weakness in the estimated gravity field solution but rather to edge effects, i.e. discrepancies at the borders of the area of investigation, which are well known in regional recovery techniques.

The three experiments demonstrate convincingly that our method of estimating variance components works well for the regularization of gravity field determinations and the combination of different kinds of satellite data. The confidence intervals computed for the unknown variance components and the unknown regularization parameters indicate that these parameters are well determined. The results show modeling errors which are present in this simulation study, but which will be present also in the analysis of real data.

Acknowledgments. The authors are indebted to Dr P. Schwintzer and two anonymous reviewers for their helpful comments.

References

Arsenin VY, Krianev AV (1992) Generalized maximum likelihood method and its application for solving ill-posed problems. In: Tikhonov A (ed) Ill-posed problems in natural sciences. TVP Science Publishers, Moscow, pp 3–12

Blinken R, Koch KR (2001) Geoid and sea surface topography derived from ERS-1 altimeter data by the adjoint method. Stud Geophys Geod 45: 235–250
Bouman J (2000) Quality assessment of satellite-based global gravity field models. Publications in Geodesy 48. Netherlands Geodetic Commission, Delft
Craven P, Wahba G (1979) Smoothing noisy data with spline functions. Numer Math 31: 377–403
Crocetto N, Gatti M, Russo P (2000) Simplified formulae for the BIQUE estimation of variance components in disjunctive observation groups. J Geod 74: 447–457
Dagpunar J (1988) Principles of random variate generation. Clarendon Press, Oxford
European Space Agency (1999) Gravity field and steady-state ocean circulation mission. ESA Publications Division, Reports for Mission Selection, ESA SP-1233(1), ESTEC, Noordwijk
Förstner W (1979) Ein Verfahren zur Schätzung von Varianz-Und Kovarianzkomponenten. Allg Vermess–Nachr 86: 446–453
GeoForschungsZentrum (2000) CHAMP – Der Blick in das Innere der Erde. GeoForschungsZentrum, Potsdam
Girard DA (1989) A fast 'Monte-Carlo cross-validation' procedure for large least squares problems with noisy data. Numer Math 56: 1–23
Golub GH, von Matt U (1997) Generalized cross-validation for large-scale problems. J Comput Graph Stat 6: 1–34
Golub GH, Heath M, Wahba G (1979) Generalized cross-validation as a method for choosing a good ridge parameter. Technometrics 21: 215–223
Grafarend E, d'Hone A (1978) Gewichtsschätzung in geodätischen Netzen, Reihe A, 88. Deutsche Geodätische Kommission, München
Helmert FR (1924) Die Ausgleichungsrechnung nach der Methode der kleinsten Quadrate, 3. Auflage. Teubner, Leipzig
Hutchinson MF (1990) A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines. Commun Stat- Sim 19: 433–450
Ilk KH, Rummel R, Thalhammer M (1995) Refined method for the regional recovery from GPS/SST and SGG. In: Study of the Gravity Field Determination using Gradiometry and GPS (Phase 2). CIGAR III/2 final report, ESA contract No. 10713/93/F/FL
Jet Propulsion Laboratory (1999) GRACE science and mission requirements document. 327–200, Rev. B, Jet Propulsion Laboratory, Pasadena, CA
Kaula WM (1966) Theory of satellite geodesy. Blaisdell, Waltham, MA
Koch KR (1987) Bayesian inference for variance components. Manuscr Geod 12: 309–313
Koch KR (1990) Bayesian inference with geodetic applications. Springer, Berlin Heidelberg New York
Koch KR (1999) Parameter estimation and hypothesis testing in linear models, 2nd edn. Springer, Berlin Heidelberg New York
Koch KR (2000) Einführung in die Bayes-Statistik. Springer, Berlin Heidelberg New York
Kusche J (2001) Implementation of multigrid solvers for satellite gravity anomaly recovery. J Geod 74: 773–782
Kusche J, Ilk KH (2000) The polar gap problem. In: Sünkel H (ed) From Eötvös to milligal. ESA/ESTEC Contract No. 13392/98/NL/GD, pp 177–206
Kusche J, Klees R (2001) On the regularization problem in gradiometric data analysis from GOCE. In: Drinkwater MR (ed) International GOCE User Workshop. ESA Publications Division, ESTEC, Noordwijk, pp 75–80
Lerch FL (1991) Optimum data weighting and error calibration for estimation of gravitational parameters. Bull Géod 65: 44–52
Marsh JG, Lerch FJ, Putney BH, Christodoulidis DC, Smith DE, Felsentreger TL, Sanchez BV, Klosko SM, Pavlis EC, Martin TV, Robbins JW, Williamson RG, Colombo OL Rowlands DD, Eddy WF, Chandler NL, Rachlin KE, Patel GB, Bhati A, Chinn DS (1988) A new gravitational model for the earth from satellite tracking data: GEM-T1. J Geophys Res 93: 6169–6215
O'Sullivan F (1986) A statistical perspective on ill-posed inverse problems. Statist Sci 1: 502–527
Ou Z (1991) Approximate Bayes estimation for variance components. Manuscr Geod 16: 168–172
Ou Z, Koch KR (1994) Analytical expressions for Bayes estimates of variance components. Manuscr Geod 19: 284–293
Rao CR (1973) Linear statistical inference and its applications. Wiley, New York
Reigber Ch, Ilk KH (1976) Vergleich von Resonanzparameterbestimmungen mittels Ausgleichung und Kollokation. Z Vermess 101: 59–67
Schwintzer P, Reigber Ch, Bode A, Kang Z, Zhu SY, Massmann FH, Raimondo JC, Biancale R, Balmino G, Lemoine JM, Moynot B, Marty JC, Barlier F, Boudon Y (1997) Long-wavelength global gravity field models: GRIM4-S4, GRIM4-C4. J Geod 71: 189–208
Tikhonov AN, Arsenin VY (1977) Solutions of Ill-posed problems. Wiley, New York
Vinod HD, Ullah A (1981) Recent advances in regression methods. Dekker, New York
Wahba G (1977) Practical approximate solutions to linear operator equations when the data are noisy. SIAM J Numer Anal 14: 651–667
Xu P (1998) Truncated SVD methods for discrete linear ill-posed problems. Geophys J Int 135: 505–514
Xu P, Rummel R (1994a) Generalized ridge regression with applications in determination of potential fields. Manuscr Geod 20: 8–20
Xu P, Rummel R (1994b) A simulation study of smoothness methods in recovery of regional gravity fields. Geophys J Int 117: 472–486