PERGAMON

Journal of Geodynamics 33 (2002) 65–74

JOURNAL OF
GEODYNAMICS

www.elsevier.com/locate/jgeodyn

Investigations on the polar gap problem in ESA's gravity field and steady-state ocean circulation explorer mission (GOCE)

S. Rudolph*, J. Kusche, K.-H. Ilk

Institute of Theoretical Geodesy, University of Bonn, Nußallee 17, 53115 Bonn, Germany

Abstract

Due to the sun-synchronous GOCE orbit two polar regions with a radius of about 700 km cannot be observed. In this contribution effects of these data gaps are investigated using a regional recovery strategy. It consists essentially of a variant of the time-wise approach, where the geopotential is represented by a system of space-localizing base functions. The ability of the three diagonal tensor components of the gradiometer instrument "to look into" the gaps is investigated. Our investigations show that $1^\circ \times 1^\circ$ mean anomalies can be recovered at reduced accuracy. Referred to a simulated gravity field model complete up to degree 360 the rms-values increase with geographical latitude from 3 mGal up to 10–12 mGal. It is not possible to overcome these problems solely by introducing additional GPS data, but if terrestrial or airborne gravity information is available, also datum defects can be successfully fixed. This work on the polar gap problem was done in the framework of the ESA contract on GOCE entitled "From Eötvös to Milligal".

1. Introduction

GOCE is the first Earth Explorer Mission (EEM) and the only one for which at this moment a sufficient budget has been allocated by ESA. The foreseen launch is either in summer or winter 2004 and GOCE will benefit strongly from the missions CHAMP and GRACE.

The aim of the GOCE mission is to provide global and regional models of the Earth's gravity field and its corresponding geoid surface with high spatial resolution and accuracy. In contrast to the other missions GOCE will be equipped with a gradiometer instrument. Assuming a measurement

* Corresponding author.
E-mail address: stephan@theor.geod.uni-bonn.de (S. Rudolph).

0264-3707/02/$ - see front matter
PII: S0264-3707(01)00055-2

accuracy of 3 mE/$\sqrt{\text{Hz}}$ in the measurement bandwidth of about 0.001–0.1 Hz the maximum resolvable degree is expected to be at or close to 250—corresponding to a resolution around 80 km.

Due to the sun-synchronous orbit with an inclination of 96.5° the two polar regions with a radius of about 700 km cannot be observed. They represent data gaps. On the other hand, the maximum of data is collected at the rim of the gaps. This is shown in Fig. 1, where the groundtracks entailing the gaps are visible. The effects of these data gaps are poorly determined frequency bands in the spectral representation of the gravity field.

GOCE will carry a three dimensional gradiometer, a capacitive type with a baseline of 50 cm. The principle is based on the measurement of the relative accelerations of test masses at different locations inside the satellite. These differential accelerations provide the second derivatives of the potential. Additionally the spacecraft will be equipped with a GPS receiver.

The polar gap problem has been studied in a number of papers, a recent investigation in view of the space-wise appoach is Tscherning et al. (2000), where more references can be found.

In this study a regional recovery strategy is used for the analysis. The geopotential is represented by a system of space-localizing base functions instead of spherical harmonics. Due to the ill-posedness of the downward continuation process a regularization technique has to be applied for the estimation of the coefficients. In Section 2 this mathematical setting is described. Section 3 deals with the simulation strategy. Our simulation investigates the capabilities of a full diagonal gradiometer to recover parameters describing the gravity field within the polar gaps. The presentation of numerical results follows in Section 4. It turns out that there are significant differences in the behaviour of the different tensor components. Additional GPS-data are not able to enhance the quality of the solution, but on the other hand a combination with terrestrial or airborne data of low noise can successfully fix datum defects. In Section 5 a short conclusion is given.

