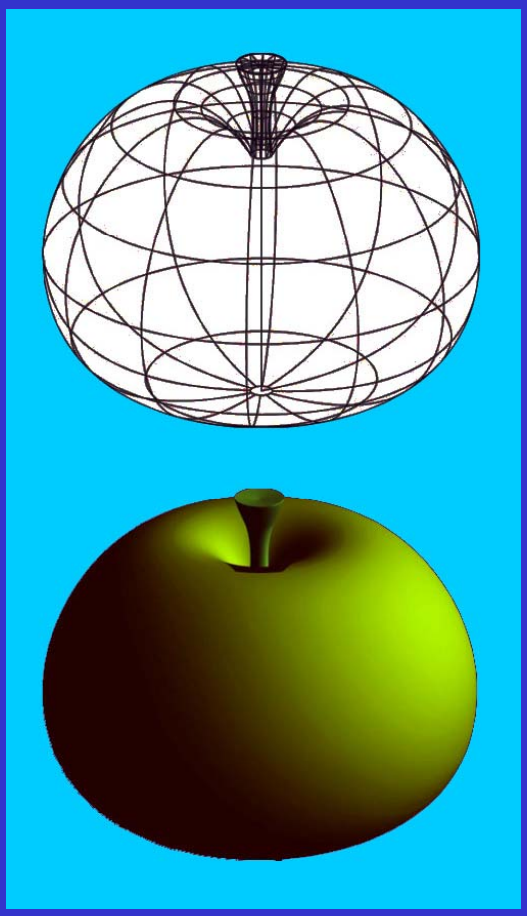

Bureau Gravimétrique International
International Geoid Service
Joint Bulletin

Newton's Bulletin

Issue n° 2, December 2004

International Association of Geodesy
and
International Gravity Field Service

ISSN 1810-8555

Newton's Bulletin N. 2

Summary

SECTION I - "Reviewed Papers"
BGI Papers:

IGeS Papers:

II Section: "Communications and News"

Gravity Field Model UCPH2004 from One Year of CHAMP Data using Energy Conservation

E. Howe, C. C. Tscherning

Department of Geophysics,

The University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen Oe, Denmark

Abstract. A gravity field model has been estimated using energy conservation and fast spherical collocation.

The energy conservation method is based on knowledge of the state vector and measurements of non-conservative forces. This is or will be provided by CHAMP, GRACE and GOCE.

Here the energy conservation method is applied to calculate gravity potential values from CHAMP data measured in July 2002 – June 2003. Precise science orbits and accelerometer data are derived from GFZ Potsdam. When estimating the loss of energy due to external forces only friction is considered, calculated using the along-track acceleration. A scale-factor for the along-track accelerometer has been estimated for each half day.

Fast Spherical Collocation requires data located equidistantly on parallels. It is much faster than general least squares collocation. Fast Spherical Collocation has been used to estimate a gravity field model to degree and order 90, UCPH2004 and a model UCPHcoll to degree and order 60 has been estimated with general least squares collocation using only 1% of data.

Evaluation of the method is made by comparison with EIGEN-2 and TUM1s. Furthermore the gravity field is compared to data from the Arctic Gravity Project (ArGP). Good agreement has been found in the comparisons.

Keywords. Energy conservation, CHAMP, gravity potential, collocation, error correlations

1 Introduction

The energy conservation method has been well documented in the past few years, see for example Han et al. (2002), Gerlach et al. (2003) and Howe et al. (2003a). Every time improvements are found and new ways to calibrate the data are used. Each group has its own way to apply the energy conservation method and to estimate the spherical harmonic coefficients.

We use the energy conservation to calculate height anomalies at satellite height. When doing so we consider the tidal potential corresponding to a rigid earth of the sun and the moon (Longman (1959)), the explicit time variation of the gravity potential in inertial space (Jekeli (1999)) and friction. We subtract the earth normal potential, U, without the centrifugal term. The equation reads

$$\varsigma = \left(\frac{1}{2} v^2 - V_{moon} - V_{sun} - \omega \left(x v_y - y v_x \right) - F - E_0 - U \right) / g \quad (1)$$

where E_0 is an integration constant.

The data used is measured in July 2002 – June 2003. The data are kindly provided by GFZ Potsdam and are Precise Science Orbits (PSO) and accelerometer data. On November 6 and December 9 - 10. 2002, orbit manoeuvres have been made these periods are removed from the data set, see Reigber (2002). We assume that the accuracy is the same before and after the orbit manoeuvre.

Only the along-track accelerometer is used and therefore only friction calculated using this information is taken into account and no other non-conservative forces are included. The friction term reads

$$F = \int |v| a_y dt \quad (2)$$

A scale factor has been estimated for each half day in order to calibrate the along-track accelerometer. The scale factor is estimated by correlating F with the difference between the calculated gravitational potential field and an a priori gravity field model. EGM96 to degree and order 24 is used as a priori gravity field model; see Lemoine et al. (1998). The scale factors for the period are shown in Figure 1. The scale factors show an interesting pattern, which warrant further investigations.

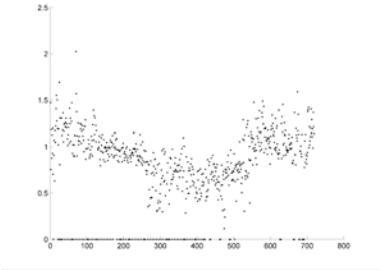

Figure 1 The scale factors for the period. The x-axis is a number corresponding to each half day of the period. The y-axis is the scale factor. The scale factors which are 0 on the figure is days with no data.

Fast spherical collocation has then been used to estimate spherical harmonic coefficients to degree and order 90. A detailed description of Fast Spherical Collocation can be found in Sansó and Tscherning (2003). Comparison have been made with coefficients estimated by general least squares collocation using 1 % of the data, EIGEN-2 (Reigber et al. (2003)), TUM1s (Gerlach et al. (2003b)) and data from the Arctic Gravity Project (ArGP) (Forsberg (2002)).

Bias parameters have been estimated as a method similar to a cross-over analysis of our data and error correlations have been calculated.

2 The gravity field model UCPH2004

EGM96 to degree 24 is subtracted from our data in order to make them statistically more homogeneous. The residual potential values are then up-/downwards continued to a common height of 413 km above the ellipsoid. The up-/downwards continuation is performed using gravity disturbances calculated from EGM96. Fast spherical collocation requires that data are located equidistantly on parallels and at the same altitude. The residual potential values are therefore gridded with 0.5° spacing using collocation. From this grid we estimate the spherical harmonic coefficients and their associated errors using fast spherical collocation. After the estimation EGM96 to degree 24 is added to get a complete set of spherical harmonic coefficients.

2.1 Analysis of the model

Previously 1 month of data was used when estimating the gravity field model UCPH2003_03 which had good agreement with other post-CHAMP models; see Howe et al. (2003b). The standard deviation of the different degrees of the old model and the new model are compared to see if an improvement is obtained by having more data, see Figure 2.

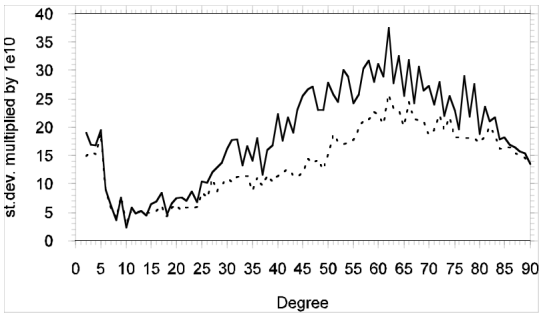

Figure 2 The standard deviation of the old model (solid) and of the new model (dotted)

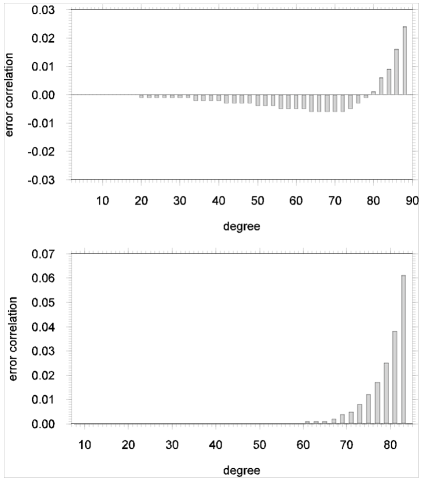

Figure 3 Typical error correlations of the data. The top figure is for degree 90 and fixed order 2. The bottom figure is for degree 85 and fixed order 7.

It can be seen that more data gives a significant improvement of the gravity field model above degree 25. In both models EGM96 to degree 24 is subtracted when estimating the model and then added again. No significant difference is expected between the two models below degree 25.

Fast spherical collocation can be used to compute the error co-variances of the spherical harmonic coefficients, see Tscherning (2004). Figure 3 illustrate some typical error correlations for fixed order. The top figure is for degree 90 and order 2 and the bottom figure is for degree 85 and order 7. The values of the error correlations show interesting patterns. But further investigations are needed in order to explain them.

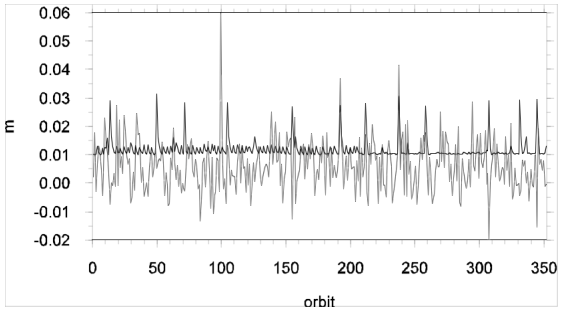

Figure 4 Bias parameters for each day (grey) and the error estimates (black)

2.2 General least squares collocation

Using general least squares collocation we need to solve a system of as many equations as unknowns each having a matrix with combinations of all of the observations. We have approximately 1 year of 10 sec data, 2567422 data points and for a useful test at least 30000 data points are needed. This gives a factor 10 between data and coefficients determined up to degree 60. There are many ways of selecting data. In this study every 100th point is selected. Another way would be to take a mean value of 100 points or to select more data at places with much gravitational variation and less data at places with little gravitational variation.

The result UCPHcoll is compared to UCPH2004. A mean difference of 0.004 m and a standard deviation of 0.50 m are seen. It was not expected to give better result than UCPH2004 since less data have been used. It can be seen though that general least squares collocation gives very good results with only a small amount of data. Further investigations with larger data sets could lead to an improvement of the model UCPH2004.

Least squares collocation gives the opportunity to make a kind of a cross over analysis, where the data do not have to be in the exact same point. It is a good test of the accuracy. A bias parameter is estimated for each day, see Figure 4. It can be seen that the estimated bias parameters are close to their error estimates. This shows consistency in the technique.

3 Evaluation of UCPH2004

The gravity field model is evaluated by comparison with other state-of-the-art gravity field models and with data from the Arctic Gravity Project.

From error analysis it can be seen that there is not much information left above degree 60. The comparisons are on this basis only made using coefficients up to degree 60. UCPH2004 is compared to Eigen-2 and TUM1s. The results are listed in table 1.

Table 1 Differences between UCPH2004 and EIGEN-2 and TUM1s in metre

	Mean	St. dev	Min	Max
EIGEN-2	-0.008	0.52	-2.76	2.57
TUM1s	0.08	0.54	-2.84	2.51

UCPH2004 is within half a meter of the two gravity field models used here. Further analysis show that the main differences are in the polar regions. This could be due to only taking friction in the direction of the velocity vector into account and not sun pressure and cross winds, which are strongest at the poles.

A comparison between the gravity field models and data from the Arctic Gravity Project is conducted to further test the accuracy of the model. If the comparison is made over the entire arctic area (64.0°N – 89.9°N, 179.9°W – 179.9°E) EIGEN-2 and TUM1s fits slightly better than UCPH2004, see table 2. The differences in how well the three models fit the arctic data are within the standard deviation for the three models. The gravity field model estimated with general least squares collocation UCPHcoll fits as well as UCPH2004 even though there is a difference of 0.5 m between the two. A new comparison is made in an area where it previously is seen that the three models disagree (70°N -80°N, 50°E - 70°E), see table 3. Here UCPH2004 has a smaller mean difference to the arctic data than EIGEN-2 and TUM1s and a slightly better standard deviation. UCPHcoll has a very small mean difference of less than 1 mGal to the arctic gravity data but the standard deviation is bigger.

Table 2 Comparison with ArGP data in the entire arctic region. Units are mGal..

	Mean	St. dev	Min	Max
UCPH2004	-0.91	24.29	-216.23	217.40
EIGEN-2	-0.25	23.53	-209.31	221.72
TUM1s	-0.16	23.47	-209.80	221.26
UCPHcoll	-0.86	24.11	-218.74	213.34

Table 3 Comparison with ArGP data in the region (70-80°N, 50-70°E). Units are mGal.

	Mean	St. dev	Min	Max
UCPH2004	-1.42	16.11	-45.93	79.29
EIGEN-2	-3.05	16.54	-53.21	81.73
TUM1s	-2.98	16.63	-52.90	80.56
UCPHcoll	-0.63	17.11	-45.77	81.69

4 Concluding remarks

It is possible with CHAMP data to estimate a state-of-the-art gravity field model using energy conservation and collocation.

We find a difference of 0.5 m between UCPH2004 and EIGEN-2 and TUM1s. The three models all have a standard deviation of about 24 mGal compared to the arctic gravity data. This shows that our gravity field model UCPH2004 is comparable to other state-of-the-art gravity field models.

In order to enhance the gravity field model the entire acceleration vector should be taken into account. Further investigations are needed. The estimated scale factors show that it would induce an error if the scale factor was assumed to be constant. It may even be a good idea to estimate a scale factor more often than each half day, one per revolution for example.

The accuracy has been improved compared to our older models due to more data. Comparison with general least squares collocation has been made. It shows that the techniques are consistent. Furthermore it is seen that general least squares collocation can improve our spherical harmonic coefficients. The only drawback is the large computer power needed and it is rather time consuming.

Acknowledgement. Thanks to the CHAMP data centre for providing the data. Thanks to Louise Sørensen and Henriette Hjorth for data processing. This paper is a contribution to the SAGRADA project sponsored by the Natural Science Council of Denmark

References

Forsberg, R. (2002). Arctic Gravity Project. At *http://earth-info.nga.mil/GandG/agp*

Gerlach, Ch., N. Sneeuw, P. Visser and D. _vehla (2003a). CHAMP gravity field recovery with the energy balance approach. *Adv. Geosciences*, 1, 73-80.

Gerlach, Ch., L. Földvary, D. _vehla, Th. Gruber, M. Wermuth, N. Sneeuw, B. Frommknecht, H. Oberndorfer, Th. Peters, M. Rothacher, R. Rummel and P. Steigenberger (2003b). A CHAMP-only gravity field model from kinematic orbits using the energy integral. *Geophysical Research Letters*, 30(20), 2037, doi:10.1029/203GL018025.

Han, S.-C., C. Jekeli and c. K. Shum (2002). Efficient gravity field recovery using in situ disturbing potential observables from CHAMP. *Geophysical Research Letters*, 29(16), doi:10.1029/2002GL015180.

Howe, E., L. Stenseng and C. C. Tscherning (2003a). Analysis of one month of CHAMP state vector and accelerometer data for the recovery of the gravity potential. *Adv. Geosciences*, 1, 1-4.

Howe E., L. Stenseng and C. C. Tscherning (2003b). CHAMP Gravity Field Model UCPH2003. *Poster presentation at EGS-AGU-EUG Joint Assembly*. Nice France. http://www.gfy.ku.dk/~eva/doc/Egs2003.pdf (20-8-2004)

Jekeli C. (1999). The determination of gravitational potential differences from satellite-to-satellite tracking. *Celestial Mechanics and Dynamical Astronomy*, 75, 85-101.

Lemoine, F., S. Kenyon, J. Factor, R. Trimmer, N. Pavlis, D. Chinn, C. Cox, S. Klosko, S. Lutheke, M. Torrence, Y. Wang, R. Williamson, E. Pavlis, R. Rapp and T. Olson (1998). The development of the joint NASA GSFC and the national imagery and mapping agency (NIMA) geopotential model EGM96. *Tech rep. NASA/TP-1998-206861*. NASA Goddard Space Flight Center.

Longman, I., (1959). Formulas for computing the tidal accelerations due to the moon and the sun. *Journal of Geophysical Research*. 64, 2351-2355.

Reigber, C., (2002). CHAMP newsletter no. 11 and 12, http://op.gfz-potsdam.de/champ/more/newsletter_CHAMP_011.html and http://op.gfz-potsdam.de/champ/more/newsletter_CHAMP_012.html.

Reigber, C., P. Schwintzer, K.-H. Neumayer, F. Barthelmes, R. König, C. Förste, G. Balmino, R. Biancale, J.-M. Lemoine, S. Bruinsma, F. Perosanz and T. Fayard (2003). The CHAMP-only Earth Gravity Field Model EIGEN-2. *Advances in Space Research* 31(8), 1883-1888, doi:10.1016/S0273--1177(03)00162-5.

Sansó F. and C. C. Tscherning (2003). Fast Spherical Collocation – Theory and Examples. *Journal of Geodesy*, 77, 101-112, doi:10.1007/s00190-002-0310-5

Tscherning, C. C., (2004). Computation of error-covariances of estimates of spherical harmonic coefficients using Fast Spherical Collocation. *Draft paper*.

A simple anisotropic model of the covariance function of the terrestrial gravity field over coastal areas

***Jonathan Chenal**[*] **and Jean-Pierre Barriot**[**]**

[*] Service de la Géodésie et du Nivellement, Institut Géographique National, Saint-Mandé, France.
[**] Bureau Gravimétrique International, Toulouse, France.

I. Introduction

The modelling of the gravity and of the geoid at the interface between continents and oceans is a very difficult challenge, as the gravity field over these regions exhibits large anisotropies. We consider in this short paper how to build simple 2D anisotropic covariance functions for collocation purposes. A detailed report is available upon request at BGI (Chenal, 2004).

II. A simple 2D geophysical model of the littoral

We consider a very simple 2D model, where the continental crust (Boillot, 1982) can be considered as a box invariant by translation along the littoral line (the $\vec{j}$ axis perpendicular to the sheet, see Fig. 1)

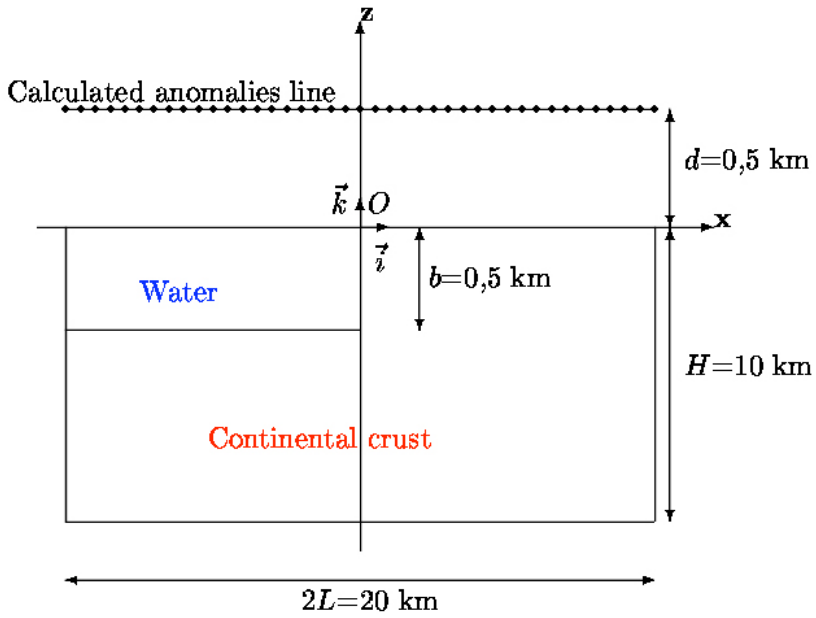

Fig. 1: 2D-model of the littoral: the sources of the gravity field are located in the box -H < z < 0 and –L < x < L (the water slice being excluded). The gravity anomalies are computed on the dotted line, from left to right.

We also consider 2D sources (i.e. sources with infinite length along $\vec{j}$). This leads to

$$\Delta g_{Tot;z}(x',d) = G\sum_{i=1}^{N}\frac{\Delta m_i\,(d-z_i)}{(d-z_i)^2+(x'-x_i)^2} \qquad [1]$$

where $\Delta g_{Tot;z}$ is the gravity anomaly along the vertical z, (x_i, y_i) is the location of each line mass, and x' is the abscissa coordinate along $\vec{i}$.

We suppose that the sources can be either randomly uniformly distributed or following a mesh of cell size e. In both cases, we use the same absolute numerical value for the masses, but with a sign uniformly distributed, to reduce computation costs.

