Fast Variance Component Estimation in GOCE Data Processing

J.M. Brockmann and W.-D. Schuh
Institute of Geodesy and Geoinformation, Department for Theoretical Geodesy, University of Bonn, Germany

Abstract

This presentation shows the convergence behavior of variance component estimation (VCE) within the iterative solution strategy for GOCE data processing. The optimized pcgma (Preconditioned Conjugate Gradient Multiple Adjustment) algorithm was designed as a tailored iterative solution strategy for the determination of the Earth's gravity field in terms of a spherical harmonic analysis. Within GOCE-HPF (GOCE High Level Processing Facility) the pcgma algorithm works with the purpose of a tuning machine in that it is used to optimize the filter design and to determine optimal variance factors for data combination (satellite-to-satellite tracking, satellite-gravity gradiometry) as well as for additional prior information about the general behavior of the gravity field, especially in polar regions. We turn special attention to the re-estimation of the variance components after only few solution steps, where beside the squared sum of the residuals also the redundancy for each group of data is of special interest to determine optimal variance components (VC) as fast as possible. Simulated GOCE data sets are used to build up a realistic scenario and to demonstrate the potential of this nested strategy.

Introduction

Starting from the linear normal equation system in terms of a least squares adjustment, the system for GOCE data processing reads as

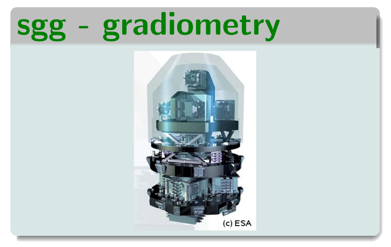
$$\underbrace{\left(\omega_{\text{sgg}}\mathbf{A}_{\text{sgg}}^T\mathbf{A}_{\text{sgg}}+\omega_{\text{sst}}\mathbf{N}_{\text{sst}}+\omega_{\text{reg}}\mathbf{P}_{\mu}\right)}_{:=\mathbf{N}}\mathbf{x} = \underbrace{\omega_{\text{sgg}}\mathbf{A}_{\text{sgg}}^T\mathbf{l}_{\text{sgg}}+\omega_{\text{sst}}\mathbf{n}_{\text{sst}}+\omega_{\text{reg}}\boldsymbol{\mu}}_{:=\mathbf{n}}, \qquad (1)$$

where

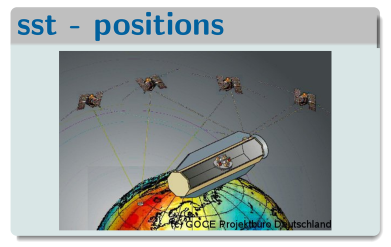
- $\mathbf{A}_{\text{sgg}}, \mathbf{l}_{\text{sgg}}$ . . . decorrelated design matrix and sgg observations,
- $\mathbf{N}_{\text{sst}}, \mathbf{n}_{\text{sst}}$ . . . pre-processed normal equation system of sst observations,
- $\mathbf{P}_{\mu}$ . . . regularization matrix, inverse covariances of prior information,
- $\boldsymbol{\mu}$ . . . prior information on parameter side,
- ω_i . . . unknown weight factors of observation groups,
- $\mathbf{x}$ . . . unknown parameters, spherical harmonic coefficients $\{c_{lm}, s_{lm}\}$.

Containing the three observation types

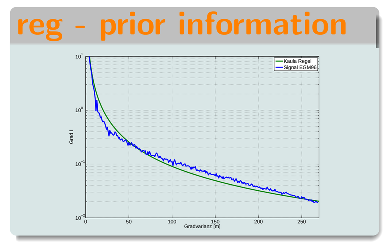
sgg - gradiometry



sst - positions



reg - prior information



There are two different ways of solving the combined system:

1. Compute the joint normal equation matrix $\mathbf{N}$ and solve $\mathbf{N}\mathbf{x} = \mathbf{n}$.
2. Avoid the computation of $\mathbf{N}$ (esp. large scaled product $\mathbf{A}_{\text{sgg}}^T\mathbf{A}_{\text{sgg}}$), solving (1) using a tailored iterative solution strategy (pcgma, Schuh 1996, Boxhammer 2006).