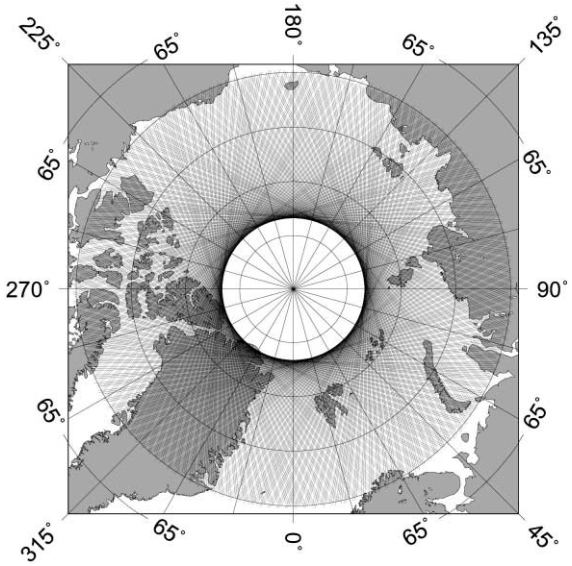

Fig. 1. (North) polar gap.

2. Mathematical modelling

It is common to model the disturbance potential by a spherical harmonic expansion,

$$T = \sum_{n=0}^{\infty} \left(\frac{R}{r}\right)^{n+1} \sum_{m=0}^{2n+1} t_{\mathrm{nm}} Y_{\mathrm{nm}}(\theta, \lambda). \tag{1}$$

$Y_{\mathrm{nm}}(\theta, \lambda)$ are the fully normalized spherical harmonics and the corresponding coefficients t_{nm} are the residuals with respect to a reference potential. Due to the global support of the spherical harmonics, in our investigation of data gaps a different parameterization had to be chosen. A regional recovery strategy—as it is used here—needs a representation with base functions of (nearly) local support. Therefore, the approximation of the gravity field is based on space-localizing base functions $\tilde{T} = \sum_{i=1}^{U} \chi_i \phi_i$, where the coefficients χ_i are the unknowns. Suitable base functions ϕ_i given by isotropic kernels. In this paper the Stokes-kernel $S(P, Q_{\mathrm{i}})$ is exclusively used. Then it is common practice to associate the unknowns with block mean gravity anomalies Δg defined on surface elements $\Delta\omega(Q_i)$ (Heiskanen and Moritz, 1967):

$$\tilde{T}(P) = \frac{R}{4\pi} \sum_{i=1}^{U} S(P, Q_i) \Delta g(Q_i) \Delta\omega(Q_i) \tag{2}$$

There are two types of observation equations used in this study: one for the gradiometer and one for the additional high-low satellite-to-satellite (SST) measurements to the GPS-satellites. The first one simply reads

$$\mathbf{b} = \mathbf{\Gamma}^{\mathrm{obs}} - \mathbf{\Gamma}^{\mathrm{ref}} = \nabla\nabla\tilde{T}. \tag{3}$$

In connection with Eq (2) the ∇-operator has to be applied twice to the Stokes-kernel $S(P, Q_{\mathrm{i}})$. Only diagonal terms $\mathbf{\Gamma}_{xx}, \mathbf{\Gamma}_{yy}$ and $\mathbf{\Gamma}_{zz}$ are used.

The observation equation in the case of SST is based on the eigenfunction expansion of the intersatellite distance (Schneider, 1984). The solution of a boundary value problem leads to an integral equation of Freholm type. A detailed treatment is given in Ilk et al. (1995). In our study the n expansion coefficients

$$b_v = r_v^{\mathrm{obs}} - r_v^{\mathrm{ref}} = -\frac{2T^2}{v^2\pi^2} \int_0^1 (\sin(v\pi\tau')\mathbf{e}_{12} \cdot \nabla\tilde{T}) \mathrm{d}\tau' \tag{4}$$