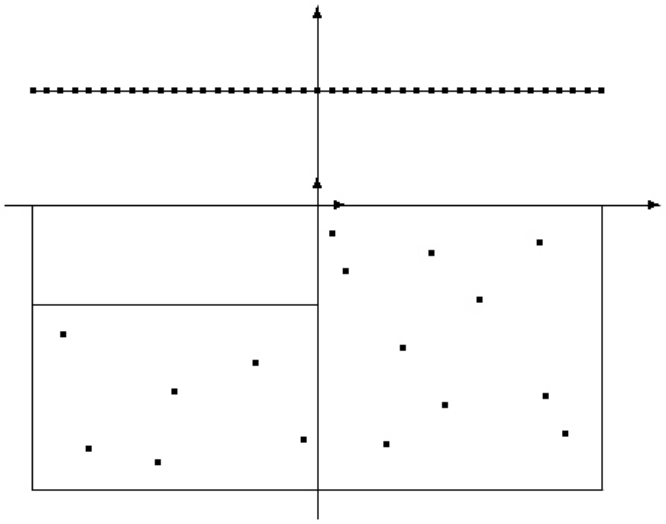

Fig. 2: Random distribution of the sources

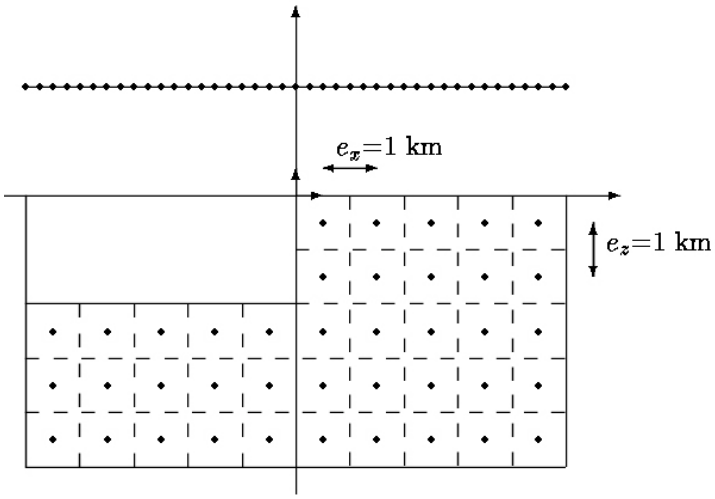

Fig. 3: Network distribution for the sources

If we consider N realizations (drawings) of the sources, the expectation of the related gravity anomaly is (Pelat, 2003):

$$E\{\Delta g_{Tot}(x', d)\} = \frac{1}{N_T} \sum_{j=1}^{N_T} \Delta g_{Tot}^{(j)}(x', d)$$

[2]

Besides, if we note

$$\Gamma^{(j)} = \begin{pmatrix} \Delta g^{(j)}_{Tot;z}(x'_1, d) \\ \vdots \\ \Delta g^{(j)}_{Tot;z}(x'_p, d) \\ \vdots \\ \Delta g^{(j)}_{Tot;z}(x'_r, d) \end{pmatrix}$$

$$\overline{\Gamma} = \begin{pmatrix} E\{\Delta g_{Tot;z}(x'_1, d)\} \\ \vdots \\ E\{\Delta g_{Tot;z}(x'_p, d)\} \\ \vdots \\ E\{\Delta g_{Tot;z}(x'_r, d)\} \end{pmatrix}$$

[3] and [4]

then the covariance matrix of the gravity anomaly over the N realizations is

$$V = E\{\Gamma\, \Gamma^T\} - \overline{\Gamma}\, \overline{\Gamma}^T$$

[5]

The most practical way to verify that this matrix is positive definite (as it should be for a covariance matrix) is to try to decompose it numerically by the Cholesky method (this also apply to the model matrices, see thereafter).

We have chosen by trials and errors two mathematical models for the anisotropic covariance function between two any points P and Q, the first one being

$$C_1(P,Q) = C_0\, exp\left(-\frac{(x_P - x_Q)^2}{A_e^2}\right)\ \left[1 + \frac{|b|}{P_h}\ \tanh\left(\frac{x_P + x_Q}{A_t}\right)\right]$$

[6]

with the free parameters C_0, A_e, P_h and A_t. The parameter x is the abscissa coordinate (see Fig. 1). The hyperbolic tangent is introduced to fit the behaviour of the covariance along the diagonal of the matrix (see thereafter), and the exponential introduces the usual decay with the distance. The second model is

$$C_2(P,Q) = C_0\, exp\left(-\frac{|x_P - x_Q|}{A_e}\right)\ \left[1 + \frac{|b|}{P_h}\ \tanh\left(\frac{x_P + x_Q}{A_t}\right)\right]$$

[7]

We note that the behaviour in the region over the sea must be equal to the one in the terrestrial region when the depth of the water is null. (i.e. when b =0).

We give the first model as information, as we rapidly discovered that it is was leading to inconsistent models of covariance matrices (i.e. non positive definite), and so we focused on the second model.

To compare the matrix models M_1 to the pseudo-experimental covariance matrices M_2, we have chosen, also after trials and errors, to minimize the criterion

$$R_M(P,Q) = \frac{|M_1(P,Q) - M_2(P,Q)|}{M_1(P,Q) + M_2(P,Q)}$$

[8]

with respect to C_0; P_h, A_t and A_e. As this criterion is strongly non-linear, the minimization was done by scanning the whole set of free parameters, which is permitted by their small number N_T.

III. Results

We first verified that the expectations of the gravity anomalies were zero over the computation profile (see Fig. 1) for a sufficient number of drawings (see Fig. 4 below).

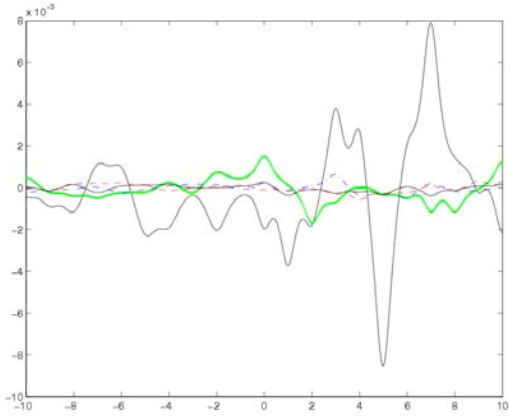

Fig. 4: Numerical values of the gravity anomalies expectations for different values of N_T (black line: 100; dotted green: 1 000; semi-dotted, semi-dashed blue: 5 000; the dotted red: 10 000; small black: 20 000. The x-coordinates are in km, the y ones in 10^{-1} mgal. A random distribution of sources was used. A similar result holds for a network distribution.

We can then consider that the gravity anomalies expectations are centred for N_T = 20 000, and we retained this value for the following computations.

III-a Results for a network distribution of sources

The "pseudo-experimental" covariance matrix obtained for the network distribution is then as follows, from formulas [3] and [4]:

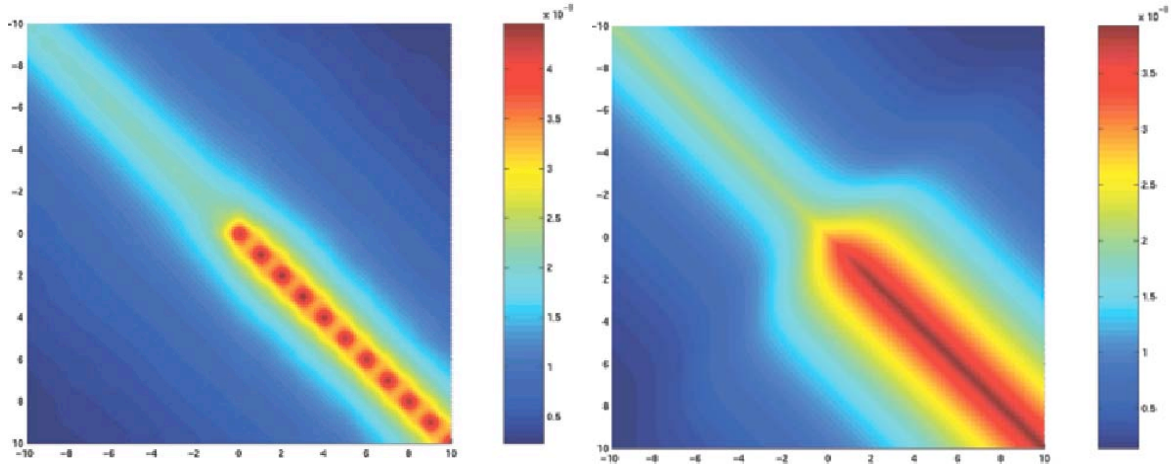

Fig. 5: "pseudo-Experimental" (left) and fitted (right) covariance matrices coming from a network of linear sources. The fit is obtained with C_0=*3.0 10*$^{-8}$; A_t=*1.93* km; P_h=*1.55* km; A_e=*6.92* km.

One can note in Fig. 5-Left that, as expected, the variances in the region over the sea (upper left corner) are lower than in the terrestrial region (lower right corner). The local maxima over the terrestrial zone correspond to mass anomalies situated just above the gravity anomalies evaluation points. The best fit for this "pseudo-experimental" matrix, by using formula [7], is shown in Fig. 5-Right.

For easy comparisons, we also defined some cuts in these matrices, as:

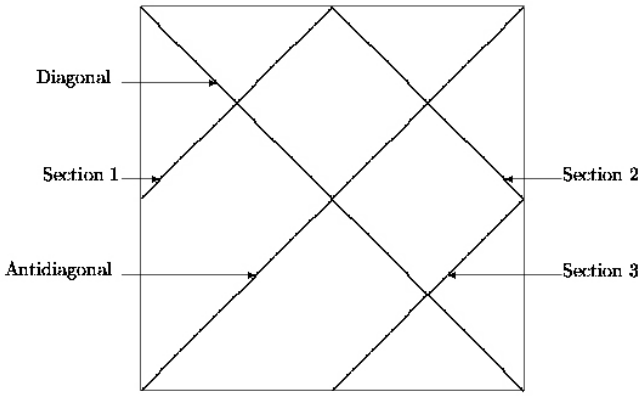

Fig. 6: Definition of the cuts over the covariance matrices.

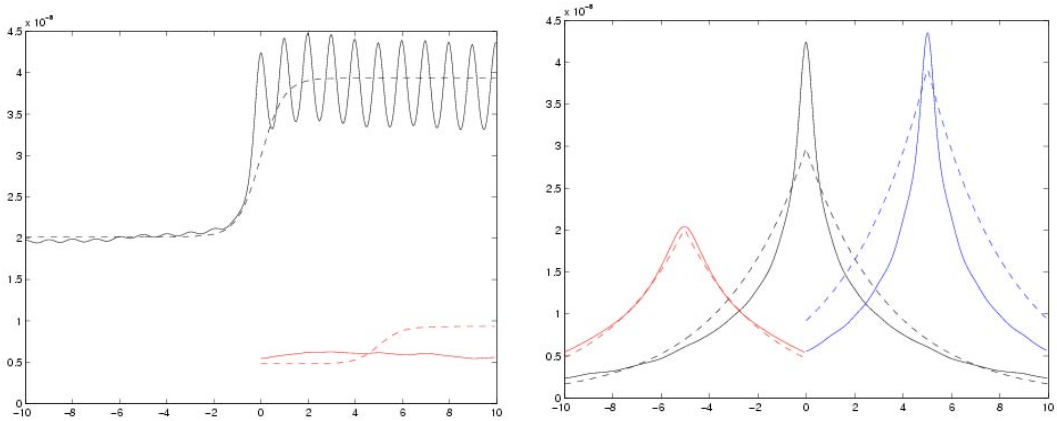

Fig. 7: Comparison over the cuts defined on Fig. 6 of the "pseudo-experimental" (plain lines) and model covariance matrices (dashed lines) for a network distribution of sources. Left: diagonal (black lines), and section 2 (red lines). Right: antidiagonal (black lines), section 1 (red lines) and section 3 (blue lines).

We can see on Fig. 8 that the fit (hyperbolic tangent) along the diagonal is very acceptable, except for the undulations caused by shallow network sources. The antidiagonal behaviour is also quite good.

III-b Results for a random distribution of sources

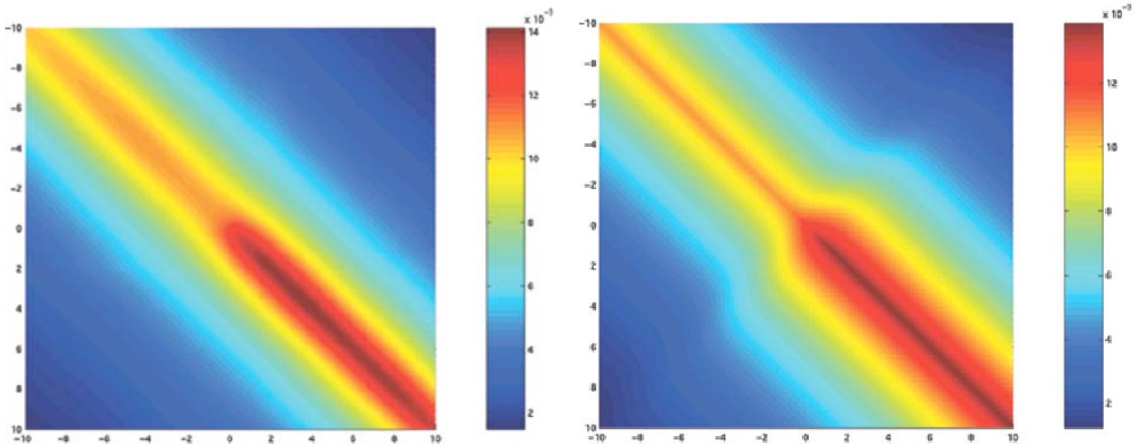

Fig. 8: "Pseudo-experimental" (left) and fitted (right) covariance matrices coming from a random distribution of linear sources. The fit is obtained with C_0=1.23 10^{-8}; A_t=1.43 km; P_h=3.88 km; A_e=8.70 km.

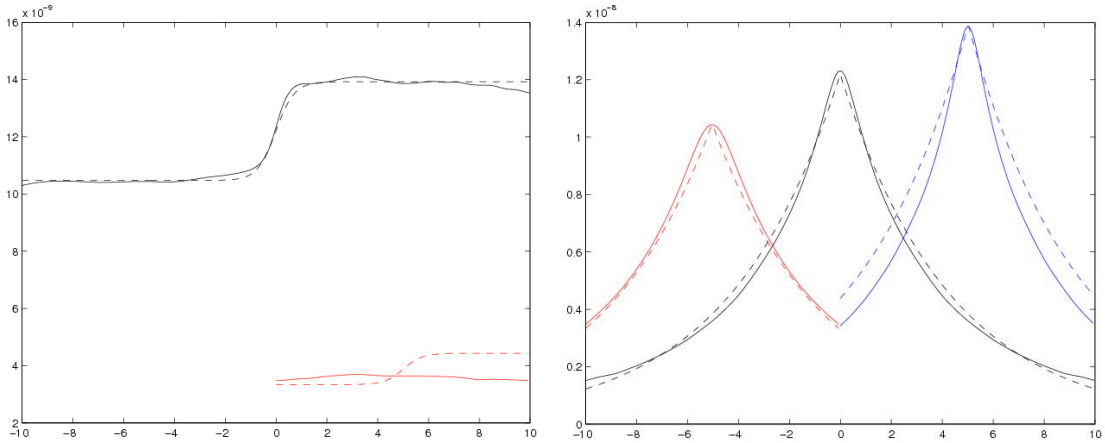

Fig. 9: Comparison over the cuts defined on Fig. 6 of the "experimental" (plain lines) and model covariance matrices (dashed lines) for a random distribution of sources. Left: diagonal (black lines), and section 2 (red lines). Right: antidiagonal (black lines), section 1 (red lines) and section 3 (blue lines).

The results for the random distribution of sources are slightly better than for the network distribution (see Fig. 8 and 9). This is obviously due to the presence of artefacts in the "pseudo-experimental" data because of the network distribution.

IV. Conclusion

This work is a first step about the modelling of anisotropic covariance function over littoral regions, and show that a simple model like function [7] can be of interest. The main difficulty will be of course to take into account real shorelines and the varying water depth of the margins.

Acknowledgements

We would like to thank Sébastien Petitjean, from the Observatoire Midi-Pyrénées in Toulouse, France, for his help and comments.

References

(Boillot, 1982) Boillot G., *Géologie des marges continentales*, Masson, 1982.

(Chenal, 2004) Modélisation des fonctions de covariance du champ de gravité en zone littorale. Rapport de stage au Bureau Gravimétrique International, DEA « Dynamique des Systèmes Gravitationnels » (Observatoire de Paris, France), 2004.

(Duquenne, 2004) Duquenne H., *Le champ de pesanteur terrestre et les méthodes modernes locales de sa détermination*, DEA « Dynamique des Systèmes Gravitationnels » (Observatoire de Paris, France), 2004.

(Moritz, 2001) Moritz H., *Advanced Physical Geodesy (Reprint)*, Civil and Environmental Engineering and Geodetic Science (Ohio State University, Columbus, Ohio, USA), 2001.

(Pelat, 2003) Pelat D., *Bruits et signaux (Introduction aux méthodes de traitement des données)*, Ecole Doctorale d'Astronomie et Astrophysique d'Ile-de-France (Observatoire de Paris, France), 2003.

[1]The nontidal gravity change at Xiangshan seismostation in Beijing

Jia Minyu① Guo Youguang② Gao Rongsheng③ Huang Dalun② Liu Shaoming① Fang Yongyuan② Xu Jinyi②

①Institute of Seismology, China Seismological Bureau, Wuhan 430071, China

②National Institute of Metrology, Beijing 100013,China

③Department of Monitoring and Predicting, China Seismological Bureau, Beijing 100006, China

Abstract

From March 1988 to March 2001, 58 absolute gravimetric station determinations have been carried out by National Institute of Metrology cooperating with China Seismological Bureau with the NIM－Ⅱ Absolute Gravimeter at Xiangshan seismostation among which 44 groundwater level observations have been carried out simultaneously. This paper studies the mechanism of gravity change at Xiangshan seismostation from various point of view. The main conclusions include: (1) the groundwater activity is the main local interferential source, the groundwater levels correlate with gravity observation values by segments and its effect on gravity can be corrected by a 5^{th} polynomial; (2) the effect of local crustal deformation is very small, thus it can be ignored; (3) the earthquake activity can result in short - term change of the gravity field and its maximum magnitude comes to 0.333 μ ms^{-2}; (4) the gravity value approximately drops linearly by 0.191 μ ms^{-2} from 1988 to 2001, the average decrease is 0.0147 μ ms^{-2} per year, and this gravity change belongs to global or regional gravity change.

Key words: Xiangshan Beijing; absolute gravimetry; nontidal gravity change

Introduction

Since the middle 20^{th} century, nontidal change of gravity field has become one of focal problem which scientists pay the most attention to. Local gravity change is mostly related to earthquake and volcano activity. The period of the change is very short and it used to be observed by relative gravity measurement. There are many achievements about this reportedly. Regional and global gravity change is related to the phenomena of plate movement, earth spin, and core rotation. The period of the change is long and it is usually observed by absolute gravity measurement. So far, there are few achievements about this reportedly. The absolute gravimetric results include not only short-term change but also long-term change. To study it not only has important significance for the study of earthquake prediction and geodynamics but also has important effect on the establishment of the dynamic datum and model of earth gravity field.

1. Measurement and its results

The Xiangshan absolute gravity station is installed in a room of the first floor of west building at Xiangshan seismostation. The pier is set up on bedrock and very stable. There is a groundwater well which is about more than 3m from the gravity station in another room to the west of the station. The Xiangshan seismostation is in a cove to the east of the Xiangshan Park, which locates at the east piedmont of West Mountain, Beijing. Here the surroundings are quiet and tasteful, and the observation condition is very good.

Form Mar. 1988 to Mar. 2001, 58 absolute gravity determinations have been carried out by

* Foundation item: The State Natural Science Foundation (49974019).[1]

National Institute of Metrology cooperating with China Seismological Bureau among which 44 groundwater level determinations have been carried out simultaneously. The measurement interval was different greatly: there were 11 determinations in 1988 and 1989 each, there was only once in 1998, the number of measurement was between 2 and 5 in other year.

The NIM-Ⅱ free-fall gravimeter which was developed by National Institute of Metrology has been used for absolute measurements. The meter has participated in the International Absolute Gravimeters Comparison held in 1980, 1985, 1989 and 1997. Its observation results were well evaluated in the word and especially its deviation from median of all meters is relatively small in 1997. In 1995 the meter was partially improved. Its uncertain range was further reduced to 0.04-0.06 μ ms^{-2}.

Around 16 consecutive drops are combined to one set, about 30 sets have been generally performed on each determination and distributed over 12-24 hours.

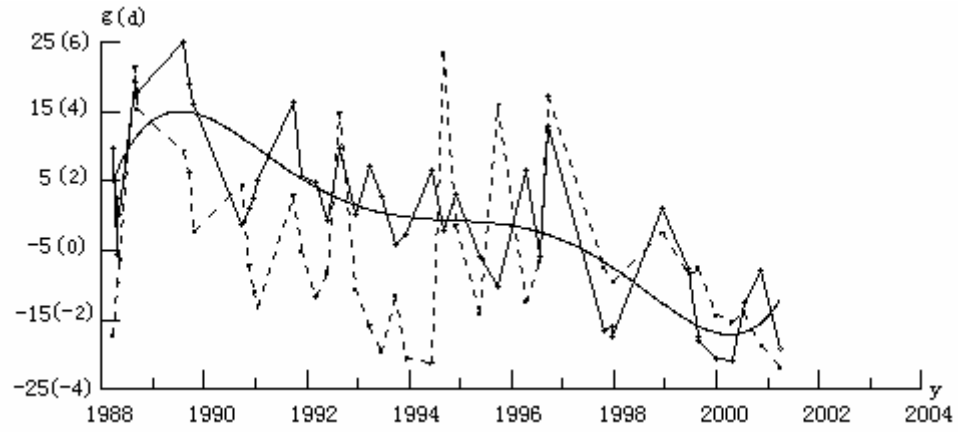

Fig.1 Absolute gravity and groundwater level measurements at Xiangshan station,
Real line is gravity, unit for g-axes is $10^{-8}ms^{-2}$; broken line is water level, unit for d axes is meter,
g or d is the differences between the observation values and their mean.