After a first parameter estimation, the weights $\omega_i^{(\iota)}$ could be determined using VCE based upon the partial redundancies and the squared sum of residuals of the observation group i (Koch and Kusche 2001). With new weights of the VCE iteration ι, eq. (1) has to be solved again.

Due to the nesting of the two iterative methods and the fact that the covariance matrix of the parameters is not computed, the algorithms have to be modified. In this presentation we will concentrate upon three parts:

1. Integration of a Monte-Carlo based trace estimator (computation of partial redundancy) in VCE algorithm combined with iterative parameter estimation.
2. Quality requirement for ω_i to get an optimal solution.
3. Convergence of VCE, esp. for trace estimator and squared sum of residuals.

The general equation to determine the variance components σ_i^2 for observation group i (e.g. Koch 2007) can be written as

$$\sigma_i^2 = \frac{\mathbf{v}_i^T\mathbf{P}_i\mathbf{v}_i}{u_i - m_i} = \frac{\Omega_i}{r_i}, \qquad (2)$$

where
- Ω_i . . . weighted squared sum of residuals of the observations,
- u_i . . . number of observations in group i,
- m_i . . . number of parameters determined by group i,
- r_i . . . partial redundancy of observation group i.

The weights of observation group i can then be written as

$$\omega_i = \frac{1}{\sigma_i^2} = \frac{m_i - u_i}{\Omega_i}. \qquad (3)$$

VCE introducing the stochastic trace estimator

Starting from (2) to determine σ_i^2 and replacing the partial redundancy by

$$r_i = u_i - \frac{1}{\sigma_i^2}\mathrm{trace}\left(\mathbf{A}_i^T\mathbf{P}_i\mathbf{A}_i\mathbf{N}^{-1}\right) \qquad \text{group i: observation equations} \qquad (4)$$

$$r_i = u_i - \frac{1}{\sigma_i^2}\mathrm{trace}\left(\mathbf{N}_i\mathbf{N}^{-1}\right) \qquad \text{group i: normal equations} \qquad (5)$$

(e.g. Alkhatib 2007) one can obtain two facts. The procedure

1. is iterative, we need an initial value for σ_i^2 resp. the weights ω_i and
2. needs the combined normal equation matrix, whose computation was avoided during parameter estimation.

If one introduces the stochastic trace estimator presented in Hutchinson (1990) as suggested by Koch and Kusche (2001) w.r.t. the fact that the matrix in the trace has to be symmetric one can reformulate partial redundancy

$$r_i = u_i - E\{\mathcal{P}_i^T\mathbf{R}_i\mathbf{A}_i\mathbf{N}^{-1}\mathbf{A}_i^T\mathbf{R}_i^T\mathcal{P}_i\} \qquad \text{with decomp. } \mathbf{P}_i = \mathbf{R}_i^T\mathbf{R}_i, \qquad (6)$$

$$r_i = u_i - E\{\mathcal{P}_i^T\mathbf{G}_i\mathbf{N}^{-1}\mathbf{G}_i^T\mathcal{P}_i\} \qquad \text{with decomp. } \mathbf{N}_i = \mathbf{G}_i^T\mathbf{G}_i. \qquad (7)$$

In these equations the random vector $\mathcal{P}$ follows a discrete uniform distribution taking the values ±1 with prohability 0.5.