for $v = 1...n$ are used as pseudo-observations. A matrix representation of the Eqs. (3) and (4) consists of the design matrix $\boldsymbol{A}$, the vector of the unknown coefficients $\boldsymbol{\chi}$, the observation vector $\boldsymbol{l}$ and a vector of random errors $\boldsymbol{\epsilon}$

$$\boldsymbol{A}\chi = \boldsymbol{l} + \boldsymbol{\epsilon}. \tag{5}$$

Due to the ill-posed nature of the downward continuation process, the presence of errors and the data gap the discretized problem (5) is ill-conditioned. Therefore, a regularization technique has to be applied. Using Tikhonov regularization with signal constraint the normals take the well-known form

$$(\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A} + \gamma^2\boldsymbol{M})\boldsymbol{\chi} = \boldsymbol{A}^T\boldsymbol{P}\boldsymbol{l} \tag{6}$$

with the regularization parameter γ^2 and the Gramian matrix $\boldsymbol{M}$. Then the solution minimizes the functional

$$\mathcal{F}_{\gamma^2} = \|\boldsymbol{A}\boldsymbol{\chi} - \boldsymbol{l}\|_{\boldsymbol{P}}^2 + \gamma^2\|\boldsymbol{\chi}\|_{\boldsymbol{M}}^2 \tag{7}$$

and the covariance matrix of the estimated unknowns $\hat{\boldsymbol{\chi}}$ reads

$$\mathrm{Cov}(\hat{\boldsymbol{\chi}}) = \hat{\sigma}(\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A} + \gamma^2\boldsymbol{M})^{-1}\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A} + \gamma^2\boldsymbol{M})^{-1}. \tag{8}$$

3. Simulation strategy

The ability of the regional recovery strategy to deal with the polar gaps is investigated in a simulation study. It starts with the computation of the GOCE orbits and the gradiometer data. The orbit parameters in Table 1 serve as input and the EGM96 model complete up to degree and order 360 was chosen as quasi true gravity field. This means that the orbits and the measurements are computed with this model and the nominal values for the gravity anomalies, too. To have a coordinate grid suitable for anomaly block formation in the north polar region the coordinate system is rotated by 90°. Fig. 2 shows this $1^\circ \times 1^\circ$ grid with approximately equal block sizes and no singularity. The simulation is extended over 31 days, which corresponds to 499 GOCE tracks in this area (see Fig. 1). The instrument is assumed to produce data at 0.2 Hz sampling rate. In connection with typical orbit arc lengths of 560 s an amount of 56,000 observations for each tensor component follows. For the SGG measurements a coloured noise model described in Koop et al. (2000) is taken into account (see Table 2). Besides, surface data could be simulated,

Table 1
GOCE orbit parameters

a	6 628 136.30 m
e	0.001
i	96.6°

Table 2
GOCE instrument parameters

f	0.2 Hz
rms	1.18 mE
T	31 Days

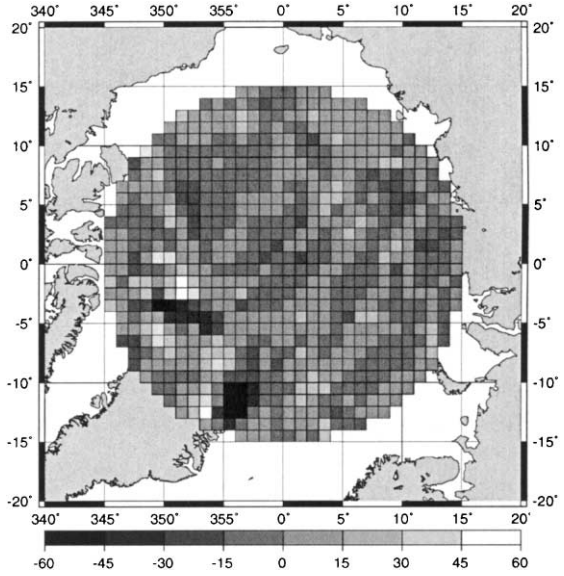

Fig. 2. Rotated coordinate grid (mGal).

too. The output of the simulation study are mean gravity anomalies, their full covariance matrix, the estimated bias, the estimated MSE matrix and true mean square errors.