In observation the effects of some factors, such as electron circuit delay, the limited light velocity, electromagnetic field, vertical level of beams and vacuum level of falling-body cabin etc, were corrected or under strict control. The observation results were corrected with solid tide, air pressure, gravity vertical gradient and polar motion. The gravity and groundwater level changes obtained by the measurements are showed in Fig.1. The smooth curve in Fig.1 represents a fitting 5^{th} polynomial of the g values.

2. The effect of groundwater activity and its correction

Groundwater activity is a most active factor resulting in gravity nontidal change. Not only its effect on the gravity field is large and ubiquitous, but also water level change is fast and the relation between the level and gravity is complicated. However, for the studies of many geoscience problems, it is only an interferential factor and must be separated from observation data. The ground water can be divided into many types, such as vadose water, phreatic water, confined water, fissure water and karstic water etc. Different types of groundwater have different effects on the gravity value of observation station. The phreatic water of them has a great effect on gravity field because it is free aquifer. Above all, we study what type of groundwater the

Xiangshan well water level represents.

In order to find an annual law of water level changes, Fig.2 is drawn, where d-axes is water level, m-axes is measured time overlooked years. It is clear that from January till May the water levels are low, by June the level begins to rise, in August and September they reach their maximum value, then, it begins to drop. This is the typical character of phreatic water changes. It originates from the effects of rainfall and evaporation. Some relatively big fluctuations in the Fig.2 result from difference between years; in other words, it results from long-term change of groundwater level. For example, some obvious lows appear in the years of 1993 and 1994 when the groundwater levels are generally on the low side (see Fig.1).

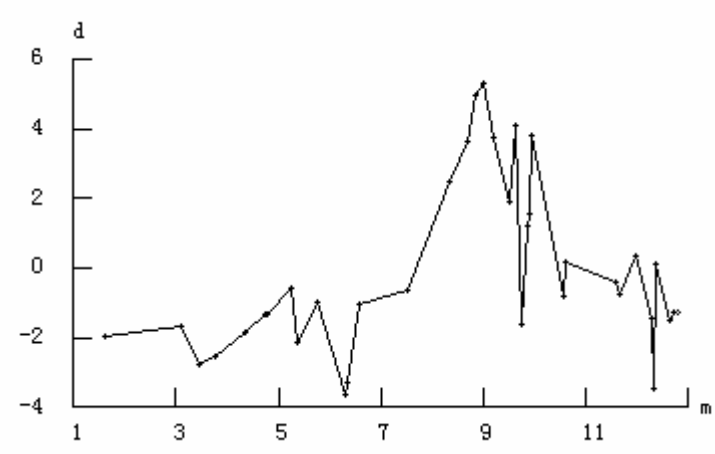

Fig.2 The characters of yearly changes of
Well water level at Xiangshan station
d is the same with Fig.1

In many cases there lies a linear relation between phreatic water changes and their gravity effects, for example, there is $\dot{g} = 42\delta\rho\delta d$ in North China Plain (Jia et al 1983). However, from Fig.1 it can be roughly seen that the relation in Xiangshan point is more complicated. It seems that the gravity changes and the water level changes both represent a tendency to long-term decrease, but the latter also occurred a bigger jump in 1994. In order to prospect for the relation between both of them Fig.3 is drawn. The fitting curve in Fig.3 is a 5^{th} polynomial, its equation is

$$\hat{g} = -2.1793 + 6.5079d + 2.5211d^2 - 0.6482d^3 - 0.1672d^4 + 0.0318d^5 \qquad (1)$$

Fig.3 and equation (1) show the curved correlation between groundwater levels and gravity values. In the two segments where d is -2.68~-1.03(low water level) and 2.87~5.05(high water level) respectively, gravity values negatively correlate with water levels, but in the segment where d is –1.03~2.87(medium water level), both of them have positive correlation. The bends at the both ends result from the cut of gravity data, a mathematical factor, so it has no physical meaning and it will not be considered. From the gravitational formula it can be known that if a water stratum is above gravity station, its gravity effect negatively correlates with the level change, and otherwise, the gravity effect positively correlates with one. Combining the characters of topography and stratum as well as the conditions of water source around Xiangshan station, the phenomenon of correlation in segments can be explained.

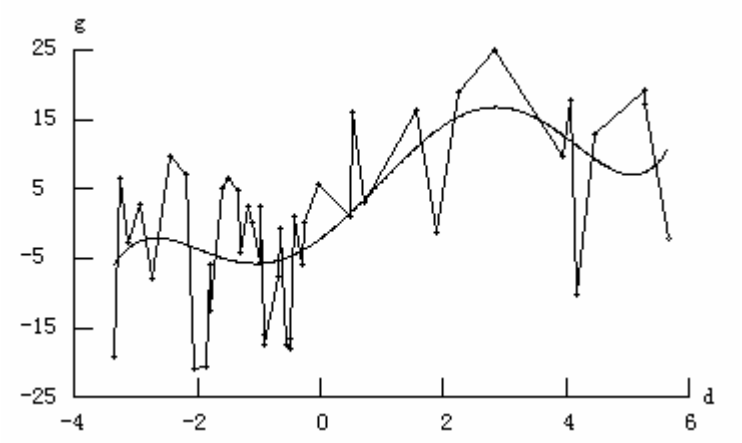

Fig.3 The relation between gravity and well water level observation values

g and d are the same with Fig.1 and Fig.2.

Because the observation station is located at the foot of east of West Mountains, the phreatic aquifer near the station and in its east (PANE) is lower than the pier, but the phreatic aquifer of the western mountainous regions (PAWM) is higher than the pier. Generally the gravity effect of PANE is bigger than PAWM, thus it comes into being that the gravity effect positively correlates with the water level change in the medium level segment. After groundwater level continually drops to the low segment, the water contained in PANE is nearly exhausted; whereas the water in PAWM is relatively abundant because of the supply of water from the mountain. Especially, when the groundwater level would be reaching its nadir in 1993~1994, an artificial pond which is about 200m in the west of the observation station and stores water 500m^3 was built, it provides more water for PAWM. Now, PAWM showed so main effect on gravity that the positive correlation in the low level segment comes into being. As for high water levels， they all occur in rainy season. Here there is plenty more surface water, but the flow velocity in mountains is bigger than plain, therefore PANE could keep being saturated status(the level didn't almost change) and the level of PAWM could change, and thus the negative correlation in high level segment comes into being. The quantitative calculation by formula (1) shows that if effective porosity of rock stratum and well water level change are the same, the gravity effects of each segment at Xiangshan station are in turn (from low to high) 21%, 55%, 42% of that at most stations in North China Plain(Jia, et al. 1983). This further proves that the aquifers in east and west of the observation station have a function of mutual counteraction and the effect of the flat is greater than the mountain. According to the prospecting on the spot the dip angle of massif rock strata is 70°, the penetration speed of the groundwater will be fast, so that it is possible that the phreatic aquifer in different area controls the well water level in different time. In addition from Fig.3 it also can be seen that the correlativity in low and high segments is weaker than medium. This indicates that the effects of other factors, such as surface water, earthquake activity, are large.

Fig.4 is mapped after the observation values in Fig.1 are corrected with groundwater according to the equation (1). The magnitude of gravity changes becomes smaller and the fitting polynomial curve becomes straighter when comparison with those in Fig.1. From the discussion in the 4th and 5th part of this paper it can be known that the gravity effect of earthquake events and the long-term change of gravity field are more prominent and reasonable, which proves that the groundwater correction is valid.

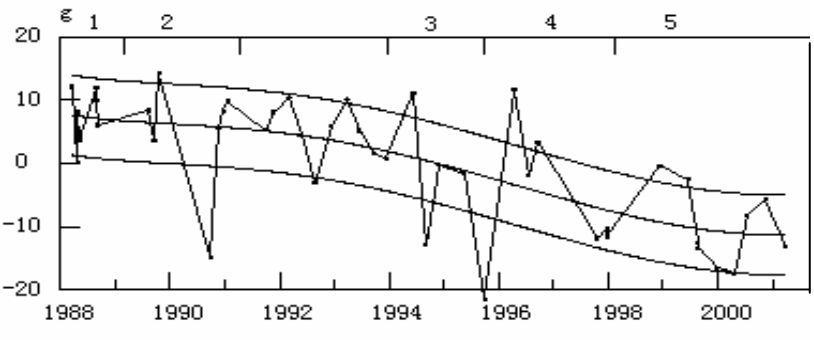

Fig.4 The gravity changes corrected with groundwater

Unit for g-axes: $10^{-8}ms^{-2}$; the numbers on the top are order number of gravity effects from earthquakes.

3. The gravity effect of local crustal deformation

This part discusses the effect of vertical displacement of the observation pier and the gravitational effect of the crustal deformation around the observation station.

In order to monitor the stability of the bedrock, the Beijing Surveying and Mapping Institute distributes a 50km measuring line in northwest Beijing. It includes more than 10 bedrock observation stations, such as Xiangshan, Yuyuantan etc, and 4 times of primary leveling were carried out in 1989, 1992, 1995 and 1998. The results show that the changes of elevation are only few mm, very small, and not continuous. So the gravity effect of elevation changes can be neglected.

At the same time, the Beijing Surveying and Mapping Institute also carried out subsidence observation in the urban area of Beijing and founded two local subsidence areas. One's center is about 22km north by east 50° from the Xiangshan station and its maximum subsidence is 385mm. Another one's center is about 44km south by east 15° from the Xiangshan station and its maximum subsidence is 363mm. When the subsidence area divided into 7 lays, the linear integration method (Talwani et al 1960, Lio, et al. 2002) is used to calculate their effects on Xiangshan absolute station. Their gravity effects are 0.13 and 0.04nms^{-2} respectively, which are very small.

In a word, the gravity changes at Xiangshan station absolutely don't result from local crustal deformation.

4. The gravity changes resulting from earthquake activity

The interference of local surroundings factors is removed in the gravity changes in Fig.4. So except the residual measurement errors they should be related to local tectonic activities or the geosciences phenomena in wider range. It is obvious that the gravity changes in Fig.4 can be divided into two parts: one part is a long-term change represented by smooth curve; another part is a short-term change described by the broken line. In this section the latter will be discussed.

The average uncertain region of 58 absolute gravity determinations is ± 0.063 μ ms^{-2}. When the smooth curve is taken as reference and ± 0.063 is taken as limit (confidence interval), two parallel curves are dropped in Fig.4. The middle part between them is an uncertainty range, and

the outside of them is viewed as a region produced the effect of tectonic activity.

Tab.1 provides the earthquakes Ms≥5.0 which occur within 300km around the Xiangshan station since 1982, and together there are 9 ones. They can be divided into 5 groups according to their occurring time. (See 2nd column of Tab.1.). It is interesting that we can find corresponding gravity changes with them one by one in Fig.4, which are also made signs from 1 to 5 on the top of that. The gravity changes usually appear earlier than earthquakes about 1-2 years, that indicates the former result from gestation of the latter, and so the gravity changes are named an anomaly for portents of earthquake.

Table.1 The earthquakes Ms≥5.0 within 300km around Xiangshan station

Order number	Group Number	Time	Ms	L	ϕ	Place
1	1	1989.10.18	5.7	113.88	39.94	Datong—Yanggao
2	1	1989.10.19	5.9	113.91	39.92	Datong—Yanggao
3	1	1989.10.19	5.5	113.87	39.92	Datong—Yanggao
4	2	1991.3.26	5.8	113.80	40.00	Datong—Yanggao
5	2	1991.5.30	5.1	118.20	39.50	Fengnan
6	3	1995.10.6	5.0	118.50	39.80	Guye
7	4	1998.1.10	6.2	114.30	41.10	Shangyi
8	5	1999.3.11	5.6	114.60	41.20	Zhangbei
9	5	1999.11.1	5.6	113.90	39.80	Hunyuan—Yangyuan

It looks as if the anomalies of portents have the following characteristics: (1) Every group anomaly contains process: the gravity value first rises, then drops, and then the earthquake occurs; the anomaly of serial number 2 also includes two rising-dropping processes, witch may be caused by 2 earthquakes happened respectively in the east and west of Xiangshang station. (2) The magnitude of the anomalies is related to an intensity of earthquake: the earthquake magnitudes of the 2nd and 4th group are big, so the anomaly magnitudes are also big (up to 0.333 μ ms^{-2}); the earthquake magnitude of the 3rd group is the smallest, so the anomaly magnitude is also the smallest. (3) The effects of earthquake events between which time interval is relatively short on gravity field mutually spliced. The first group of earthquakes is an earthquake swarm and the magnitudes are all relatively big. However, its anomaly form is quite incomplete. The incomplete rising segment is caused by truncation data. The short and small dropping segment results from interference of the 2nd group of earthquakes. Also the short and small rising segment of the 2nd group results from the interference of the dropping segment of the1st group, i.e. the dropping segment of the 1st group and the rising segment of the 2nd group mutually superpose.

5. The long-term changes of the earth gravity field

The smooth curve in Fig.4 is a 5th polynomial. Its equation is:

$$\tilde{g} = 1.7090 - 1.9786t - 0.1980t^2 + 0.0187t^3 - 0.0030t^4 - 0.0002t^5 \qquad (2)$$

Where $\tilde{g}$ is the long-term change of gravity, and t is the time represented by subtracting 1992.8890 from the year corresponding to $\tilde{g}$ (the month and day corresponding to $\tilde{g}$ are reduced to decimal fraction which takes year as unit). The curve in Fig.4 shows the long-term drop process of gravity and this process approximate to linearity. Using Equation (2) we obtained that the gravity decreases by 0.191 μ ms^{-2} altogether from March 1988 to March 2001 and the average decrease is 0.0147 μ ms^{-2} per year. This result is consistent with people's theory estimate of gravity long-term change.

Such long-term changes of gravity field have local or global significance. The chief purpose of International Absolute Gravity Basestation Network is monitoring the global gravity long-term changes, especially the gravity changes caused by earth spin and earth core rotation (Boedecker et al 1986, 1993). The observation result at Xiangshan will make contribution to this worldwide concerned study. The author will also further study the mechanism of gravity long-term changes at Xiangshan when there are more results published.

This paper is in agreement with JILAG-3 results (Torge et al. 1999, Jia et al. 1998).

Acknowledgments

The authors wish to express their gratitude to China Seismological Bureau, National Institute of Metrology and The State Natural Science Foundation, China, which generously sponsored the project. We greatly appreciate the Beijing Surveying and Mapping Institute for the provision of crustal deformation data. We gratefully acknowledge the help during the research or writing this paper and the provision of additional information by Gao Facheng of director of Xiangshang Seismostation, Dr. Ludger Timmen.

Reference

Boedecker, G., Th.Fritzer, 1986, International Absolute Gravity Basestation Network, I.A.G—SSG 3.87 Status Report March 1986. Veroff Bayer. Komm, fur die internat. Erdmessnng der Bayer. Akad. d. Wissensch., Astron.—Geod. Ard.,Heft Nr. 47, Munchen.

Boedecker, G. 1993, IAGBN Catalogue of Stations and observation, IAG conference report in Aug 1993 in Beijing.

Jia M., You Z., et al. 1983, Influences of the motion of the underground water on precision measurement of the gravity and the method of eliminating them. Crustal deformation and earthquake, Vol.3, No.1, 50—67.

Jia, M., Xing, C., et al. 1998, Gravity changes with time in Yunnan and Beijing observed by absolute gravimetry, Bur. Grav. Int., Bull. d'Inf., 83, 37-50. Toulouse.

Lio Sh., Gao R., et al. 2002,The Local Crustal Deformation around Xiangshan Absolute Station and Its Effect on the Gravimetric Result, Journal 0f Geodesy and Geodynamics, 22(2), Wuhan (in Chinese).

Talwani, M. And Ewing, M., 1960. Rapid computation of gravitational attraction of three—dimensional bodies of arbitrary shape, Geophysics, 25:203—225.

Torge W., Schnull M., et al. 1999, Absolute and Relative Gravimetry 1990/1992/1995 in the Western Yunnan Earthquake Prediction Experimental Area, Deutsche Geod. Komm., Reihe B, Muenchen.

Reasoning about the presence of outliers in the European quasi-geoid data set

Damiano Triglione
Politecnico di Milano, Polo Regionale di Como, Italy
damiano.triglione@polimi.it

Abstract
The purpose of this work is studying the (possible) relation between the distribution of the outliers found in the European quasi geoid data set and the magnitude of the slope (gradient) in that points, either in the same signal or in the topographic heights' signal. As we will see, this relation is highlighted by a Montecarlo method. A particular attention has been paid to the fact that, in a so wide area, the east-west step (used to calculate the gradient magnitude in the gridded data) has to be adapted to the changing curvature of the ellipsoid.

A definition of "Outlier"

An intuitive definition of outlier (taken from [14]) can be: "an observation that stands so away from the other ones that it leads to suspect that it could have been generated by a different mechanism" (another insight is presented also in [2]).
Therefore the "evidence" of the deviation presented by the erroneous data is the indicator of their probable incompatibility with the remaining data. Historically the error rejection was performed by people who were responsible of data preprocessing: watching the globality, they were able to identify the suspected minority. Nowadays an automatic procedure is desirable.
A deep exam of anomalous data is imposed by the awareness that just one outlier can "pollute" a big set of measures.

Introduction

In IGeS archives, several geoid data sets are collected. For the european area (North=77°, South=25°, East=67.5°, West=-35°), there are available two kinds of data: QuasiGeoid and Geoid (see for reference [9],[10],[26],[27]). Since a new project for the European Geoid re-computation is under definition, it was considered as particularly interesting to examine the old geoid data, chasing for outliers. We have then decided to apply our own software (hereafter described) to the two datasets, finding that about 90% of the outliers in each data set are common. So we decided to focus our attention only on the QuasiGeoid.
The software "r.outdet", used for the current outlier rejection, is an evolution of "r.outldetect", developed a few years ago. Both the modules were devoleped by us in the last years, exploiting also interesting routines offered in [19]. The actual release extends the capabilities of the open source GIS G.R.A.S.S., a software environment full of tools for processing spatial distributed data of any kind.
The software's core method of processing is placed in the analysis of every point of the data set, around which a moving window is opened. The size of this window is user-defined, and determines how many surrounding points will be involved in the test; the target is to decide whether the central point is an outlier.
More precisely: for a rectangular window of side $(2k+1)\delta_x$ x $(2k+1)\delta_y$ (of course it should be an odd number, while k is any integer larger than zero), when the data are gridded, there are

$$N_k = (2k+1)^2 - 1 = 4k\,(k+1)$$

observations $h_{obs,i}$ around the central h_{obs}. With an appropriate interpolation model, and using only the surrounding values, the central value $\hat{h}$ is estimated and compared with the central observation h_{obs}, in order to implement a statistical test on the hypothesis

$$H_0: E[\Delta h] = E[\hat{h} - h_{obs}] = 0$$

The statistics used to perform this test is obtained by a Least Square Estimation approach for the coefficients of a polynomial that constitutes the interpolation model. For example, if the user chooses the constant function as local model (see [15]), there is only one coefficient to estimate:

$$h_{model}(x,y) = a_0$$

If every observation $h_{obs,i}$ is assumed, as independent, with normal distribution (mean a_0 and variance s_0^2), we have

$$\hat{a}_0 = \frac{1}{N_k}\sum_{i=1}^{N_k} h_{oss,i} \sim N[a_0, s_0^2 / N_k]$$

A correct estimator of s_0^2 is

$$\hat{s}_0^2 = \frac{1}{N_k - 1}\sum_{i=1}^{N_k} (h_{oss,i} - \hat{a}_0)^2 \sim \frac{s_0^2}{N_k - 1}\chi^2_{(N_k - 1)}$$

so, since

$$\Delta h = h_{oss} - \hat{a}_0 \sim N[0, s_0^2 (1+1/N_k)]$$

we have that the test is performed by

$$\sqrt{\frac{N_k}{N_k + 1}}\frac{\Delta h}{\hat{s}_0} \sim \frac{Z}{\sqrt{\frac{\chi^2_{N_k - 1}}{N_k - 1}}} = t_{(N_k - 1)}$$

Remark

Indeed real errors with respect to simple models like these are not expected to be neither independent nor normally distributed. The experience, however, says that traditional tests do tend to be conservative, i.e. to identify as outliers data that could not be so with respect to more realistic distributions. Since outliers as such are not “false” data but just data with a different statistical signature and have to be more closely analyzed, we think that the proposed procedure is still acceptable.

See [4] to read about another conservative approach to outlier detection, while the subjective nature of outlier rejection procedures is in [7].

The outlier rejection strategy

Actually, our outlier rejection is based on the following polynomials, well known in literature on approximation and reliability, that are a straightforward generalization of the above elementary example.