transformation		product		system of eq.		parameter estimation
$\bar{\mathbf{p}}_{ik} = \mathbf{G}_i^T\mathbf{p}_{ik}$, $\bar{\mathbf{p}}_{ik} = \mathbf{A}_i^T\mathbf{R}_i^T\mathbf{p}_{ik}$, $\mathcal{P}_i \Rightarrow \mathbf{p}_{ik},\ k \in 1..K$	$\Rightarrow$	$\mathbf{N}^{-1}\bar{\mathbf{p}}_{ik}$, $\bar{\mathbf{p}}_{ik}$: transformed random vector	$\Rightarrow$	$\mathbf{N}\mathbf{a}_{ik} = \bar{\mathbf{p}}_{ik}$, additional r.h.s.	$\Rightarrow$	$\mathbf{a}_{ik} = \mathrm{cg}(\mathbf{N},\ \bar{\mathbf{p}}_{ik})$ additional parameters $r_i = n_i - \frac{1}{\sigma_i^2}\frac{1}{K}\sum_{k=1}^{K}\bar{\mathbf{p}}_{ik}^T\mathbf{a}_{ik}$

Effect of an error in the trace estimation

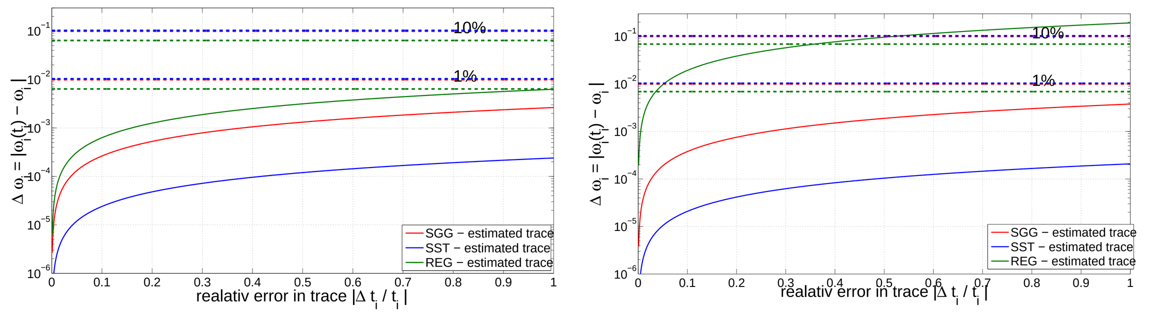
Introducing an error Δu_i in the trace estimation to eq. (3), the erroneous estimation for the weight factor could be written as

$$\tilde{\omega}_i = \frac{n_i - (u_i + \Delta u_i)}{\Omega_i} = \frac{n_i - u_i}{\Omega_i} - \frac{\Delta u_i}{\Omega_i} = \omega_i - \frac{\Delta u_i}{\Omega_i} = \omega_i - \Delta\omega_i. \qquad (8)$$

As relative error one can obtain

$$\frac{\Delta\omega_i}{\omega_i} = \frac{\frac{\Delta u_i}{\Omega_i}}{\frac{n_i - u_i}{\Omega_i}} = \frac{\Delta u_i}{\Omega_i}\frac{\Omega_i}{n_i - u_i} = \frac{\Delta u_i}{n_i - u_i} = \frac{1}{\frac{n_i}{u_i} - 1}\frac{\Delta u_i}{u_i}, \qquad (9)$$

demonstrating a linear effect with the factor $a_i := 1/\left(\frac{n_i}{u_i} - 1\right)$ on the weights.



(a) For model up to d/o 200. (b) For model up to d/o 270.

Figure : Absolute error of weight estimation caused by a relative error in the trace.

The investigations can be summarized with three statements:

1. The effect of an error is highly dependent on degree of the expansion,
2. The effect on SGG and SST is negligible, assuming the trace term to be zero leads in both resolutions to an error in the weights less than 1%.
3. the regularization group is the group for which the quality in trace estimation is important for increasing degrees.