Of utmost importance in the practical implementation of regularization techniques is the choice of an optimal regularization parameter γ^2. There are two strategies: comparison to the true solution, which is only possible in a simulation, or applying a parameter choice rule. In the latter case the value of γ^2 can be obtained e.g. by the so called L-curve method (Hansen, 1992).

4. Numerical results

First the information content of the different tensor observations is investigated. In Fig. 3 the main diagonal entries of the normal matrix $(\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A})_{ii}$ are plotted versus the geographical latitude. The weight matrix $\boldsymbol{P}$ was built according to the noise model covariance function. As expected the information content strongly degrades when passing the rim of the data gap. Because of the density of observations the maximum is reached near the rim. Comparing the tensor components the earth-pointing direction shows the largest values in the data area and the smallest in the gap area. For the along- and the cross-track component one can expect a worse accuracy below 83°, but within the gap both the horizontal components provide most information.

For the determination of the optimal regularization parameter the L-curve is used (Hansen, 1992). The L-curve is a plot of the weighted residual norm versus the norm of the regularized solution for the all valid regularization parameters. The optimum is located at the point of maximum curvature, this means the typical corner of the curve in a double-logarithmic plot. Fig. 4 shows the results for each tensor component in a common plot. In all three cases there is a significant corner with the desired parameter. In contrast to Xu (1998) we found this a posteriori parameter choice rule working excellently.

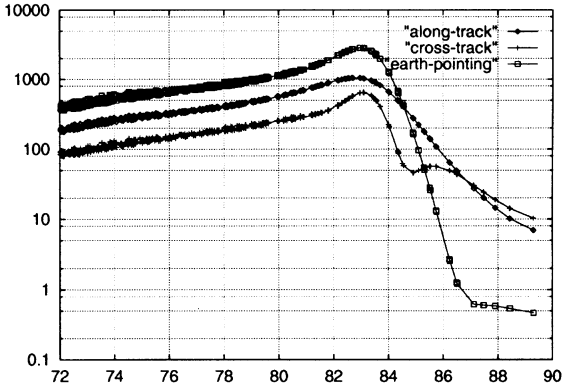

Fig. 3. Main diagonal entries of normal matrix versus geographical latitude.

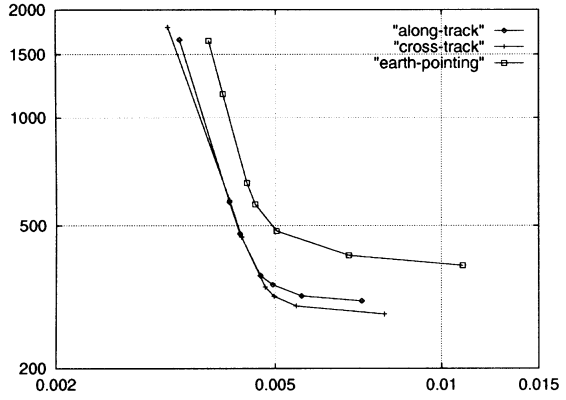

Fig. 4. L-curves for different tensor components, $\|A\hat{\chi} - l\|_P^2$ versus $\|\hat{\chi}\|_M^2$.

The separate solutions of the three components are not presented in this paper. We refer to Kusche and Ilk (2000), where it is shown that cross-track gradiometric observations yield the best result in the gap area, but a ranking seems by far not as distinct as expected from Fig. 3. In Fig. 5 the estimated mean square errors are plotted according to the following matrix

$$\mathrm{MSE}(\hat{\chi}) = \mathrm{Cov}(\hat{\chi}) + \beta\beta^T \tag{9}$$

with $\mathrm{Cov}(\hat{\chi})$ defined in (8). $\boldsymbol{P}^{-1}$ is the variance–covariance matrix of the data and β is the bias, which is predicted from

$$\beta = \gamma^2(\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A} + \gamma^2\boldsymbol{M})^{-1}\boldsymbol{M}\bar{\chi}, \tag{10}$$

where the unknown anomalies are replaced by an estimated mean value $\bar{\chi}$. True mean square errors obtained from the known spherical harmonic model are given in Fig. 6 for comparison. The results displayed in Figs. 5 and 6 compare very well. The main difference is the behaviour at the rim of the gap. The decrease of accuracy is more rapid for the estimated MSEs than in the