- Bilinear (4 coefficients):

  $h_{bil}(x,y) = a_0 + a_1 \cdot x + a_2 \cdot y + a_3 \cdot xy$
- Bicubic (16 coefficients):

  $h_{bic}(x,y) = a_0 + a_1 \cdot x + a_2 \cdot y + a_3 \cdot xy + a_4 \cdot x^2 + a_5 \cdot y^2 + a_6 \cdot x^2 y + a_7 \cdot xy^2 + a_8 \cdot x^2 y^2 +$
  $+ a_9 \cdot x^3 + a_{10} \cdot y^3 + a_{11} \cdot xy^3 + a_{12} \cdot x^3 y + a_{13} \cdot x^2 y^3 + a_{14} \cdot x^3 y^2 + a_{15} \cdot x^3 y^3$

Since it is important to keep a high overdetermination, for the first case we chose a window size of 3x3, while for the second one a size of 7x7.
The test procedure is then performed by observing that, in both cases,

$$\hat{h} = \hat{a}_0$$

Since $\hat{a}_0$, in the case of gridded data, is still given by

$$\hat{a}_0 = \frac{1}{N_k} \sum_{i=1}^{N_k} h_{oss,i}$$

for the bilinear interpolator and by a more complicated formula – coming from the Least Square Adjustment – for the bicubic interpolator, by exploiting a standard testing theory, we can write

$$\Delta h = h_{oss} - \hat{h} \sim N[0, s_0^2 (1+q)]$$

where q depends on the number of unknown parameters (in other words, the degree of the polynomials plus 1) and the size of the moving window. It can be found that for both the methods (bilinear with 3x3 and bicubic with 7x7) the value of q is 1/8.
Therefore the test statistics applies (see for instance [3])

$$\frac{h_{oss} - \hat{h}}{\hat{s}_0 (9/8)} \sim t_n$$

where t_ν denotes the student's t distribution with ν degrees of freedom (the number of observations reduced by the number of unknown parameters). For the bilinear case (with window size 3) ν=8-4=4, while for the bicubic case (window size 7) ν=48-16=32.

In table 1 there are the results, according to different significance levels α.
It shows the number of suspected outliers for the present test, to be compared with the number of suspected outliers expected because we fix a certain α risk of the first kind (for a normal distribution, given α, we expect αN data to be suspected).
The results point to a significant non normality of the data; we have just seen that in the 3x3-bilinear interpolation the redundancy (number of degrees of freedom, ν) is 4 while in the 7x7-bicubic interpolation is 32; therefore, with α=1%, the corresponding t values are 4.60 and 2.75. This explains why we find more outliers in the bicubic case.
The choice of the α value to be used (in our case α=0.1%) has been performed by considering that the suspected outliers should be in any way more than the expected, but not too many.

Method	***size***	***Alfa***	***"Outliers"*** (out of 127920 data)	***Expected "Outliers"*** (out of 127920 data)
Bicubic	7x7	5%	15061	6396
Bicubic	7x7	1%	2290	1279
Bicubic	7x7	0.1%	179	128
Bicubic	7x7	0.05%	79	64
Bilinear	3x3	1%	290	1279
Bilinear	3x3	0.1%	17	128

Table 1

Focusing on the 179 outliers found with the third option, their distribution is shown in figure 1; figure 2, instead, shows them against the corrected gradient of the signal.

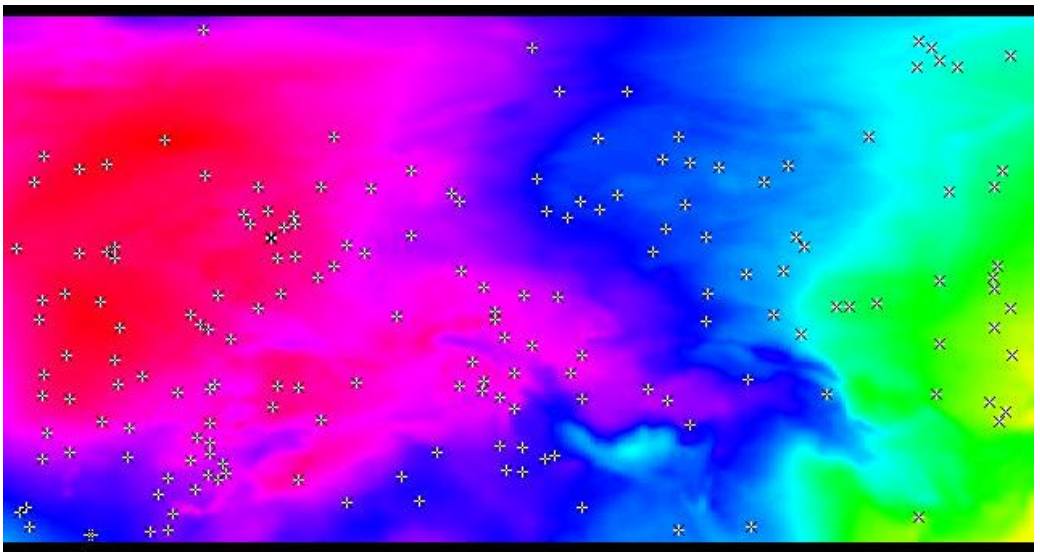

Figure 1 – 179 outliers against the QuasiGeoid signal

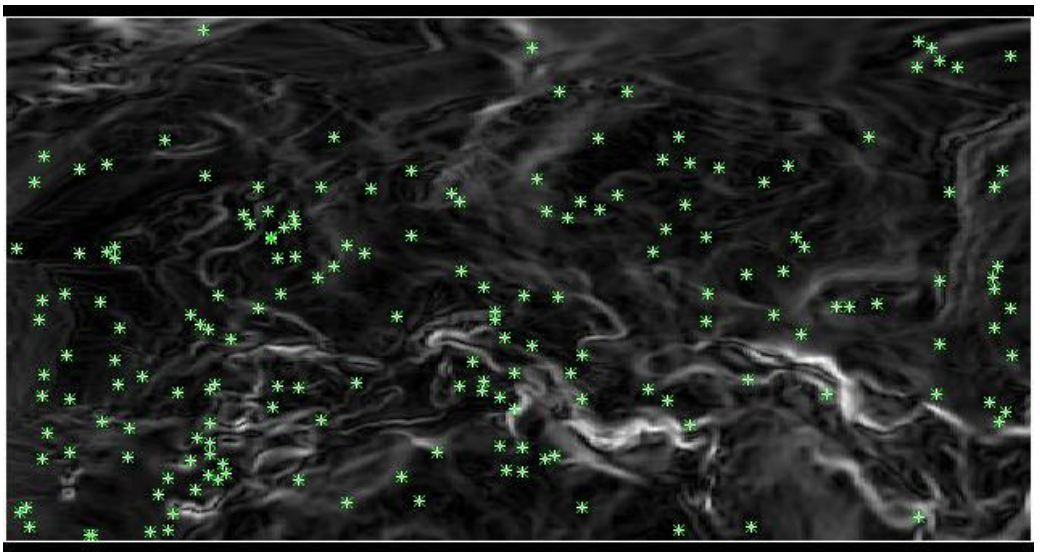

Figure 2 – 179 outliers against the gradient magnitude of the QuasiGeoid signal

It should be noted that the corrected gradient is computed by the following formula

$$\text{Grad}^*\ G(\varphi_i,\lambda_\kappa) = \sqrt{\left(G^*_{NS}\right)^2 + \left(G^*_{EW}\right)^2} \cong \sqrt{\left(G_{NS}\right)^2 + \frac{1}{\cos^2(\varphi_i)}\left(G_{EW}\right)^2}$$

where

$G(\varphi,\lambda)$ means the geoid signal evaluated in (φ,λ);

* means approximately corrected according to the local curvature of the ellipsoidal coordinate lines;

$G_{NS} = 1/(8\ \Delta\varphi)\cdot$

$\{G(\varphi_{i-1},\lambda_{\kappa-1})+2G(\varphi_{i-1},\lambda_\kappa)+G(\varphi_{i-1},\lambda_{\kappa+1})-G(\varphi_{i+1},\lambda_{\kappa-1})-2G(\varphi_i,\lambda_\kappa)-G(\varphi_{i+1},\lambda_{\kappa+1})\}$

$G_{EW} = 1/(8\ \Delta l)\cdot$

$\{G(\varphi_{i-1},\lambda_{\kappa+1})+2G(\varphi_i,\lambda_{\kappa+1})+G(\varphi_{i+1},\lambda_{\kappa+1})-G(\varphi_{i-1},\lambda_{\kappa-1})-2G(\varphi_i,\lambda_{\kappa-1})-G(\varphi_{i+1},\lambda_{\kappa-1})\}$

It's worth to investigate what kind (if any) of relationship there is between the location of the "outliers" and the gradient magnitude. At first sight, the geographical distribution of the outliers does not follow any kind of criterion based on the slope of the geoid. One way to have a confirmation of this statement is to proceed with a Montecarlo Method (for reference, [13]), by creating a suitable index that shows the general behaviour of the gradient magnitude against the behaviour of the same value when applied to the outliers. More precisely, we used the index

$$\Phi^{\mathbf{k}} = \sum_{i=1}^{\#\text{of outliers}} \log\left(1 + \left|\nabla N(P_i)\right|\right)$$

where in this case number (#) of outliers is 179, while k (the repetition index) runs from 1 to 10000 and every point P_i is sampled from a uniform distribution on the knots of the grid. As it is evident, the index increases when points have systematically higher and higher gradient magnitude. The reader could ask why we did not choose a simpler expression, such as for instance

$$\Phi^{\mathbf{k}}_{simple} = \sum_{i=1}^{\#\text{of outliers}} \left|\nabla N(P_i)\right|$$

The reason stands in the limit of the "double precision" representation of float numbers: a computer cannot store, in its memory, big numbers having an arbitrary size. Using the logarithmic mapping, instead, we reduce the amplitude of the sum, without loosing any information because an increasing monotonic map preserves the ordering of the values. More over, adding one is required to make every addendum positive. In any way, since the boot-strap method applied here works with any distribution, we can in principle use any function.

The frequency distribution of Φ is compared with the single value

$$\Phi^{*} = \sum_{i=1}^{\#\text{of outliers}} \log\left(1 + \left|\nabla N(P_i^{*})\right|\right)$$

where $\{P^{*}_i\}$ denotes the set of points where the outliers were found.

As it can be seen in figures 3a, 3b, 3c, 3d, the index Φ^{*} does not seem to be correlated with the gradient magnitude (in figures 3c and 3d it is close to the average of the empirical distribution) or slightly correlated, but in the opposite sense we expected (in figures 3a and 3b it is on the left, as if a smooth signal could lead to find more outliers).

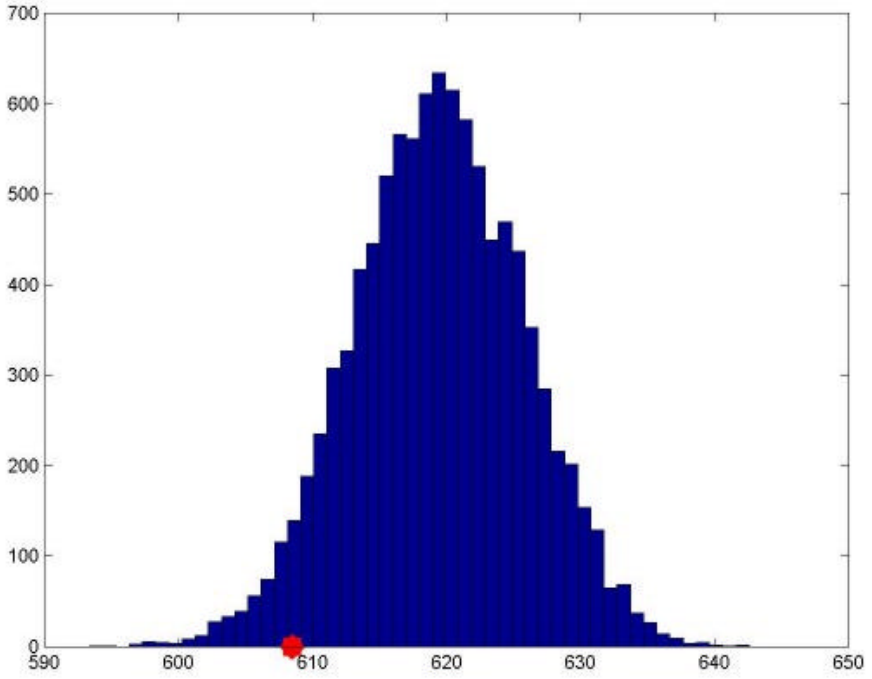

Figure 3a – Φ^* with: bicubic interpolation, 7x7 window size, significance level 0.05%

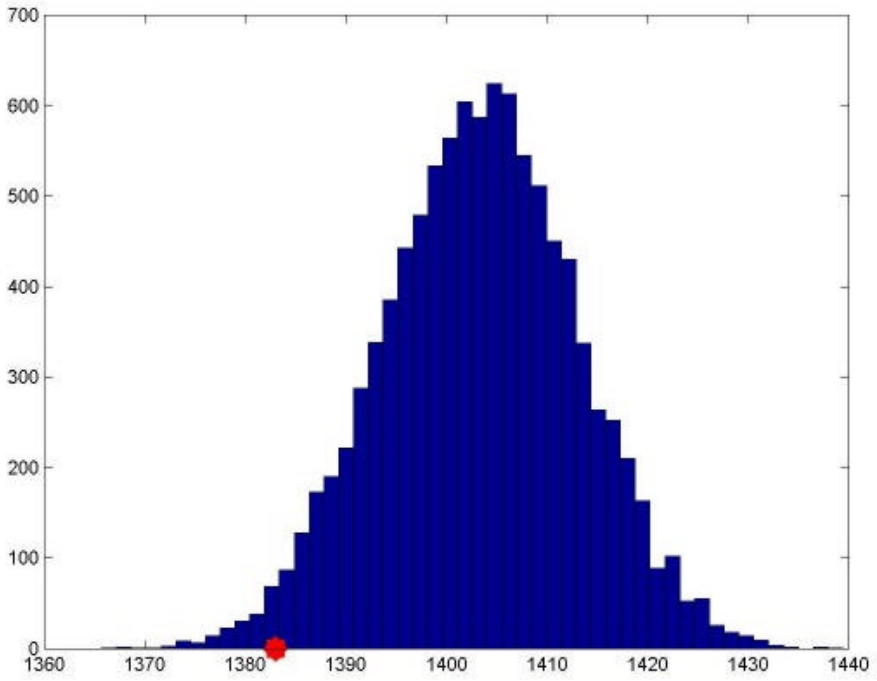

Figure 3b – Φ^* with: bicubic interpolation, 7x7 window size, significance level 0.1%

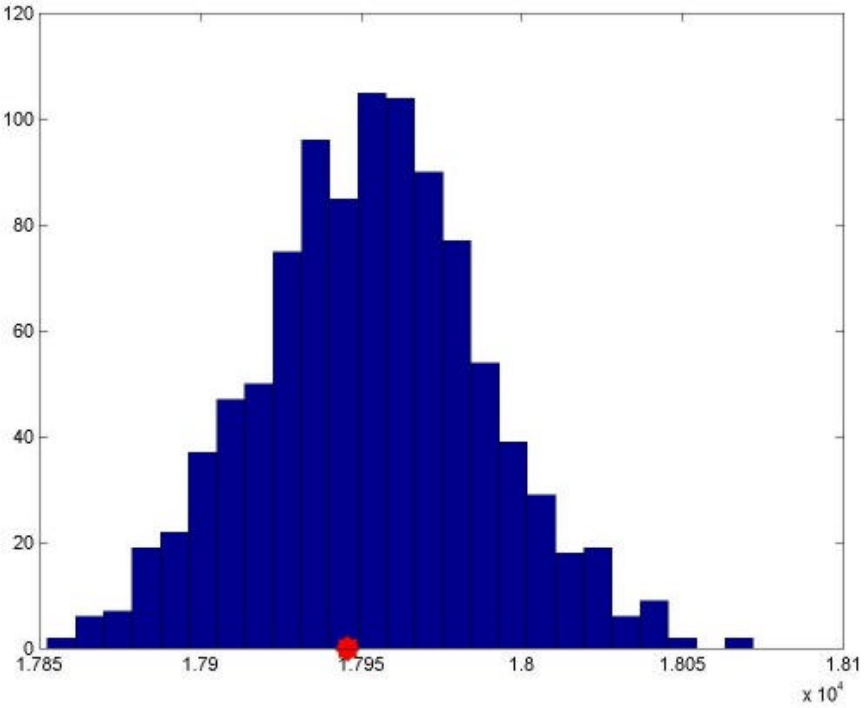

Figure 3c – Φ^* with: bicubic interpolation, 7x7 window size, significance level 5%

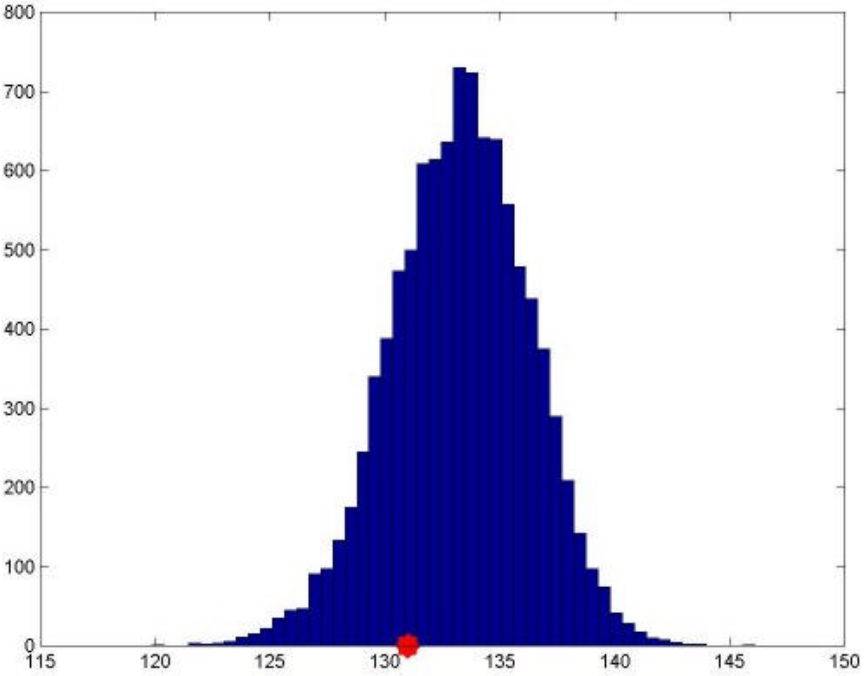

Figure 3d – Φ^* with: bilinear interpolation, 3x3 window size, significance level 0.1%

Going deeper into the rejection criterion

This apparent lack of correlation, contrary to our intuition, can be studied more deeply. In fact we suspect that our "outliers" fall into two cathegories; in one we have outliers with large absolute residuals (we can call these "relevant outliers"), which we can consider as true outliers, in the other we have outliers just because the $\hat{s}_0$ resulting from the adjustment of the bicubic surface to the surrounding values is too small.
Then we decided that we were more interested in looking into the relation between relevant outliers and inclination of the geoid. So, after inspecting the histogram of the "modulus" of the outliers (see figure 4), we decided to compare the distribution of the Φ index with the value of Φ for the most relevant ones. More precisely, we compared the value of Φ for the 10 "extreme outliers" (characterized by largest absolute residuals) with the distribution of Φ for 10 points randomly selected, on a 10000 sample.
The result is shown in figure 5, where finally the exceptionality of the index Φ for the outliers is quite evident.

This has convinced us that the pure statistical rejection criterion should be modified; as a matter of fact the other 169 outliers have a smaller residual absolute value and they are seen as outliers by the testing procedure only because the bicubic model was particularly good for that specific window.
This is also clarified by looking (figure 6) at the distribution of $\hat{s}_0$ on the 179 suspected where a relative increase of the frequencies is verified above the threshold of 15 cm and in fact our extreme outliers are all above 20 cm.
All that points towards the need of implementing a new s_0^2 estimator based on an average value in the area. In any way, in order to confirm our guess that the 10 "largest" outliers are really related to a physical phenomenon, we have investigated also their connection to the gradient of the undelying topographic surface, giving some graphical representations in figures 7, 8, 9, 10.