Effect of a non optimal converged squared sum of residuals

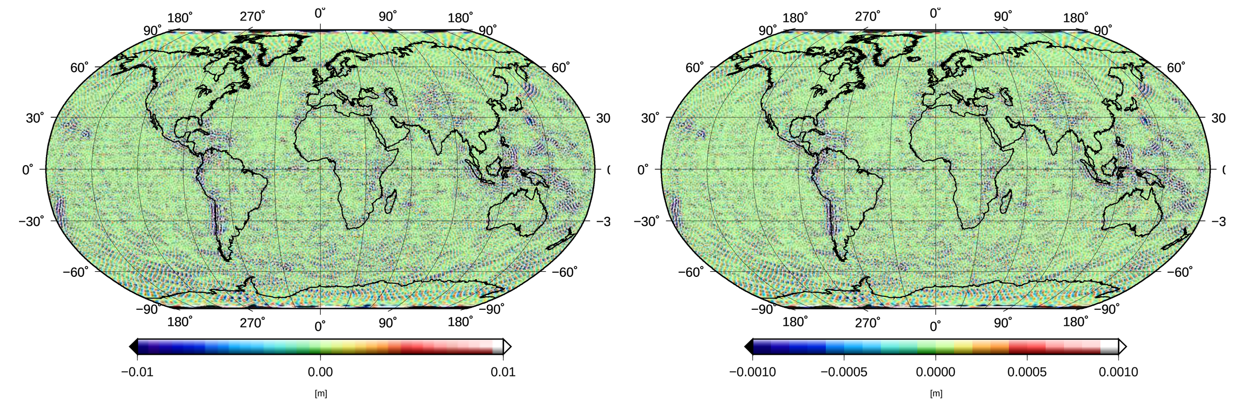
An error $\Delta\Omega_i$ in the squared sum of residuals (e.g. not converged CG) could be introduced as $\tilde{\omega}_i = \frac{n_i - u_i}{\Omega_i + \Delta\Omega_i}$. Linearization according to Taylor leads to a relative error in the weights

$$\frac{\Delta\omega_i}{\omega_i} \approx -1\frac{\Delta\Omega_i}{\Omega_i} + 1\left(\frac{\Delta\Omega_i}{\Omega_i}\right)^2 \mp \mathcal{O}\left(\left(\frac{\Delta\Omega_i}{\Omega_i}\right)^3\right). \qquad (10)$$

Here we see, that a relative error of not converged residuals in the first approximation causes an error scaled by 1 in the weights. In contrast to the trace estimator, the absolute error in $\Delta\Omega_i$ can be larger than Ω_i. Therfore the relative error is not limited. But we have to keep in mind that an update of the weights does not slow down the convergence of the residuals. In addition we can take advantage of the typical behavior of the CG algorithm where the major part of the residuals is minized in the first iteration steps.

Effect of errors in the weights on the solution

The critical point in variance component determination is the weighting factor of the regularization group, whose convergence is clearly the worst. A simulation shows the effect of $\Delta\omega_{\text{reg}} = 0.1$ and $\Delta\omega_{\text{reg}} = 0.01$. This gives us a feel for the error convergency of the weight factors w.r.t. the first resp. the second digit.



(a) $\Delta\omega_{\text{reg}} = 0.1$

Sektor	min	max	mean	rms
±83.5°	−0.075	0.091	0.000	0.009
±90.0°	−0.127	0.112	−0.000	0.012

(b) $\Delta\omega_{\text{reg}} = 0.01$

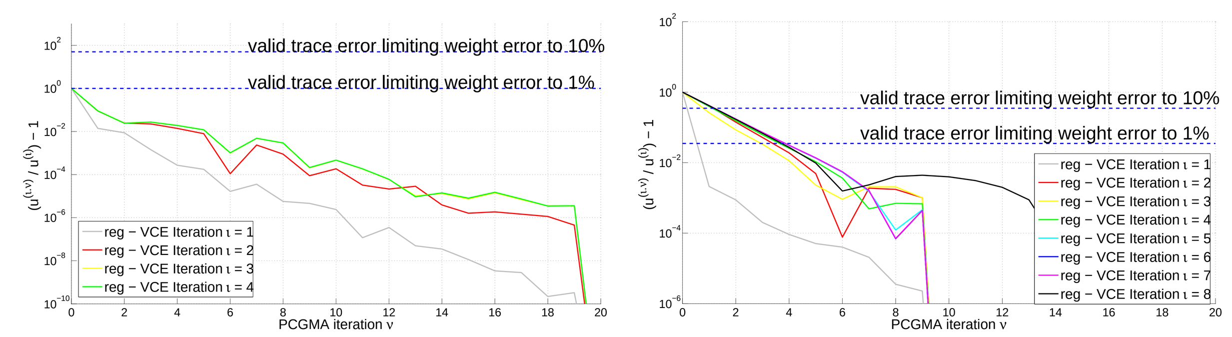
Sektor	min	max	mean	rms
±83.5°	−0.007	0.009	−0.000	0.001
±90.0°	−0.012	0.014	0.000	0.001