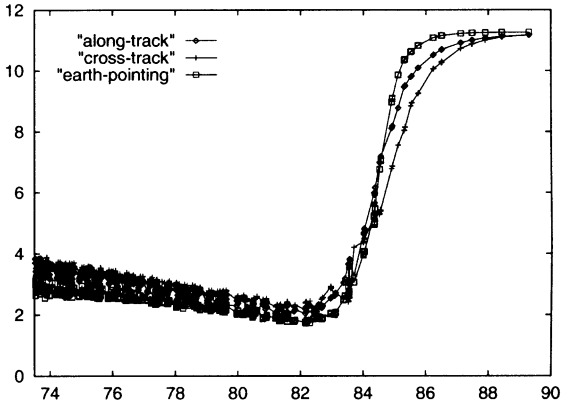

Fig. 5. Estimated MSEs (mGal) versus geographical latitude.

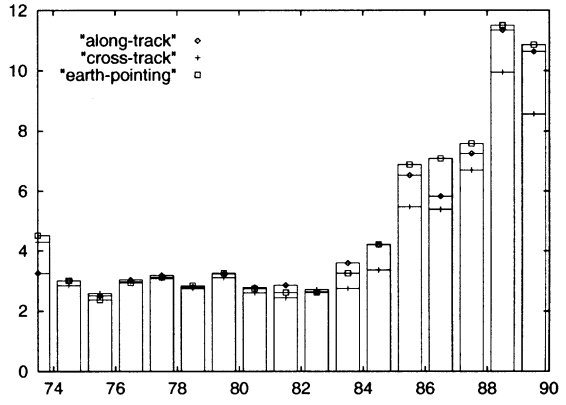

Fig. 6. True MSEs (mGal) versus geographical latitude.

case of true MSEs. In both of them the cross-track component shows slightly the best performance. The coincidence of the pictures indicates that the MSE estimate according to (9) is quite appropriate.

Finally a full-tensor solution is computed by merging the normal matrices. Fig. 8 shows the result of the gravity field recovery in the polar gap area compared to the true field in the sense of the spherical harmonic model (Fig. 7). It is obvious that the fine structures of the field are invisible for the gradiometer. Within the polar gap features on spatial scales less than 500 km get lost, although the coarse part can be recovered. The rms-value in the area of dense data is about 3 mGal compared to a signal rms of 13.4 mGal. The differences of Figs. 7 and 8 are plotted in Fig. 9.

GOCE will be continuously tracked by GPS, but one expects that these data will contribute significantly only to the long-wavelength part of the gravity field. Therefore, uncorrelated range measurements with a variance of $(3\ \mathrm{cm})^2$ had been computed. The accuracy of a solution based on GPS only is very poor with a rms value of 8–9 mGal. It is not possible to determine the field with the chosen $1^\circ \times 1^\circ$ resolution. In a combined solution the improvement is below 0.1 mGal in the rms-value.