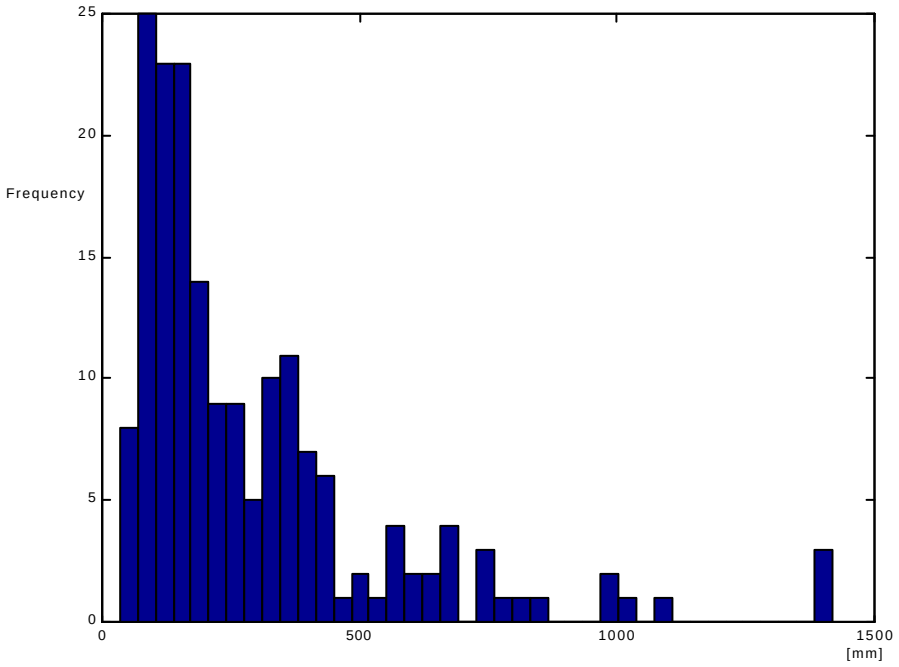

Figure 4 – Histogram of absolute residuals limited to the 179 points classified as outlier

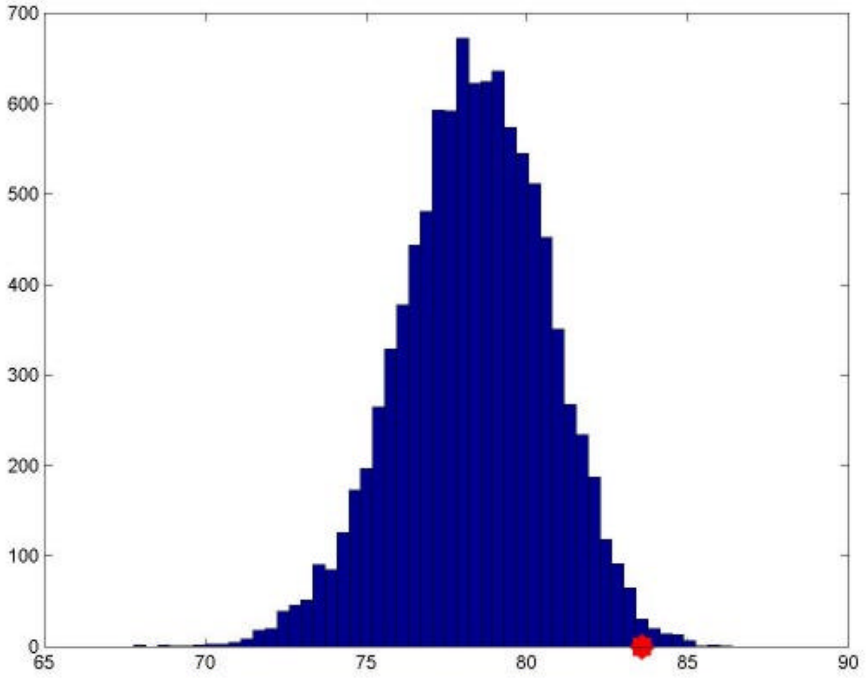

Figure 5 – $\mathbf{F}^*$ for the 10 extreme outliers

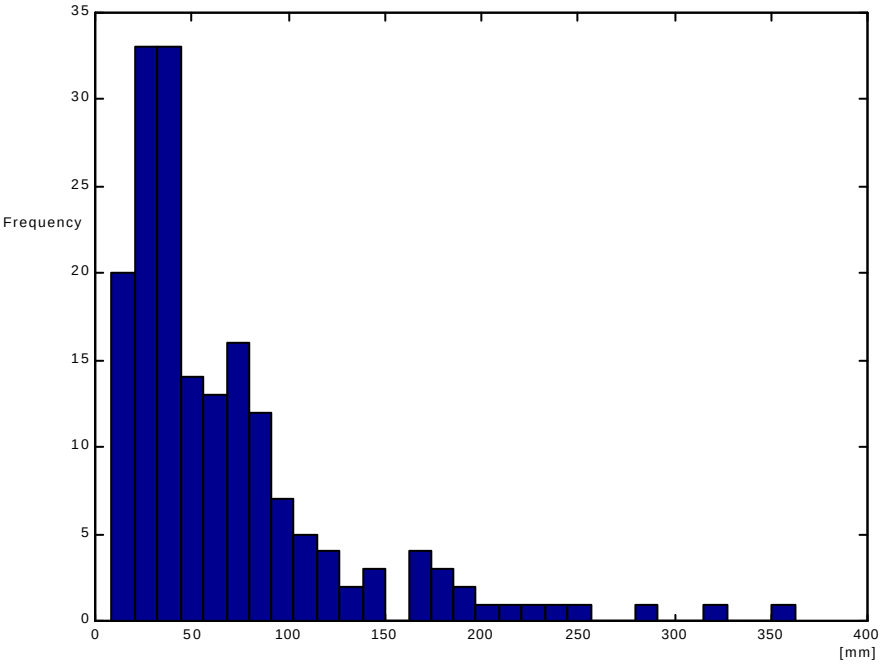

Figure 6 – Histogram of $\hat{s}_0$ limited to the 179 points classified as outlier

Conclusions

Our first conclusion is that the software implemented for the outlier rejection has to be improved, regarding to the s_0^2 estimation, which has to be estimated more specifically to avoid the detection of an unreasonable number of "irrelevant outliers". By the way, the Monte Carlo method, proposed to compare pointwise quantities (like outliers) with area distributed fields, seems to give significant results.
In particular, the relation, between the relevant outliers and large values of geoid gradients, is well illustrated in figure 8: most of the relevant outliers fall in areas characterized by high gradient magnitude and a complicated pattern.
Another conclusion is that these outliers do not seem to be so much related to mountainous areas, but rather to coastal areas with sudden jumps of sea floor. This points towards a good functioning of our terrain correction algorithm in rugged areas, but a poor functioning in coastal areas with depth contrasts, as anybody who has computed a gravimetric geoid knows by experience.

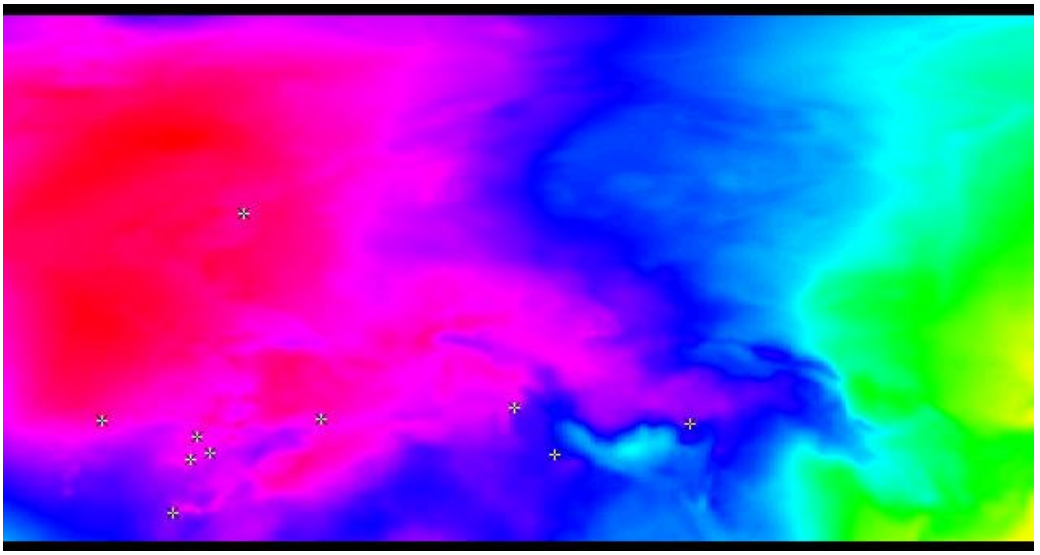

Figure 7 – 10 outliers against the QuasiGeoid signal

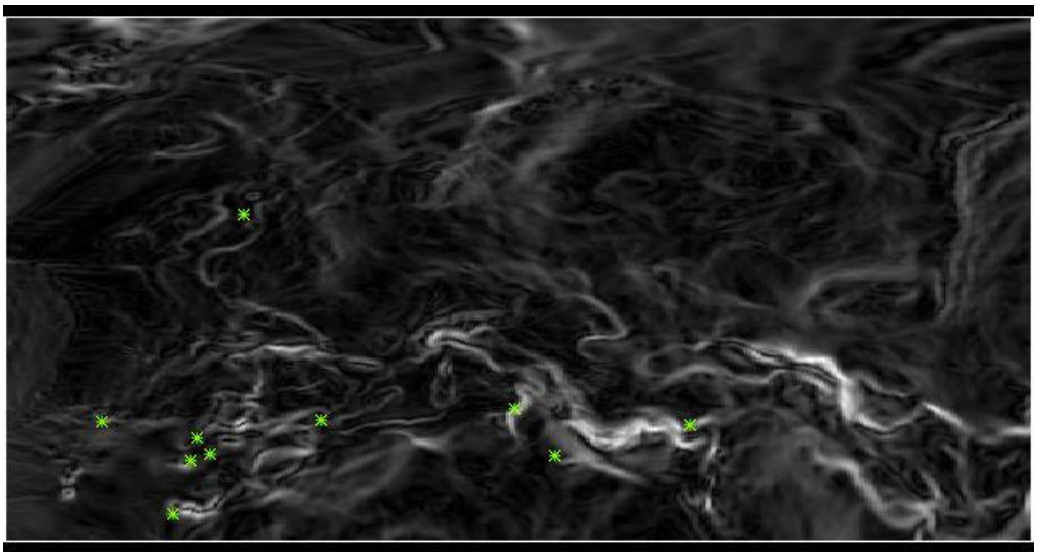

Figure 8 – 10 outliers against the gradient magnitude of the QuasiGeoid signal

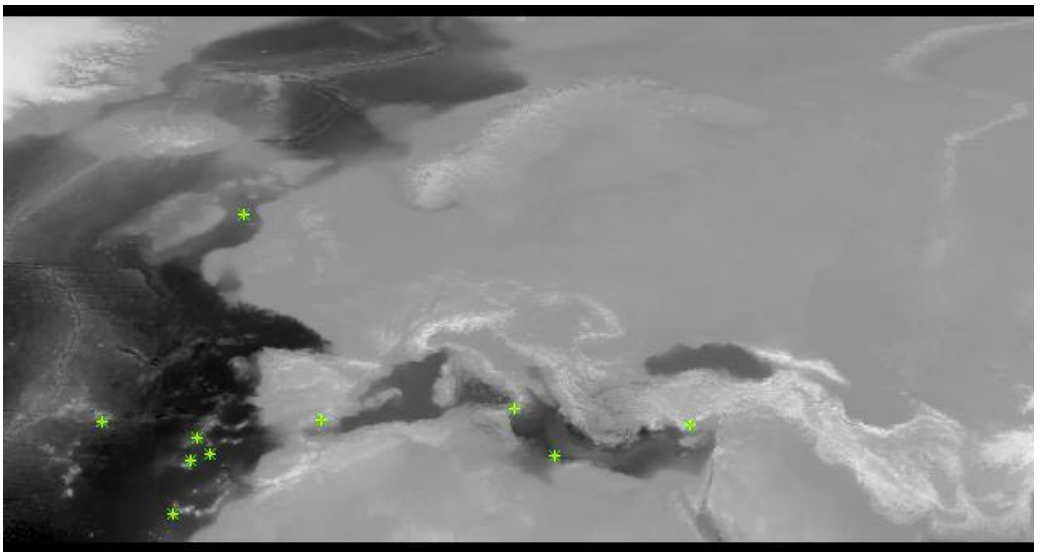

Figure 9 – 10 outliers against the Topographic heights

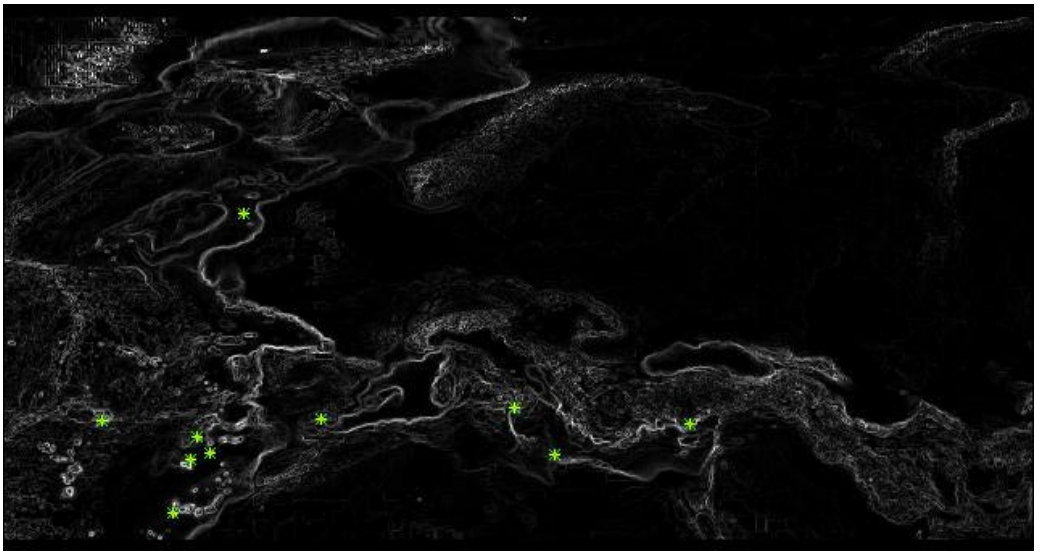

Figure 10 – 10 outliers against the gradient magnitude of the Topographic heights

Bibliography

[1] M. Barbarella, L. Mussio (1985), A strategy for Identification of Outliers in Geodetic Sciences, Statistic and Decisions, Supplement Issue n.2, R. Oldenbourg Verlag, Monaco.

[2] V. Barnet and T. Lewis (1978), Outliers in statistical data, John Wiley and Sons, New York.

[3] T. Bašiæ, R.H. Rapp (1992), Oceanwide prediction of gravity anomalies and sea surface heights using Geos-3, Seasat, and Geosat altimeter data and ETOPO5U bathymetric data. Dept. Geod. Sci. aNd Surv., Rep. 416, Columbus.

[4] B. Benciolini, L. Mussio, F. Sansò (1982), An Approach to Gross Errors Detection more Conservative than Baarda Snooping", Int. Archives of Photogrammetry and Remote Sensing, Vol. 24, Helsinki.

[5] P.A. Burroghs (1986), Principles of Geographic Information Systems for land resources assesment, Oxford University Monographs on Soil and Resources Survey, n. 12, Clarendon Press, Oxford.

[6] G. Casella (1990), R.L. Berger, Statistical Inference, Wadsworth & Brooks, Pacific Grove.

[7] D. Collett, T. Lewis (1976), The subjective nature of outlier rejection procedures, Applied Statistics, 25, 228-237.

[8] N. Cressie (1991), Statistics for Spatial Data, Wiley, New York.

[9] H. Denker, D.Behrend, W.Torge (1994), European Gravimetric Geoid: Status report 1994, Proceed. IAG Symp. No. 113, Gravity and Geoid, Graz, Austria, Sept. 11-17, 1994, Springer.

[10] H. Denker, D.Behrend and W.Torge (1995), The European Gravimetric Quasigeoid EGG95, International Association of Geodesy, IGeS Bulletin 4

[11] W.J. Dixon (1953), Processing data for outliers, Biometrics.

[12] J.H. Ellenberg (1973), The joint distribution of the standardized least squares residuals from a general linear regression, Journal of the American Statistical Association, 68, No.344, 941-3.

[13] D. Gamerman (1997), Markov Chain Monte Carlo, Chapman & Hall.

[14] D.M. Hawkins (1980), Identification of Outliers, Chapman and Hall.

[15] K. Koch (1987), Parameter Estimation and Hypothesis Testing in Linear Models, Springer-Verlag.

[16] G.S. Lingappaiah (1976), Effects of outliers on the estimation of parameters, Metrika.

[17] R.B. Murphy (1951), On tests for outlying observations, PhD thesis, Princeton University.

[18] A.J. Pope (1975), The Statistics of Residuals and the Detection of Outliers, presented paper at IUGG XVI General Assembly, Grenoble.

[19] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery (1996), Numerical Recipes in C, Cambridge University Press.

[20] C.P. Robert and G.Casella (1999), Monte Carlo Statistical Method, Springer.

[21] L. Sachs (1984), Applied Statistic – A Handbook of Techniques, Second Edition, Springer-Verlag, New York.

[22] F. Sansò (1988), Il trattamento statistico delle misure, CLUP, Milano.

[23] F. Sansò (1991), Il trattamento statistico dei dati, CittàStudi.

[24] F. Sansò (1996), Quaderni di trattamento statistico dei dati, CittàStudiEdizioni.

[25] M.A. Tanner (1996), Tools for Statistical Inference, Springer.

[26] W. Torge, G.Weber, H.G. Wenzel (1982), Computation of a high resolution European gravimetric geoid. Proc. 2nd Int. Symp. On the Geoid in Europe and Mediterranean Area, Rome.

[27] W. Torge and H. Denker (1990), Possible Improvements of the Existing European Geoid, Proceed. IAG Symp. No. 106, Determination of the Geoid - Present and Future, Milan, Italy, June 11-13, 1990, Springer.

[28] H.Wackernagel (1995), Multivariate geostatistics, Springe-Verlag, New York.

[29] V. Vapnik (1982), Estimation of Dependences Based on Empirical Data, Springer-Verlag, New York.

An enhanced method for validating gridded data sets

Damiano Triglione
Politecnico di Milano, Polo Regionale di Como
damiano.triglione@polimi.it

Abstract
The purpose of this work is to propose a reliable method to detect outliers, continuing the work done in a previous article. The previous work was affected by a serious defect related to a rather unstable local estimate of a dispersion index. This paper relies on a robust estimation of a relevant parameter that, before being used for testing the data, is processed with a low-pass filter. The many advantages, obtained in this way, are highlighted in an experiment on a real D.T.M. dataset.

Introduction

The adopted method for outlier rejection in a gridded points data set is the following one: to validate the value $u(t_k)$, a window [whose dimension is $(2s+1)\delta_x$ x $(2s+1)\delta_y$] is opened around the position t_k; the observation points t_j falling in the window [whose number is $N_s = (2s+1)^2 - 1 = 4s\,(s+1)$] are considered in order to estimate the coefficients of an interpolating model. With the interpolating model (which, we underline, is tuned only on the surrounding points), an estimated value of $u(t_k)$ is compared with the observed $u(t_k)$ itself, leading to test the hypothesis (with significance level α)

$$H_0: E[\hat{u}(t_k) - u(t_k)] = 0 \quad .$$

If the local model is a constant function, there is only one coefficient to estimate: $\hat{u}(t_k)$ itself. In the past we proposed fitting polinomials with coefficients estimated by a Minimum Least Square criterion. As it is known (see [18]), we found that these methods have many disadvantages; in particular they are unreliable when there are more outliers in the window and in certain cases they are too much sensitive.
The enhanced method we propose now is robust against outliers, because we use

$$\hat{u}(t_k) = \underset{j \neq k}{Median} \{u(t_j)\} \quad ;$$

therefore t_k is considered an outlier if

$$|Z_{emp}(t_k)| > Z_{\alpha/2} \quad ,$$

where

$$Z_{emp}(t_k) = \frac{u(t_k) - \hat{u}(t_k)}{\sqrt{\left(1 + \frac{\pi}{2N_s}\right)\frac{\pi}{2}} \cdot MAE_k} \quad ,$$

$$MAE_k = \frac{1}{N_s} \sum_{j \neq k} \left| u(t_j) - \hat{u}(t_k) \right| \quad ,$$

while $Z_{\alpha/2}$ is the abscissa value of the standard gaussian distribution, corresponding to a right tail of area $\alpha/2$.[1]

The estimated σ of $[\hat{u}(t_k) - u(t_k)]$ is typical of robust estimation literature (cfr. [8], [6] and [17]). In our conservative approach the normal (gaussian) approximation is used, but – as discussed in [18] – this does not represent a problem.

It should be noted that $Z_{emp}(t_k)$, depends on the size of the moving window. Therefore an algorhytm to find an acceptable dimension for the windows has been developed, too. The size of the window depends essentialy on *s*, since δ_x and δ_y are led by the kind of acquisition process and the type of the cartographic projection of the area to which the points belong.
A good choice of *s* can be made by analyzing the quality of the interpolation model on a global scale. We performed this analysis with the index "global MAE":

$$\text{gMAE}(s) = \frac{1}{N}\sum_k \left|u(t_k) - \hat{u}(t_k)\right| \quad .$$

Its behaviour (that depends on *s* since $\hat{u}(t_k)$ comes from the median of values that are enclosed in the moving window) is this: starting from *s*=1, gMAE decreases until, for a value *s*=*s**, it attains the minimum value gMAE_{min}; for *s*>*s**, gMAE increases indefinitely with *s*. Therefore this optimal value *s** can also be used for the computation of every local MAE_k.
Though we are not able to justify this result, we found experimentally that, when data are regularly gridded, then *s**=1 has always been the optimal value (gMAE is minimized for a window of size 3x3).