Figure : Effect of error in weights on geoid heights [m] as a difference to the optimal solution.

Convergence of the first digits causes maximal errors in the magnitude of 10 *cm*. This is in the magnitude of the mean error that the optimal solution has w.r.t. the "'true"' EGM96 model (closed loop simulation). Convergence of ω_{reg} in the second digit limits the maximal error to a about 10 *mm* w.r.t. optimal solution (one magnitude decreased).

Convergence of the trace estimation

From creating a reference solution in a kind of a "brute force" strategy, where many CG iterations are done during a single VCE iteration, one can assume the last trace estimation in each VCE iteration to be the best one can achieve with the weights $\omega_i^{(\iota)}$.



(a) Trace convergence in 200 model. (b) Trace convergence in 270 model.

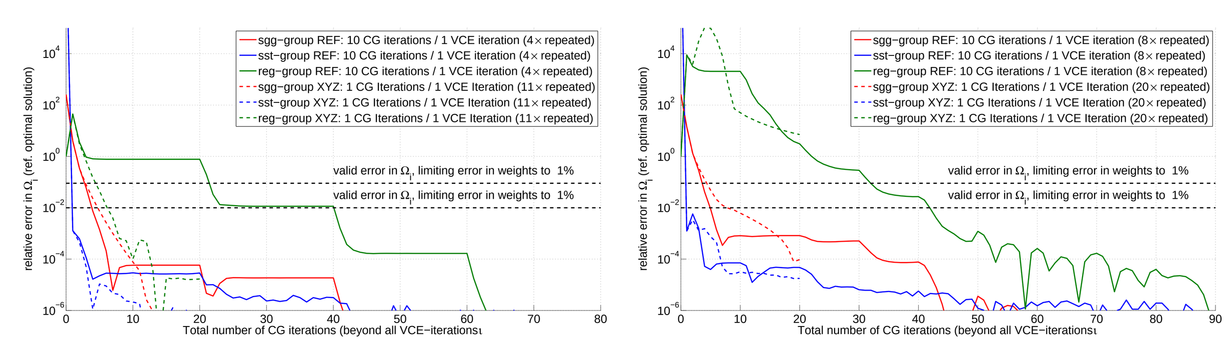
Figure : Relative error in the trace estimation (reference: last estimation of each VCE iteration).

The above figures show the relative errors of the trace estimation during the CG iteration process for every VCE iteration. We selected the trace estimation of prior information, because the special sensitivity of the weight factors was shown above. A comparison for a degree and order (d/o) 200 model and a d/o 270 model reflects the sensitivity w.r.t. the resolution.

Whereas for a d/o 200 model one iteration is sufficient, the d/o 270 model needs about four iterations to reduce the relative weight error under the 1% level.

Convergence of residuals

Assuming Ω_i after the last iteration in the reference solution to be minimal, one can show the remaining relative error over all CG iterations done in the reference solution.