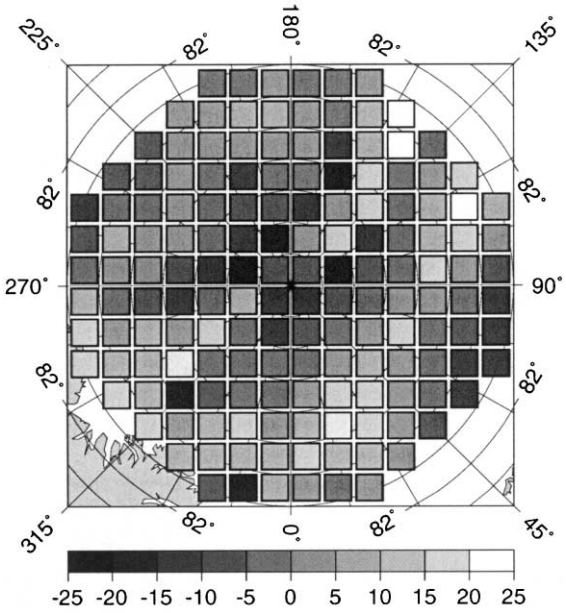

Fig. 7. True mean gravity anomalies (mGal).

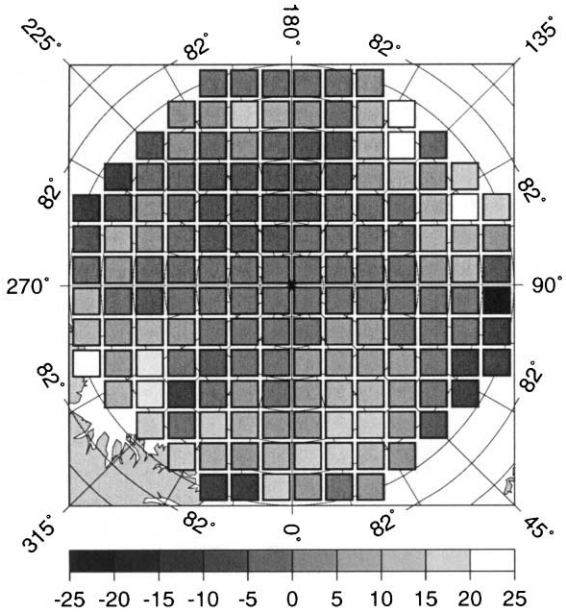

Fig. 8. Full-diagonal recovery (mGal).

Recent developments suggest that the polar regions will be almost completely covered with airborne data in a few years. But one might expect that these data sets will suffer from inconsistencies, biases or long-wavelength errors. The possibilities of improving a regularized GOCE gravity field solution with additional data are investigated. Therefore, we derive "true" $1^\circ \times 1^\circ$

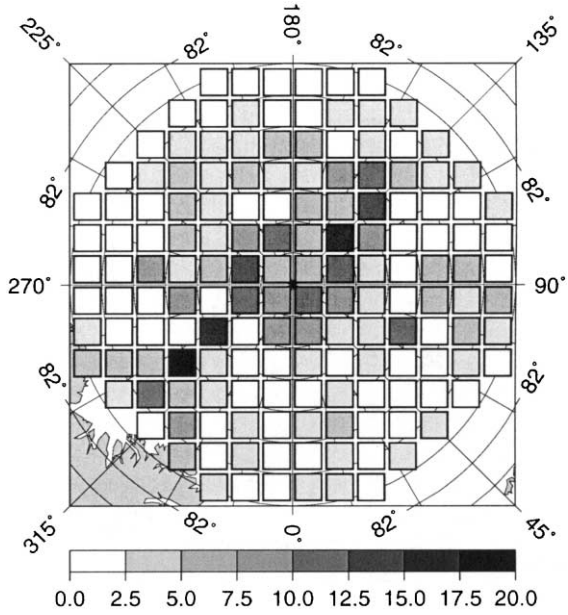

Fig. 9. Differences (mGal).