Inadequacy of the simple local MAE calculated as above
When the points t_j, around t_k , are well fitted by their median (better than the mean index for not being "outlier-prone" - see also [1], [2] and [3]), then a small difference between $u(t_k)$ and its estimation leads to consider t_k an outlier.
For instance, consider 8 points in a grid with exactly the same height and a central point with 5 cm height difference; of course, since the MAE in this case is exactly zero, this point will be flagged as an outlier. We notice already that this can happen exactly because here and there 8 points can really bear similar values; when we increase the number of surrounding points, this effect could become milder.
In any way the formula for outlier rejection needs a revision, in order to label as outliers only points whose values are sensibly different from their estimations.
Looking back at

$$Z_{emp}(t_k) = \frac{\left|\hat{u}(t_k) - u(t_k)\right|}{\sqrt{\left(1+\frac{\pi}{2N_s}\right)\frac{\pi}{2}}\cdot MAE_k} \quad ,$$

we find that, to decrease $Z_{emp}(t_k)$, under equal conditions, the value of local MAE_k should be increased, avoiding critical situations like that described above. This

[1] A partial demonstration of the above formulae is the asymptotic behaviour of Z_{emp}: when $N\rightarrow\infty$, Mean Absolute Error (MAE) is close to Standard Deviation (σ) and Median (MD) is close to Mean (μ). So, having defined $y=u(t_k)-MD$, we discover that $y=N[0,\sigma^2]$, $\mu[|y|]= (2/\pi)^{1/2}\,\sigma$ and therefore $\frac{y}{\sqrt{\frac{\pi}{2}}MAE_k} \approx Z$.

operation must be accomplished with caution, since the rule that is going to be introduced has to be valid under fairly general conditions, not case by case.
We decided that a suitable way to "modulate" the local MAE_k is to smooth it, as if it were a signal. From the general theory of signals, when a signal is smoothed its minimum increases, its maximum decreases, its mean is preserved and the variance decreseases. Moreover, these behaviours are stressed as far as the size of the moving average window increases.

The smoothing strategy
A simple way to smooth our "MAE signal" is using a traditional bi-dimensional moving average on a window of size (2*s*+1)x(2*s*+1). It is known that when a matrix **M** (nr)x(nc) is convolved with a kernel **K** (2*s*+1)x(2*s*+1), the result is a matrix **R** (nr+2*s*)x(nc+2*s*) because **M** is padded with a frame of zeros for s times. But the significant values of the operation are in the submatrix **S** (nr-*s*)x(nc-*s*) of **M**, since only the most internal values are calculated with all the elements covered by **K**. Naturally the choice of the dimension *s* in this case is very important.
In the next paragraph we show some experiments with real data from a D.T.M. that have given us some hints on this choice.
On the other hand at least two two things are clear: altough an optimal window to estimate the central value is often of dimension 3x3, yet the corresponding MAE (computed on that window only) has a by far too large probability of having a too small value, thus generating false outliers. On the other side, if we used the average of the local MAEs over the whole data area that would be too large to detect real outliers. So we decided to base our choice on experiments.

The experiment
In order to test the quality of the method of outlier rejection and to calibrate the one on which local MAEs have to be averaged, we chose an area (in the northern part of Italy) where we have a Digital Terrain Model characterized by a flat zone and a mountainous one. There is also a lake (Lake of Como) whose form is a "Y" upside down. The description of the region is the following (in degrees with their decimals):

```
north:      46.1326129
south:      45.491
east:       9.2928936
west:       9.043
e-w resol:  0.00277659556
n-s resol:  0.00194428152
```

therefore the matrix M is composed by 330 rows and 90 columns, as shown in figure 1.

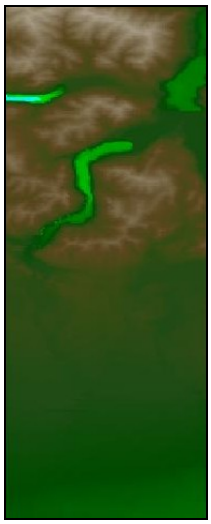

Figure 1 – The area of study

An outlier rejection performed with our method on the above dataset (with significance level of α=0.1%) and using local MAE indexes without any smoothing, leads to identifying 40 outliers.
It should be noted that, in order to avoid border effects, a frame of width 9 all around the boundary has been dropped in every case because the largest kernel is 19x19 (so s=9) and we wanted to have results capable of overlapping.
So instead of validating 330x90=29700 points, we effectively validated 312x72=22464 points.

By smoothing the local MAE with a simple kernel 3x3, the number of outliers found is 1; this shows that smoothing in a so restricted area provoked an increasing of the local MAE (in every point) and therefore to consider good points the ones previously labelled as outliers. Using wider and wider kernels, the number of found outliers goes around 50 (as can be seen in the table 1). This can lead to think that this is the reasonable number of outliers.

MAE smoothing's Window	# Outliers	Min MAE	Max MAE	Mean MAE	Var MAE
None	40	0	210,88	32,69	1306,8
3x3	1	0	162,06	32,69	1167,7
7x7	30	0	117,12	32,69	1023,4
11x11	46	0,535	101,665	32,68	945,9
15x15	48	0,551	94,96	32,64	900,1

Table 1 – Interesting indicators with different smoothing windows

Another way of perceive the process is by plotting an histogram of the entire MAE dataset in the different cases.
In figure 2 it can be seen that – smoothing with a wider and wider mask – the number of very small MAEs decreases while the number of MAE around the mean increases.

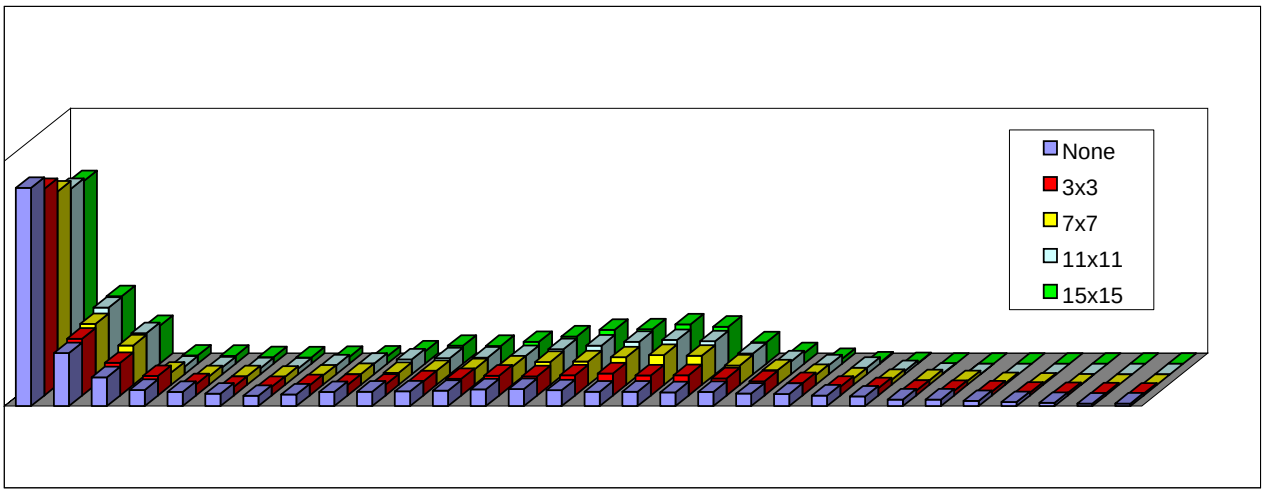

Figure 2 – Histogram of MAE dataset in different cases

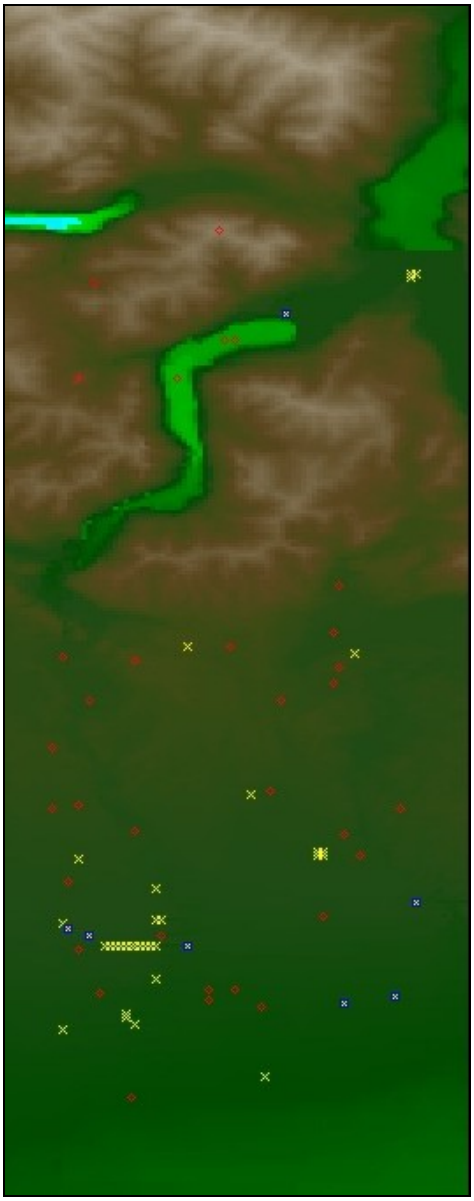

Figure 3 – Relevant points according to two approaches

As it can be seen from Figure 3 (where outliers according to original MAE are plotted with red asterisks, outliers according to mask 9x9 with smoothed MAE are plotted with yellow, and the 7 outliers in common with blue boxes), by smoothing the MAE, we are able to find "true" outliers (for example very interesting the row of 11 outliers in the south west, which gives the impression of some systematic error) and ignore many "false" ones, for example those that are in the flat zone.

The smoothing process of the local MAEs produces a positive effect that comes from the combination of two effects: the increasing of the minimum MAE (that leads to consider good points the ones that with calculated MAE were considered outlier, because of their small deviation from the estimation) and the decreasing of the maximum MAE (that leads to consider outlier points some that previously were considered good ones). The resulting effect is that a reasonable number of outliers is detected by a reliable process. This insight is confirmed by figure 4 (and figure 5, that is a zoom of figure 4), where – in a diagram local MAE vs absolute differences – with red circles (o) are shown outliers according to original MAE, while with blue crosses (x) are shown outliers according to smoothed MAE with mask 9x9.
Indeed, circles with very low MAEs and very low absolute differences[2] are ignored by smoothing the MAE; also circles with very high local MAE and very high absolute differences are not considered outlier, according to the enhanced process. The 7 outliers in common are perceivable because each of the 7 circles has its own cross on the same value of absolute difference.
It can also be noted that, in a gridded dataset, it is always $N_s=8$. So, having fixed $\alpha=0.1\%$, is easy to calculate the gradient of the dotted line that defines the limit over which we find outliers: since the theoretic value of Z is 2.5758,

$$\frac{\left|\hat{u}(t_k)-u(t_k)\right|}{\sqrt{\left(1+\frac{\pi}{16}\right)\frac{\pi}{2}}\cdot MAE_k} > Z_{\alpha/2} \Leftrightarrow \left|\hat{u}(t_k)-u(t_k)\right| > 3.53118\, MAE_k \quad .$$

In conclusion, the simple idea of smoothing the MAE is a powerful tool to detect outliers as described above. The smoothing window can be fixed to dimension 9x9 altough close by values would not dramatically affect the result. There is still work left to conceive a more general method suitable for not gridded datasets.

[2] With the term "absolute difference" we mean the absolute value resulting from the difference between the value of the observed point and its estimation (the median of the surrounding points).

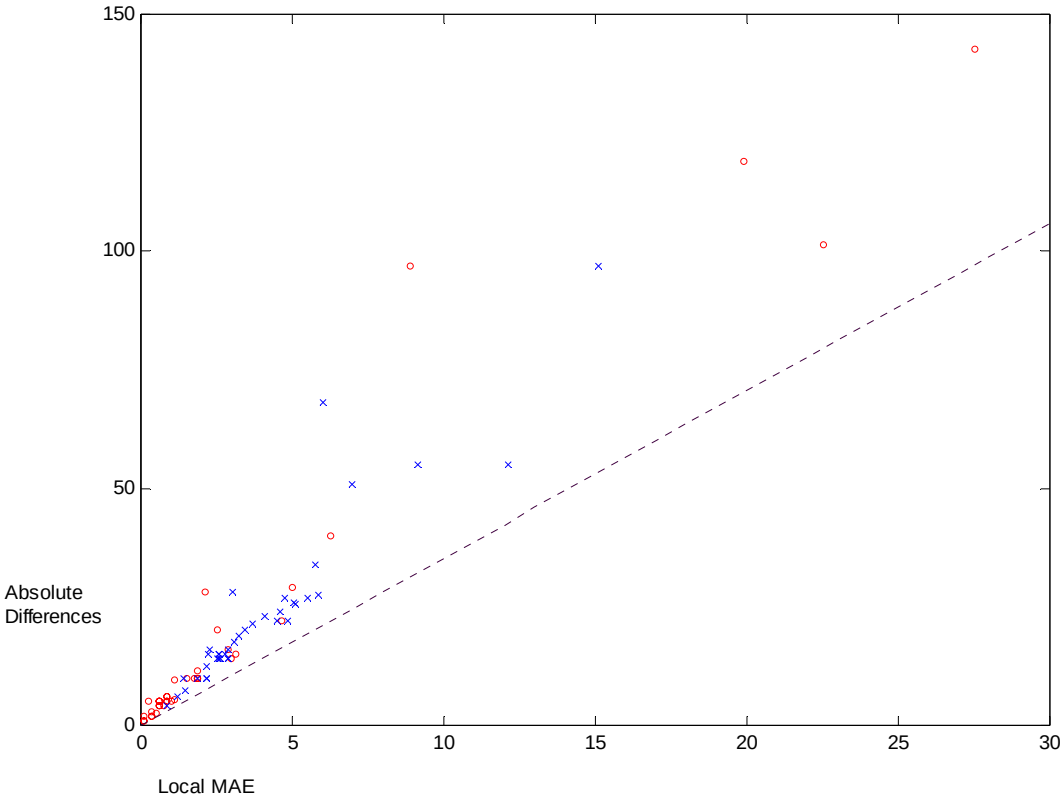

Figure 4 – Relevant points according to two approaches

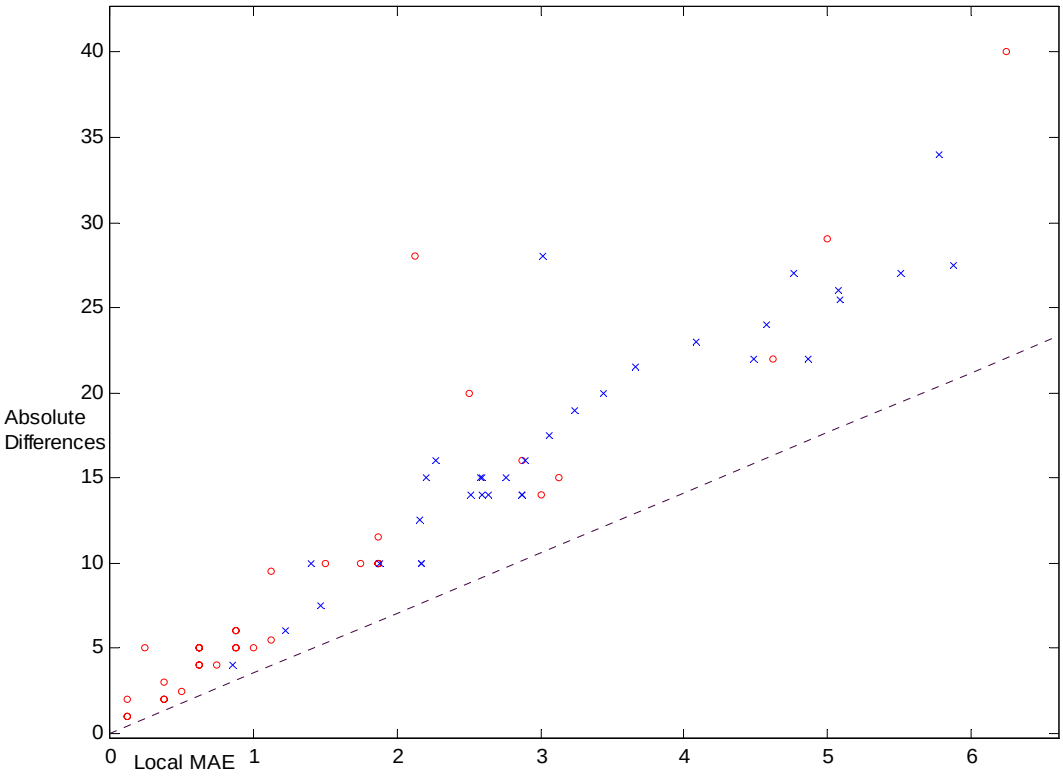

Figure 5 – Particular of figure 4

Bibliography

[1] M. Barbarella, L. Mussio (1985), A strategy for Identification of Outliers in Geodetic Sciences, Statistic and Decisions, Supplement Issue n.2, R. Oldenbourg Verlag, Monaco.

[2] V. Barnet and T. Lewis (1978), Outliers in statistical data, John Wiley and Sons, New York.

[3] B. Benciolini, L. Mussio, F. Sansò (1982), An Approach to Gross Errors Detection more Conservative than Baarda Snooping", Int. Archives of Photogrammetry and Remote Sensing, Vol. 24, Helsinki.

[4] G. Casella (1990), R.L. Berger, Statistical Inference, Wadsworth & Brooks

[5] D. Collett, T. Lewis (1976), The subjective nature of outlier rejection procedures, Applied Statistics, 25, 228-237.

[6] H.A. David and H.N. Nagaraja (2003), Order Statistics, Third Edition, Wiley.

[7] A. Dermanis, A. Grün and F. Sansò (2000), Geomatic Methods for the Analysis of Data in the Earth Sciences, Springer – Verlag.

[8] E.R. Dougherty and J. Astola (1994), An Introduction to Nonlinear Image Processing, SPIE Optical Engineering Press.

[9] R.O. Duda and P.E. Hart (1973), Pattern classification and scene analysis, John Wiley & Sons.

[10] D.M. Hawkins (1980), Identification of Outliers, Chapman and Hall.

[11] A. Jain (1989), Foundamentals of digital image processing, Prentice Hall.

[12] L. Kaufman and P.J. Rousseeuw (1990), Finding Groups in Data. An Introduction to Cluster Analysis, John Wiley & Sons.

[13] K. Koch (1987), Parameter Estimation and Hypothesis Testing in Linear Models, Springer-Verlag.

[14] J.S. Lim (1990), Two-dimensional Signal and Image Processing, PTR Prentice Hall.

[15] A.J. Pope (1975), The Statistics of Residuals and the Detection of Outliers, presented paper at IUGG XVI General Assembly, Grenoble.

[16] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery (1996), Numerical Recipes in C, Cambridge University Press.

[17] L. Sachs (1984), Applied Statistics – A Handbook of Techniques, Second Edition, Springer-Verlag, New York.

[18] D. Triglione, Reasoning about the presence of outliers in the European quasi-geoid data set, Same Number, In print.

A COMPUTER PROGRAM FOR AN ADJUSTMENT OF COMBINED GPS/LEVELLING/GEOID NETWORKS: CASE OF STUDY: NORTH OF ALGERIA

BENAHMED DAHO Sid Ahmed (National Centre of Space Techniques, Geodetic Laboratory - BP 13 Arzew- 31200- Algeria. Phone: 00 213-41-47-22-17, Fax : 00 213-41-47-36-65, Email : d_benahmed@hotmail.com)

Abstract: The combined use of global positioning system (GPS), levelling and geoid height information has been a key procedure in various geodetic applications and provides at the same time a very attractive evaluation scheme for the accuracy of gravimetric geoid models. The main goal of this paper is to propose a tool of adjustment of combined GPS/Levelling/Geoid networks. For this kind of adjustment, generally, two main types of unknowns are estimated; the gravimetric geoid accuracy and 2D spatial field that describe all the datum/systematic distortions between the available height data sets. Accordingly, two modelling alternatives for the correction field are programmed, namely a pure deterministic parametric model, and a hybrid deterministic and stochastic model.

In this context, and for a first attempt, a Fortran program has been developed translating in application this methodology using the Hirvonen model as analytic covariance function of the reduced signals and the four parameter model to describe all possible datum inconsistencies and others systematic effects between the available height data sets. The necessary data used to test this program, the adopted methodology, the computation procedure, and the obtained results will be presented.

Key words: Covariance function, Hybrid deterministic and stochastic model, levelling.

1. Introduction

As an important result of development in high technology, satellite based positioning system has become to use in geodesy and surveying professions. These developments made the measurement works more accurate, more practical and more economic.

In practice, a high accurate gravimetrically determined geoid is often computed by using the technical Remove-Restore. Such geoid can have a very high resolution and very high relative accuracy in the sense of the difference of the geoidal height. However, its absolute accuracy in the sense of the determined geoidal height itself, is currently very poor due to the systematic errors caused by the difference in reference systems, the long wave-length errors of the geopotential model used as reference in the computation of the geoid, the biases existing in the gravity data and in the digital terrain model (DTM), etc.

In the other hand, we can now measure by means of the space techniques, on land through a combination of GPS positioning and precise Levelling and at sea through satellite altimetry, the geoid on some points on the earth's surface with very high absolute and relative accuracy, especially when the GPS coordinates have been attached directly to VLBI or SLR stations.

Consequently, the gravity solution, which has very high resolution and relative accuracy but poor absolute accuracy, and the GPS levelling solution, which has poor resolution but very high accuracy can be combined in the same adjustment (Jiang & al., 1996).