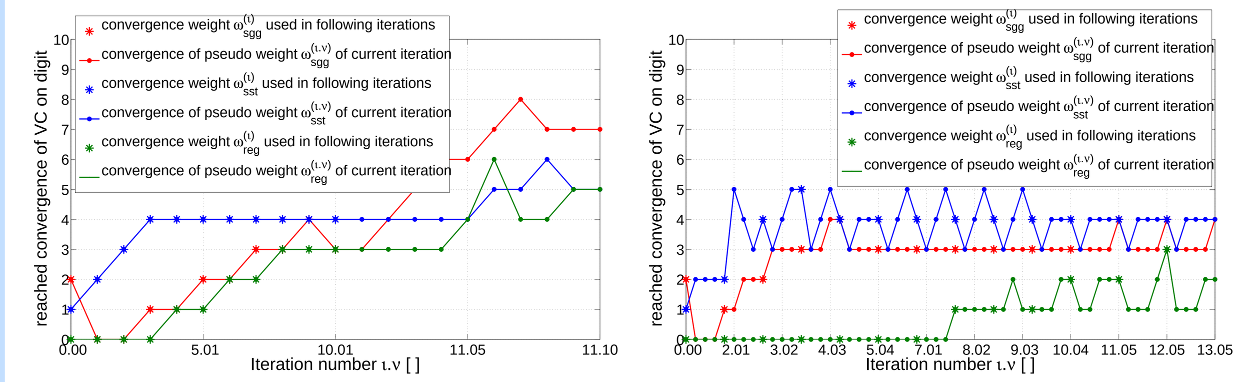
(a) In model up to d/o 200. (b) In model up to d/o 270.

Figure : Relative error in Ω_{reg} referring to minimal value after last CG iteration.

In the left figure (d/o 200 model) the convergence of Ω_i after only a few steps after weight adaption can be seen; whereas in the right figure especially the slow convergence rate for the prior information group can be noticed.

Convergence of variance components in optimized nested approach

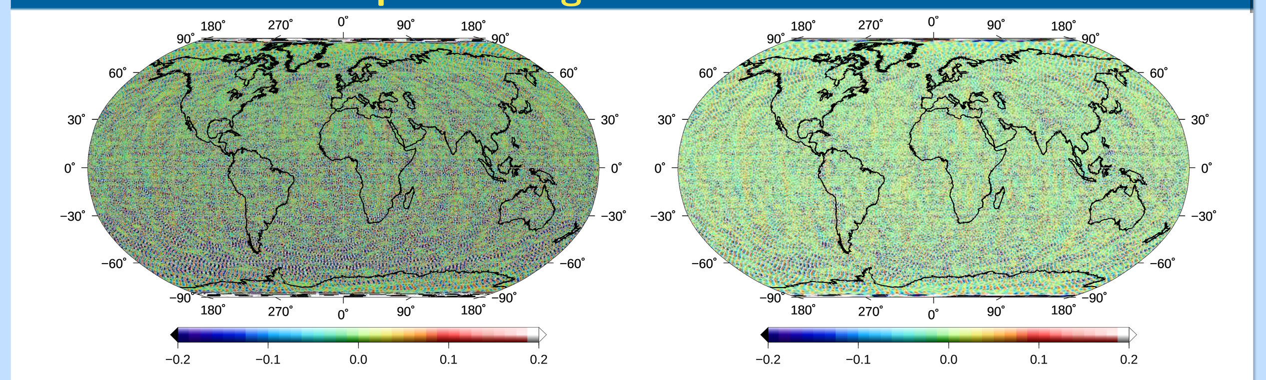
Applying the knowledge about the convergence behavior, one can see that we only need a single CG iteration in each VCE iteration w.r.t. the d/o 200 model. Therefore we can stop the nested procedure after 5–6 VCE iterations to achieve a variance component accuracy up to the second digit. In the d/o 270 model, the convergence behavior is in general worse; we need about 4 CG iterations in each VCE to make sure the trace converges. In total 40 CG iterations are necessary.



(a) d/o 200 (10 VCE iterations, 1 CG iteration each). (b) d/o 270 (11 VCE iterations, 4 CG iteration each).