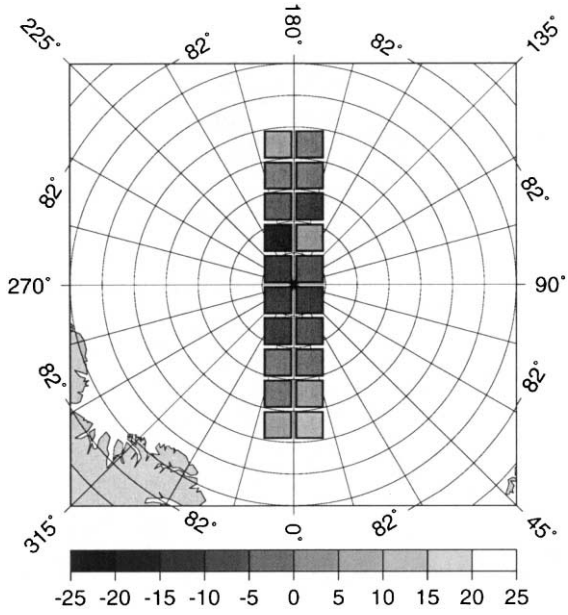

Fig. 10. Additional data (mGal).

mean anomalies from the EGM96 model. They serve as additional surface data with a high weight. Fig. 10 shows these data in an example, where the supplementary measurements are available only within the gap (above 85°). An extended example consists of surface data bridging the gap (above 75°). Also for the type of data two cases are considered: at first ''relative gravity''

in the form of differences of the block values and secondly "absolute gravity", where the block values are directly used. The main results can be summarized as follows (for a detailed description see Kusche and Ilk, 2000). Surface data within the polar gap can solely locally improve the satellite solution. In contrast, it is impossible to obtain overall improvements if only few additional data sets are given. In the relative case datum defects of the surface data can be fixed, in the small case with an accuracy of 1 mGal and in the extended case within an accuracy of 0.1 mGal. So, a combination with relative data seems sufficient, but the coverage of surface data should also extend into the stable non-polar part of the satellite solution.

5. Conclusion and outlook

Our investigations show that $1^\circ \times 1^\circ$ mean anomaly blocks within the polar gaps can be recovered, but at reduced accuracy only. From the rim to the center of the gap the performance reduces from 3 mGal rms up to 10–12 mGal. Measurements in cross-track direction contain the most information, although the differences between the tensor components are small. We found that the estimated MSEs represent a good approximation for the true errors. Additional high-low data from GPS do not improve the results in this regional scenery. The implementation of airborne or terrestrial gravity data can help to improve the recovery results, only if a complete coverage of the polar gap is available.

References

Hansen, P.C., 1992. Analysis of discrete ill-posed problems by means of the L-curve. SIAM 561–580.

Heiskanen, M., Moritz, H., 1967. Physical Geodesy. W.H. Freeman and Company, San Francisco.

Ilk, K.H., Rummel, R., Thalhammer, M., 1995. Refined method for the regional recovery from GPS/SST and SGG. In: CIGAR III/2 Study of the Gravity Field Determination using Gradiometry and GPS, ESA contract No. 10713/93/F/FL. European Space Agency.

Koop, R.M., Visser, P., Klees, R., 2000. Detailed scientific data processing approach. In: Sünkel, H. (Ed.), From Eötvös to mGal, Final Report ESA/ESTEC contract No. 13392/98/NL/GD WP2:29-64.

Kusche, J., Ilk, K.H., 2000. The polar gap problem. In: Sünkel, H. (Ed.), From Eötvös to mGal, Final Report, ESA/ESTEC contract No.: 13392/98/NL/GD, WP5:177–206.

Schneider, M., 1984. Observation equations based on expansions into eigenfunctions. Manuscripta Geodaetica 9, 169–208.

Tscherning, C.C., Forsberg, R., Albertella, A., Migliaccio, F., Sanso, F., 2000. The polar gap problem: space-wise approaches to gravity field determination in polar areas. In: Sünkel, H. (Ed.), From Eötvös to mGal, Final Report, ESA/ESTEC contract No. 13392/98/NL/GD, WP5:331-336.

Xu, P., 1998. Truncated SVD methods for discrete linear ill-posed problems. Geophys. J. Int. 135, 505–514.