However, the fitting the gravimetric geoid to a set of GPS levelling points by using the Least Squares adjustment permits to estimate the residuals v_i which are traditionally taken as the final external indication of the network accuracy. The main problem under this method is that the v_i terms will

contain a combined amount of GPS, levelling and geoid random errors that need to be separated into individual components for a more reliable geoid assessment.

The objective of this paper is to propose a tool of adjustment of combined GPS/ Levelling/ Geoid networks. For this purpose, two modelling alternatives for the correction field are programmed namely a pure deterministic parametric model, and a hybrid deterministic and stochastic model (Kotsakis & al. 1999). The program developed in this framework allows to estimate the gravimetric geoid accuracy and to compute the 2D spatial field parameters that describe all the datum/systematic distortions between the available height data sets.

2. Mathematical model for the combined adjustment

In this section, the theory of general adjustment model will be reviewed briefly in order to describe the methodology adopted in the setting of this work. For detailed aspects of combined adjustment of different heights data sets can be found in Kotsakis & al. (1999).

Let us assume that at each point P_i of a test network composed the m points, we have a triplet of height observations (h_i, H_i, N_i), or equivalently one synthetic observation:

$$\mathbf{l_i} = \mathbf{h_i} - \mathbf{H_i} - \mathbf{N_i} = \mathbf{a_i^T}.\mathbf{x} + \mathbf{s_i} + \mathbf{v_i^h} - \mathbf{v_i^H} - \mathbf{v_i^N} \qquad \textbf{(1)}$$

where $\mathbf{h_i}$, $\mathbf{H_i}$ and $\mathbf{N_i}$ denote the available observed values for the GPS, orthometric and geoid height respectively. The $\mathbf{v_i}$ terms describe the zero mean random errors, for which a second-order stochastic model is available:

$$\mathbf{E}\{\mathbf{v_h}\mathbf{v_h^T}\} = \mathbf{C_h} = \sigma_h^2\mathbf{Q_h}\ ,\ \mathbf{E}\{\mathbf{v_H}\mathbf{v_H^T}\} = \mathbf{C_H} = \sigma_H^2\mathbf{Q_H} \quad \text{and} \quad \mathbf{E}\{\mathbf{v_N}\mathbf{v_N^T}\} = \mathbf{C_N} = \sigma_N^2\mathbf{Q_N} \qquad \textbf{(2)}$$

and Q_h, Q_H and Q_N denote the cofactor matrices.

For the orthometric height, the covariance matrix $\mathbf{C_H}$ is determined from the adjustment of the levelling network. In the same way, $\mathbf{C_h}$ can be computed from the adjustment of the GPS surveys performed at the levelled benchmarks. However, the covariance matrix C_N is determined by means of the error propagation from the original noisy data used in the geoid solution (Kotsakis & al., 1999).

By means the matrix notation, the equation **(1)** can be written under the form:

$$\mathbf{L} = \mathbf{AX} + \mathbf{s} + \mathbf{Bv} \qquad \textbf{(3)}$$

where $l = [l_1 \ldots\ldots l_i \ldots\ldots l_m]^T$; vector of measured quantities set.

$\mathbf{s} = [\mathbf{s_1} \ldots\ldots \mathbf{s_i} \ldots\ldots \mathbf{s_m}]^T$; vector of the signals.

$\mathbf{v} = [\mathbf{v_h^T}\ \ \mathbf{v_H^T}\ \ \mathbf{v_N^T}]^T$; residual random noise. $\mathbf{B} = [\mathbf{I_m}\ \ -\mathbf{I_m}\ -\mathbf{I_m}]^T$; I_m : m x m unit matrix.

A is a given (m x n) matrix expressing the effect of the parameters **X** on the observation $\mathbf{l_i}$; it is sometimes called ″ sensitivity matrix″. The expression **AX** is usually obtained by linearizing an originally non-linear function of the (n) parameters; it corresponds in our case to all necessary reductions that need to be applied to the original data in order to eliminate the datum inconsistencies and other systematic errors in heights data sets. Another function s, the ″signal″, irregularly oscillating

about zero; it is assumed that this quantity has an expected value equal to zero. Finally, X is a (n x 1) vector of unknown non-random parameters.

The problem is to determine this curve AX + s by means of discrete observation l, which are furthermore affected by observational errors v. It clearly appears that the adopted general adjustment model is analogous to collocation model with parameters; it is a synthesis between adjustment and prediction.

The solution of the general adjustment model (3) satisfy the minimum condition:

$$\mathbf{s}^T\mathbf{Q}_s^{-1}\mathbf{s} + \mathbf{v}_h^T\mathbf{Q}_h^{-1}\mathbf{v}_h + \mathbf{v}_H^T\mathbf{Q}_H^{-1}\mathbf{v}_H + \mathbf{v}_N^T\mathbf{Q}_N^{-1}\mathbf{v}_N = \min \qquad \textbf{(4)}$$

with $\mathbf{Q}_s^{-1}$ being an appropriate weight matrix for the unknown correction signals.

One of the main difficulties in this approach is that the mean value $\mathbf{m}_s = \mathbf{E}\{\mathbf{s}\}$ of the stochastic signals will not necessary be zero, due to the systematic behaviour that is supposed to exist in their values. In order to avoid such problem, one can initially solve the system (3) using (4) with a unit signal weight matrix. The initial solution for the signal part (Kotsakis & al., 1999),

$$\begin{aligned} \mathbf{Q} &= \left(\mathbf{Q}_h + \mathbf{Q}_H + \mathbf{Q}_N + \mathbf{I}_m\right)^{-1} \\ \mathbf{W} &= \mathbf{I}_m - \mathbf{A}\left(\mathbf{A}^T\mathbf{Q}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{Q} \\ \hat{\mathbf{S}}_{init} &= \mathbf{Q}\mathbf{W}\mathbf{l} \end{aligned} \qquad \textbf{(5)}$$

Now, we can create the following reduced observations and signals:

$$\mathbf{l}_r = \mathbf{l} - \mathbf{m}_s \quad \text{and} \quad \mathbf{s}_r = \hat{\mathbf{s}}_{init} - \mathbf{m}_s \qquad \textbf{(6)}$$

It is now clear that the reduced signals s_r have zero mean. Furthermore, and in order to compute the cofactor matrix Q_{s_r} and to predict the signal s at points other than the measuring points, it is necessary to have an analytical covariance function of signals. In this context, the numerical values s_r will be used for an empirical determination of a covariance function model describing the average spatial behaviour of the reduced signals s_r and to select consequently the appropriate analytical covariance model.

In this way, we can repeat the adjustment of the model (3) using a new version for the stochastic model of the correction signals:

$$\begin{aligned} &\mathbf{l}_r = \mathbf{A}\mathbf{X} + \mathbf{s}_r + \mathbf{B}\mathbf{v} \\ &\mathbf{E}\{\mathbf{s}_r\} = \mathbf{0} \quad \text{and} \quad \mathbf{E}\{\mathbf{s}_r\mathbf{s}_r^T\} = \mathbf{C}_{s_r} = \sigma_{s_r}^2 \cdot \mathbf{Q}_{s_r} \end{aligned} \qquad \textbf{(7)}$$

Finally, the solution of the adjustment model (1) by using (7) will be given by the following unbiased estimators:

$$Q = \left(Q_h + Q_H + Q_N + Q_{s_r} \right)^{-1}$$

$$W = I_m - A\left(A^T Q A \right)^{-1} A^T Q$$

$$\hat{X} = \left(A^T Q A \right)^{-1} A^T Q l_r \qquad \textbf{(8)}$$

$$\hat{v}_h = Q_h Q W l_r \; , \; \hat{v}_H = -Q_H Q W l_r \; , \; \hat{v}_N = -Q_N Q W l_r , \; \hat{S}_r = Q_{s_r} Q W l_r$$

We note that the deterministic approach is obtained from the collocation approach by omitting the presence of the residual correction signals s. In this case, the general adjustment model (3) will be reduced to the form:

$$L = AX + Bv \qquad \textbf{(9)}$$

The final solution of equation (9) is deducted of the solution given by (8) while putting $Q_{s_r} = 0$. This solution satisfies two different but equivalent minimum conditions, both of which have been given already by Gauss: least squares and minimum variance. The well known least squares condition for the adjustment model (9) is:

$$v_h^T Q_h^{-1} v_h + v_H^T Q_H^{-1} v_H + v_N^T Q_N^{-1} v_N = \min \qquad \textbf{(10)}$$

3.Correction surface model

In practice, the various wavelength errors in the gravity solution may be approximated by two kinds of functions in order to fit the quasigeoid to a set of GPS levelling points. The first model is based on a general 7-parameter similar datum shift transformation with its simplified 4-parameter model for only the geoid determination purpose. The second is a polynomial regression with its simplified case, a planar regression.

In this work and in the aim to minimise the long wavelength errors, the systematic datum differences between the gravimetric geoid and the GPS/levelling data were removed by using the following four-parameter transformation equation:

$$a_i^T x = x_0 + \cos(\rho_i).\cos(\lambda_i)\,\Delta X + \cos(\rho_i).\sin(\lambda_i)\,\Delta Y + \sin(\rho_i)\,\Delta Z \qquad \textbf{(11)}$$

where x_0 is the shift parameter between the vertical datum implied by the GPS/levelling data and the gravimetric datum, and ΔX, ΔY and ΔZ are the three translation parameters in X, Y, Z axes.

Furthermore, and in order to perform the general adjustment model using the collocation approach in which the signal part is considered as additional stochastic parameters, an analytical expression of the covariance function of reduced signals is more convenient. For this purpose and in our case, the Hirvonen model is adopted as optimal analytic covariance function of the reduced signals, which is given by:

$$C(\psi) = \frac{C_0}{1 + \left(\frac{\psi}{\zeta} \right)^2} \qquad \textbf{(12)}$$

where C_0 : variance of the reduced signals,

ζ : denote the correlation distance.

The use of this model as a local covariance function requires the estimation of two parameters: the variance of the reduced signals and the correlation distance. These parameters are obtained by fitting the Hirvonen function to a number of empirical covariance values employing the least squares adjustment.

The empirical covariance function of the reduced signals was computed with a new program developed in the framework of this work using the following formula:

$$C_{ss}(\psi) = \frac{1}{M}\sum s_i . s_j \qquad (13)$$

where M is the number of combinations, ψ is the spherical distance between Q_i and Q_j and s is the reduced signal.

The summation was made for all the combinations of the data points Q_i and Q_j whose distance was comprised between $(\psi - \Delta\psi / 2)$ and $(\psi + \Delta\psi / 2)$, and here $\Delta\psi$ = 0.5 minutes. The value of the sampling interval size $\Delta\psi$ represents the minimum distance between the benchmarks stations.

The obtained results of the empirical covariance function are identical to these calculate by the **EMPCOV** program of the **GRAVSOFT** software.

4. Numerical tests

4.1. Description of program

A Fortran program **ADJ_GLG** has been developed at National Centre of Space Techniques by using the general adjustment model described above. This program has two main objectives. The first aim is to adapt the gravimetric quasigeoid to a set of levelled GPS reference points, and the second aim is to proceed to meticulous and reliable analysis gravimetric geoid accuracy.

The present version will perform three different adjustments by using:

- The Deterministic approach,
- The collocation approach, and
- The Least squares adjustment without information a priori on the accuracy of height data sets.

In the last case, only the 2D spatial field that describes all the datum/systematic distortions between the available height data sets were estimated.

Special attention has been paid to the organisation of data of gravity geoid grid and the GPS levelling point to establish the observation equation matrix, and to save memory and to speed up computation.

The input data necessary to perform the **ADJ_GLG** program were: the grid of the geoid in standard and binary format, the levelled GPS benchmarks coordinates, and the variances-covariance matrices of the GPS/levelling and Geoid networks if they are available, otherwise the uniform accuracy of the ellipsoidal, orthometric and geoid heights will be used.

4.2. Data used

4.2.1. GPS surveys

Many GPS campaigns have been carried out in the past years in Algeria. Furthermore, in the framework of the TYRGEONET (*TYRhenian GEOdetic NETwork*) project, two sites located in the North of Algeria have been determined in the WGS84 system, which have been used later for the densification and improvement of accuracy of the local geodetic networks. The number of stations GPS used in this investigation was 41, which 18 are benchmarks of the first order levelling network, and the others belong to the second levelling network. All of these points are located in the north of Algeria whose 34 points are close to the station of Arzew (see fig. 2). The GPS observations were performed with four ASHTECH Z-12 dual frequency receivers with baseline length ranging from about 1 to 1000 km, and the BERNESE software with precise ephemerides was used to process the GPS data. The computed ellipsoidal heights were referred to WGS84 system and their standard deviations do not exceed 3 cm. So, in order to make possible the estimation of N (geoid undulation) in these points, all GPS stations have been connected to the national levelling network, which consists of orthometric heights. The accuracy of the levelling heights may be estimated to about 6 cm depending on the type of connection measurements.

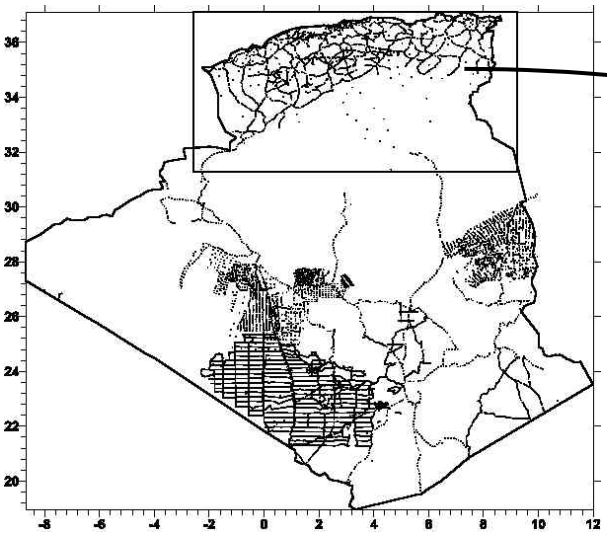

Fig 1. Geographical distribution of BGI gravity Measurements.

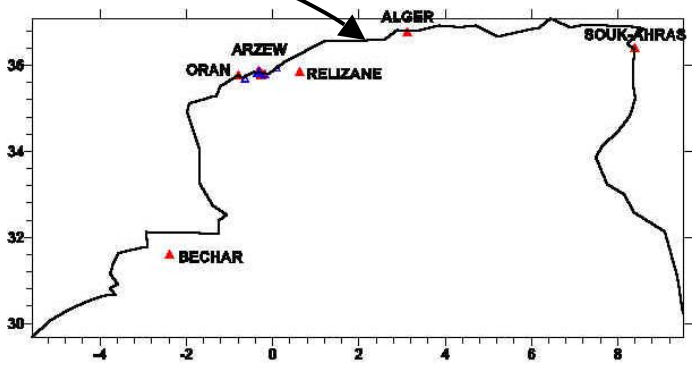

Fig 2. Geographical distribution of GPS stations (Δ : Benchmark, + : Control point)

4.2.2. Local geoid

In view the use of the GPS for the orthometric height computation, the National Centre of Space Techniques through the national projects of research, has recently focused a part of the current research on the precise geoid determination using different methods. In 1999 a new gravimetric geoid published in the IGes Bulletin, was computed over the whole of Algeria by the present author (Benahmed, 2000). This solution is based on the validated gravity data supplied by the BGI, topographic information and the optimal geopotential model OSU91A, which were combined using the remove-restore technique in connection with the Fast collocation method. The final result is a gravimetric geoid on a 5' x 5' grid in the area with the limits $20° \leq \varphi \leq 37°$ and $-7° \leq \lambda \leq 10°$. The Fig.3 shows a map of the quasi-geoid solution in Algeria contoured with 1m interval.

The selected zone to test this program is located in north of Algeria. The choice was emphazised by the relatively high density of the gravity points, and by the availability of the precise GPS stations covering the whole of the area (see fig. 1.).

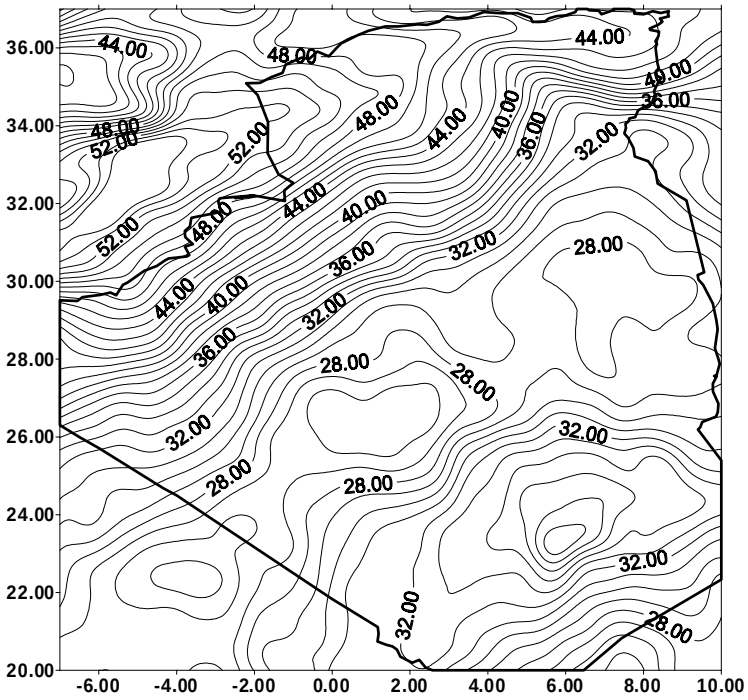

Fig. 3. Quasi-geoid solution in Algeria (m)

4.2.3. Practical results

Before applying the general adjustment model, a first computation by removing the long length effects using the four parameters transformation permits to confirm the existence of one suspect blunder in the GPS levelling measurements. The result of the pre-adjustment shows that the combined method is very effective for detecting the blunder errors in the GPS levelling measurements a condition that the short wavelength of the geoid are very modelling.

After one rejection, 40 GPS levelling points have been selected but only 18 well distributed GPS levelling points are used as benchmarks points, and all the other points were excluded from the combined adjustment in order to estimate the real accuracy given by the comparison between the adjusted values and the known ones. Moreover, and since the variance-covariance matrix of the GPS and levelling networks adjustments necessary for this kind of combined adjustment are not available, we have used the a priori uniform accuracy of the ellipsoidal, orthometric fixed to 3 cm, 6 cm respectively according to networks accuracy. Also and in absence of the prior geoid error model, we have used a unit weight matrix and get an estimate for the a posteriori unit geoid variance.

The following tables give the obtained results of the general adjustment model by using the collocation approach. For the very long-wavelength errors, the four-parameter similitude datum shift transformation model was determined. The values of parameters are presented in the table 1. The fig. 4 shows the empirical and analytic covariance functions of the reduced signals in benchmarks points. The bad behaviour of the empirical function is principally due to statistical character of the reduced signals in benchmarks and to the total number of the GPS stations used in the computation of the empirical values, which is too small relatively to experimental area size. Furthermore, the most GPS points are close to Arzew station.

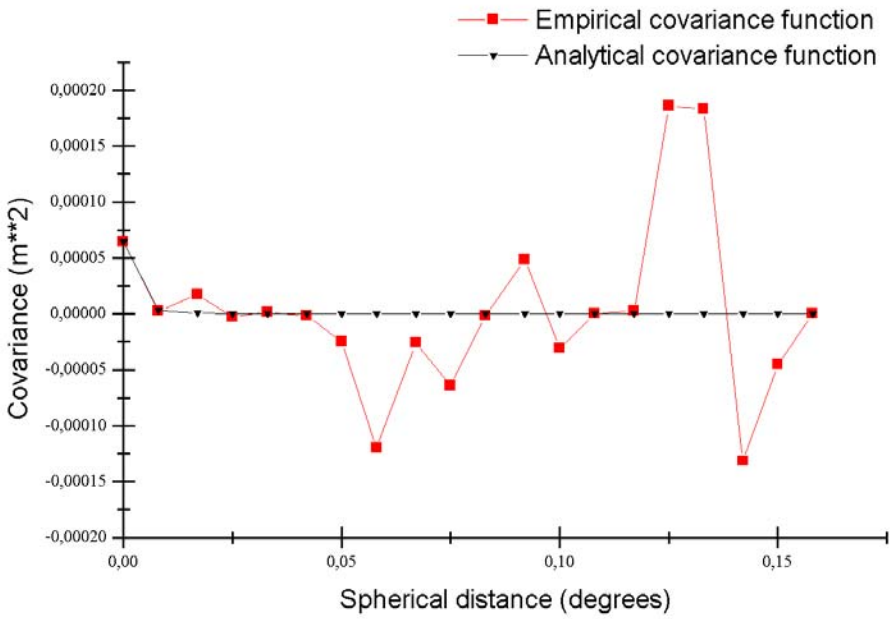

Fig 4. Covariance functions of reduced signals

The tables 2 and 3 summarise respectively the statistics of the differences between the geoid height determined by GPS/Levelling and gravimetric quasigeoidal heights after fitting out the systematic biases using a four-parameter transformation, and the individual components of adjusted GPS, levelling and geoid residuals. The first statistics show that a good fit between the gravimetric quasigeoid and GPS/Levelling has been reached, but we do not believe that the rms value is the real accuracy of the determined geoid, this provided proof that the combined adjustment can optimally fit the gravity geoid to the GPS levelling point in the least square sense. However, the second statistics prove that the residuals in benchmarks are due particularly to gravimetric quasigeoid errors.