Figure : Converged digits of VCE (reference: last VCE of optimal solution)

General benefit of optimal weight determination

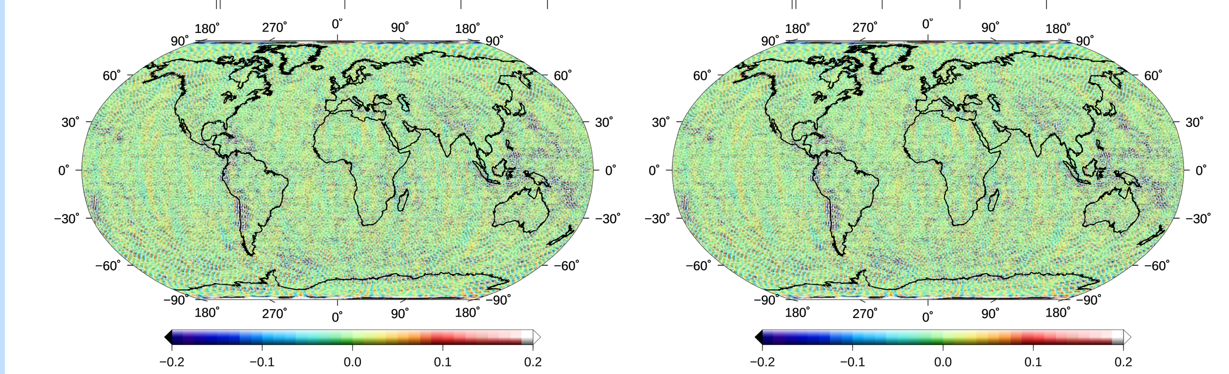


(a) Difference (geoid heights (m)) $\iota = 1$

sector	min	max	mean	rms
±83.5°	-3.190	2.594	0.000	0.282
±90.0°	-2057.030	2098.570	5.207	149.173

(b) Difference (geoid heights (m)) $\iota = 3$

sector	min	max	mean	rms
±83.5°	-0.970	1.004	-0.001	0.153
±90.0°	-3.698	3.571	0.012	0.357



(e) Difference (geoid heights (m)) $\iota = 5$

sector	min	max	mean	rms
±83.5°	-1.237	1.258	-0.001	0.129
±90.0°	-3.047	3.654	0.017	0.338

(f) Difference (geoid heights (m)) $\iota = 8$

sector	min	max	mean	rms
±83.5°	-1.242	1.262	-0.001	0.128
±90.0°	-3.041	3.658	0.017	0.337

Figure : General benefit of optimal weight determination. Optimal solutions with weights determind in different VCE iterations. Closed loop simulation with EGM96 as reference field.

Acknowledgments

Part of this work was financially supported by the BMBF Geotechnologien program GOCE–GRAND II (Grant 03F0421B) and the GOCE High–level Processing Facility (HPF). Parts of the simulations were performed using JUMP at Jülich Supercomputing Center. The computing time was granted by the John von Neumann Institute for Computing (project 1827).

References

Alkhatib, H. (2007) *On Monte Carlo methods with applications to the current satellite gravity missions*. Dissertation. Institut für Geodäsie und Geoinformation der Universität Bonn.

Boxhammer, C. (2006) *Effiziente numerische Verfahren zur sphärischen harmonischen Analyse von Satellitendaten*. Dissertation. Institut für Theoretische Geodäsie der Universität Bonn.

Brockmann, J.M. (2008) *Effiziente Varianzkomponenentenschätzung über iterative Techniken bei der GOCE–Daten Prozessierung*. Diplomarbeit. Institut für Geodäsie und Geoinformation der Universität Bonn.

Hutchinson, M.F. (1990) A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines. *Commun. Stat. Simula.*, 19(2):433–450.

Koch, K. R. and J. Kusche (2001) Regularization of geopotential determination from satellite data by variance components. *Journal of Geodesy*, 76(5):641–652.

Koch, K.R. (2007) *Introduction to Bayesian Statistics*. Springer, Berlin - Heidelberg - New York, 2. edition.

Schuh, W.-D. (1996) *Tailored numerical solution strategies for the global determination of the earth's gravity field*. Number 81 in Mitteilungen der Universität Graz. Universität Graz.

Institut of Geodesy and Geoinformation
Department of Theoretical Geodesy
University of Bonn

Contact:
Nussallee 17
D-53115 Bonn
www.igg.uni-bonn.de/tg/