The standard deviation of the discrepancies between the gravimetric quasi geoid solution and GPS levelling geoid undulations at control points amounts to ± 5cm after fitting. This fact confirms the good fit in the test area between the Algerian quasigeoid and GPS/levelling data has been reached.

Finally, we note that the three approaches described above give the similar values of parameters and the residuals. It is due to the fact that the variance-covariance matrices of different heights data sets used in this investigation are the diagonal matrices and the estimated signals are the very small quantities.

Parameters	Value	RMS
ΔX (m)	-750.482	14.925
ΔY (m)	-40.376	1.149
ΔZ (m)	-497.213	9.155
Scale factor	0.00014090	0.000003

Table 1. Four-parameter transformation model.

Mean	Min.	Max.	RMS
0.000	-0.030	0.055	0.020

Table 2. Residuals after fitting the Gravimetric geoid to 18 GPS Levelling points (in meters)

	Mean	Min	Max	RMS
v_h	0.000	-0.001	0.001	0.000
v_H	0.000	-0.003	0.002	0.001
v_N	0.000	-0.051	0.028	0.017

Table 3. Individual components of GPS, levelling and geoid residuals (in meters).

Mean	Min	Max	RMS
0.017	-0.074	0.104	0.047

Table 4. Statistics of differences between the gravimetric geoid undulations and GPS levelling at 22 control points (in meters).

Conclusion

This paper presents the new tool for an adjustment of combined GPS/levelling/geoid networks. The proposed methodology is interesting specially when we envisage to use the GPS techniques for levelling purposes with respect to a local vertical datum, and when we want to proceed to a meticulous and reliable analysis of relative observations to different heights. Two modelling alternatives for the correction field are programmed namely a pure deterministic parametric model, and a hybrid deterministic and stochastic model. So, the developed program can be used for testing the reliability of preliminary geoid error models, which have been derived via internal error propagation from the source data and their noise used in the gravimetric solution.

However, the difficult step in the application of general adjustment model using the collocation approach is the estimation of the covariance function of reduced signals and subsequently the selection of its corresponding analytic representation. In this context, the special attention will be made for the optimal choice of a local covariance model of the reduced signals capable to describe their spatial behaviour.

In the field experiment introduced above, the results of the numerical test show that:

- The estimated signals don't have any influence on the obtained results since their magnitude varies between 10^{-6} and 10^{-7} meters.
- A good fit between the Algerian quasigeoid and GPS/levelling has been reached, it proves clearly that the combination of GPS/levelling and geoid models is capable to produce orthometric heights with an accuracy acceptable for the low order levelling network densification
- The residuals in benchmarks are due mainly to gravimetric quasigeoid errors in the test area.

Finally, the results obtained were satisfactory, so in the near future the new adjustment will be performed for a more reliable geoid assessment. This will include an accurate gravimetrically geoid computed in the whole of Algerian territory by integrating the maximum of new gravimetric, topographic and geodetic informations and the new data of GPS/levelling. Furthermore, it is necessary to add in the general adjustment model the variance-covariance matrices of different height data sets in order to take account of the correlation between data.

Acknowledgements

I would like to thank Prof. C. Kotsakis and F. Sanso for their encouragements and their support.

References

Balmino G., Sanso F. (1995) New Geoids in the World. IGeS Bulletin N° 4

Benahmed Daho S. (2000) The New Gravimetric Geoid in Algeria. IGeS Bul. N 10

Jiang Z., Duquenne H (1996) On the combined adjustment of gravimetrically determined geoid and GPS levelling stations. Journal of Geodesy 70: 505-514

Knudsen P. (1987) Estimation and Modelling of the Local Empirical Covariance Function using gravity and satellite altimeter data. Bulletin Géodésique, vol. 61.

Kotsakis C., Sederis M. (1999) on the adjustment of combined GPS/levelling/geoid networks. Journal of Geodesy 73: 412-421

Moritz H. (1976) Covariance Function in Least Squares Collocation, Report N° 240, Ohio State University.

Moritz H. (1980) Advanced Physical Geodesy, H. Wichmann-Abazcus Press, Karlsruhe-Tundridge Wells.

Rao CR (1971) Estimation of variances components – MINQUE theory. J Multivar Statist 1:257-275

Rapp R. H., Sanso F. (1995) Determination of the geoid present and future. International Association of Geodesy symposia, N° 106.

Sjoberg L (1984) Non-negative variance component estimation in the Gauss-Helmert adjustment model. Manuscrita Geodetica 9:247-280.

A NEW QUASI-GEOID COMPUTATION FROM GRAVITY AND GPS DATA IN ALGERIA

S. A. Benahmed Daho (1), J. D. Fairhead (2)

(1) National Centre of Space Techniques, Geodetic Laboratory - BP 13 Arzew - 31200 - Algeria. Email: d_benahmed@hotmail.com /Fax: +213-4147-3665

(2) GETECH-University of Leeds – Department of Earth Sciences – Leeds – LS2 9JT –United Kingdom. Email: jdf@getech.leeds.ac.uk /Fax: +44-113 2429234

Abstract

Due to the rapid increasing use of GPS heighting, which already gives the same accuracy as levelling over some 10 to 100 km, there is an urgent need to provide the ″cm quasi-geoid″ to geodesists and surveyors. In this context, the Geodetic Laboratory of National Centre of Space Techniques has recently focused a part of the current research on the precise geoid determination using different methods. In 1997, the first determination of a preliminary geoid in a small zone in the north of Algeria was calculated by using the least squares collocation and the Gravsoft software. Nowadays, an improved quasi-geoid has been computed over the whole of Algeria. This solution was based on the validated gravity data supplied by the Geophysical Exploration Technology Ltd (GETECH), topographic information and optimal geopotential model, which were combined by using the remove-restore technique in connection with the Fast Fourier Transformation. The GLOBE 30″ digital elevation model has provisionally been chosen as the DEM to be used for computation of the effects of the topography according to the RTM reduction modelling method. However, the spherical harmonic coefficient set OSU91A, complete to a degree and an order 360, was adopted as a reference in order to eliminate the long wavelengths of the gravity field. In this paper, the main features of the Algerian quasi-geoid solution are summarized, and extended tests of this solution are undertaken using the new GPS and levelling data collected from the TYRGEONET project and the local GPS/Levelling surveys. The comparisons based on different GPS campaigns provide, after fitting by using the four-parameter transformation, an RMS differences ± 11cm especially for the north part of the country over distances of 1 to 1000 km and proves that a good fit between the new quasi-geoid and GPS/levelling data has been reached.

Key words: Fast collocation method, GPS/levelling, Geopotential model, TYRGEONET project

1. Introduction

Geoid determination is one of the most fundamental problems in geodesy. The precise model of the geoid not only enable us to transform satellite-derived heights to physically meaningful heights based on The Earth's gravity field, but also plays an important role in geophysics and oceanography. The computation of this surface can be done in a fast and efficient way by means of FFT techniques or Fast collocation which lead to geoid estimates over large areas, so avoiding sub-areas computation and patching of sub-solutions.

Furthermore, accurate global geopotential models and, over some regions, detailed DTM allows an effective reduction of the gravity data, thus implying an optimal application of the "remove-restore" procedure.

The first determination of a preliminary geoid in a small zone in the north of Algeria was done in 1997 (Benahmed Daho et al., 1998), using the Least Squares Collocation (LSC) method and the Gravsoft software. In second gravimetric solution over the whole of Algeria territory, the BGI gravity data and topographic information were included (Benahmed Daho, 2000).

Nowadays, an improved quasi-geoid has been computed over the whole of Algeria between the limits $16° \leq \varphi \leq 40°$ for latitudes and $-10° \leq \lambda \leq 14°$ for longitudes, in a grid with mesh of 5'x 5'. This solution was based on the validated gravity data supplied by the Geophysical Exploration Technology Ltd (GETECH), topographic information and tailored geopotential model, which were combined using the remove-restore technique in connection with the Fast Fourier Transformation (FFT). The computation of the topography effects according to the RTM reduction modelling method is based on a global topographic model GLOBE of 30″ x 30″. However, the spherical harmonic coefficient set OSU91A, complete to a degree and an order 360, was adopted as a reference for to remove and restore the long wavelength components of the gravity and the geoid respectively.

In order to test the quality and the accuracy of the quasi-geoid calculated, the results of this computation are compared to GPS/Levelling data collected from the international TYRGEONET project and the local GPS/Levelling surveys.

2. The available data

In order to fulfil the requests of the FFT-based algorithms used in this paper for geoid height computation, we did not restrict ourselves to the continental area of Algeria when selecting the data, but we considered also the surrounding sea region. Therefore, the data selection area is bounded by limits $16° \leq \varphi \leq 40°$; and $-10° \leq \lambda \leq 14°$. The prediction area is also bounded by same limits, but it is noted that the results at the borders of the test area are affected by errors, generally due to the lack of data around the evaluation points.

2.1. Gravity data

For this work, the pre-processed free air anomalies on a 5' grid in the area bounded by the limits mentioned above, derived by merging terrestrial gravity data and satellite altimetry data, have been provided by GETECH through the agreement between the National Centre of Space Techniques/Geodetic Laboratory and University of Leeds/GETECH without any information on the accuracy of different values. These data have been acquired in the framework of African Gravity Project (AGP) from Bureau Gravimétrique International (BGI), Defence Mapping Agency and from oil companies and many national academic and non-academic organisations in different countries. All of these gravity measurements are adjusted at GETECH to IGSN71 by using "Latin American Gravity standardisation Net 1977" and are referred to the Geodetic Reference System 1980. Figure 1 gives a graphical representation of the gravity data coverage in the computation area. From the figure it becomes clear that the coverage with gravity observations is not sufficient for some land areas particularly in the south of the country and new measurements are needed to accomplish a homogeneous coverage.

Furthermore and since the history of GETECH gravity data processing is not clearly understood, it is very difficult to estimate the accuracy of the resulted geoid heights. However and considering that the BGI gravity data on the Algerian territory have been included for the generation of the previous 5' grid of anomalies, we have proceed to assess the prediction accuracy by comparing the each validated BGI gravity observation with the a value predicted from the GETECH grid using the Spline interpolation. Table 1 summarises the statistics of the differences. The analysis of the results shows the large discrepancies between the original BGI gravity data and predicted ones and proves that gridded gravity set was not derived for precise geoid determination purpose, but for other geophysical applications such as regional geological interpretations.

Anomalies	Mean	Sd	Min.	Max.
Δg (Obs)	4.887	26.845	-82.590	136.200
Δg (Pred)	4.380	26.184	-89.650	107.825
Differences	4.657	0.502	-29.923	69.666

Table 1. Statistics of the differences (mGals)

2.2 Geopotential model

The tailored high-degree global geopotential model OSU91A (Rapp & al., 1991), complete to degree and order 360, was adopted as a reference in order to eliminate the long wavelengths of the gravity field. Replacing the global model OSU91A by the most recent model EGM96 do not improve significantly the present solution, because no new gravity data from Algerian territory was incorporated in EGM96. Gravity anomalies and geoid undulation can be computed in a spherical approximation from a geopotential coefficient set by:

$$\Delta g = \frac{GM}{r^2}\sum_{n=2}^{N_{max}}(n-1)\sum_{m=0}^{n}\left[\overline{C}_{nm}.\cos m\lambda + \overline{S}_{nm}.\sin m\lambda\right]\overline{P}_{nm}(\cos\theta) \qquad (1)$$

$$N = \frac{GM}{r\gamma}\sum_{n=2}^{N_{max}}\sum_{m=0}^{n}\left[\overline{C}_{nm}.\cos m\lambda + \overline{S}_{nm}.\sin m\lambda\right]\overline{P}_{nm}(\cos\theta) \qquad (2)$$

where θ, λ are the geocentric colatitude and longitude of the point where Δg and N will be determined ; $\overline{C}_{nm}, \overline{S}_{nm}$ are the fully normalised spherical geopotential coefficients of the anomalous potential ; $\overline{P}_{nm}$ are the fully normalised associated Legendre functions ; N_{max} is the maximum degree of the geopotential model.

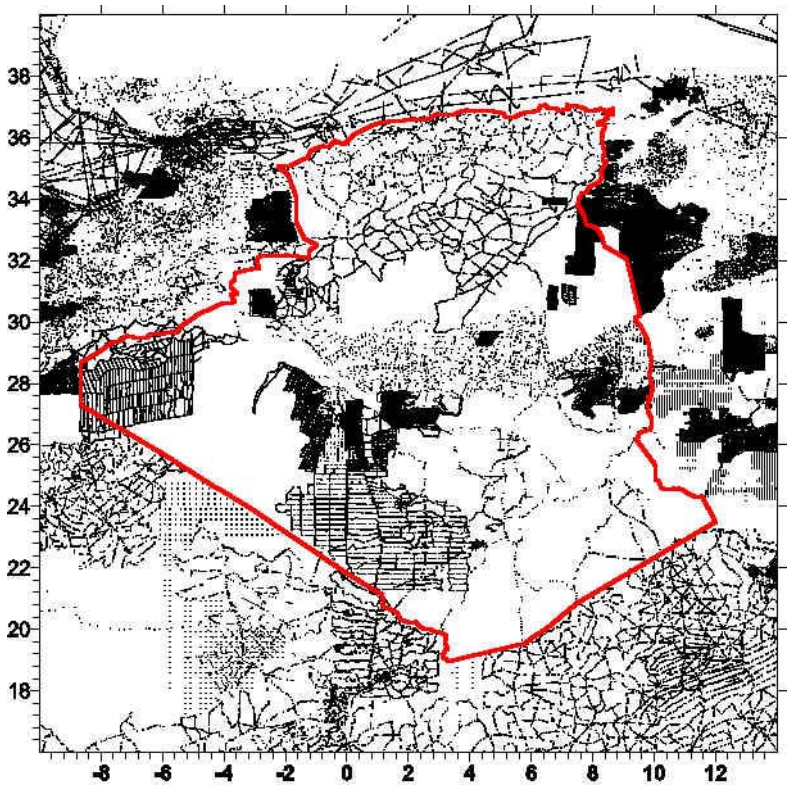

Figure 1. Geographical distribution of GETECH gravity measurements.

2.3 Topographic data

The computation of the terrain effects on the quasi-geoid required a detailed DTM model. For this purpose, The GLOBE 30″ digital terrain model has provisionally been chosen as the DTM to be used for computation of the effects of the topography according to the RTM reduction modelling method.

3. Computation procedure

The computations were done using the "GRAVSOFT" package, developed during a number of years at the National Survey and Cadastre (KMS). For actual solution, the Fast Fourier Transformation (FFT) and the Remove-Restore procedure were used to compute the quasi-geoid estimate. In order to smooth the gravity field, the previous gridded gravity data must be corrected for the effect of the atmospheric and reduced for the effect of the spherical harmonic model and the topography. The spherical harmonic coefficient set OSU91A has been used to remove the long wavelength component of gravity data. The computation of the effects of topography according to the RTM reduction modelling method is based on global topographic model GLOBE of 30″ x 30″ which were used up to a distance of 200 km. The reference surface of 10' x 10' needed for the RTM reduction has been obtained by means of a moving average applied to the detailed one. Table 2 shows the statistics of the reduced gravity data.

Anomalies	Mean	Sd	Min.	Max.
Δg_{Obs}	2.43	28.81	- 172.96	218.84
$\Delta g - \Delta g_{OSU91A}$	-2.29	14.94	-122.90	192.28
$\Delta g_r = \Delta g_{Obs} - \Delta g_{OUS91A} - \Delta g_{RTM}$	-1.55	14.62	-120.12	823.63

Table 2. Statistics of reduced gravity data (mGals).

From the results of Table 1, it is obvious that the OSU91A reference field fits well the gravity in the area under consideration, and the smoothing of the gravity data is considerable after the removal of the topographic effect if we take into account only the mean and standard deviation values. However, we will note that all the gravity residuals values are less than to 277.08 mGals excepting the maximum value which is too large with respect to the statistics of the observed gravity. Probably, this is due to the errors produced by the FFT technique and the global topographic model GLOBE used in the computation of the topographic correction at borders of the test area.

The residual quasi-geoid (ζ_r) has been evaluated using FFT technique, implemented in the GEOFOUR program written by Rene Forsberg. The statistics of residual undulations in a 5' x 5' grid (289 x 289 values) are presented in Table 3 while the plot of the values is shown in Figure 2.

Mean	Sd	Min.	Max.
0.00	1.42	-7.05	5.69

Table 3. Statistics of the quasi-geoid residuals ζ_r (m)

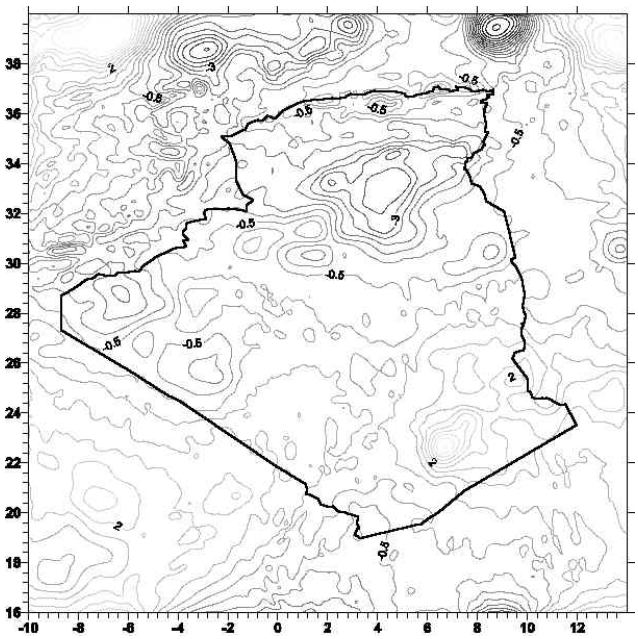

Figure 2. Residual quasi-geoid (m)

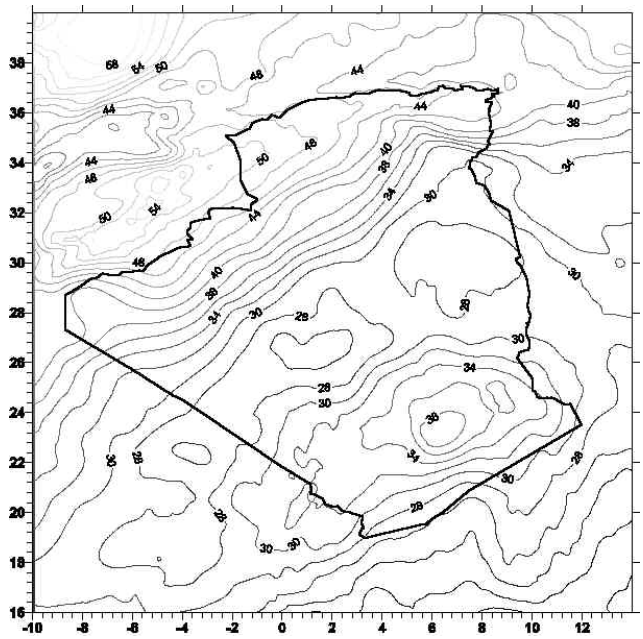

Figure 3. Quasi-geoid solution in Algeria (m)

The final quasi-geoid was obtained by adding the model and the residual terrain effect on the 5' x 5' residual quasi-geoid grid. The values of the geopotential model range from 15.85 m to 56.0 m and yield the major part of the quasi-geoid. The standard deviation and maximum values of the contribution from the gravity data are 1.42 m and 5.69 respectively, while the corresponding values of the RTM effects are 0.07 m and 0.86 m. Figure 3 shows a map of the quasi-geoid solution in Algeria. The statistics of the total quasi-geoid values are summarised in Table 4.

Mean	Sd	Min.	Max.
35.18	8.83	17.69	60.65

Table 4. Statistics of the quasi-geoid (m)

4. Comparison with GPS/Levelling data

Many GPS campaigns have been carried out in the past years in Algeria. Furthermore, in the framework of the TYRGEONET project, two sites located in the North of Algeria have been determined in the WGS84 system, which have been used later for the ALGEONET (*ALGerian GEOdynamical NETwork*) project and for densification and improvement of accuracy of the local geodetic networks. The number of stations GPS used in this investigation was 258, which 16 are benchmarks of the first order levelling network, and the others belong to the second levelling network. All of these points are located in the north of Algeria and distributed as in Figure 5 whose the most are close to the station of Arzew. So, in order to make possible the estimation of N (geoid undulation) in these points, all these GPS stations are connected to the national height system through spirit levelling. The GPS observations were performed with four ASHTECH Z-12 dual frequency receivers with baseline length ranging from about 1 to 1000 km, and the BERNESE software with precise ephemerides was used to process the GPS data.

Among 258 GPS levelling points only 16 well distributed GPS levelling points are used as benchmarks points, and the others were excluded in order to estimate the real accuracy given by the comparison between the adjusted values and the known ones. These values have been compared also to the BGI solution which is based on a set of 12183 validated point free air gravity anomalies supplied by the BGI, two elevation grids; 1 km x 1 km digital terrain model for the north of Algeria and the ETOPO5 for the rest of the area, and the OSU91A geopotential model, which were

combined using the remove-restore technique in connection with the Fast collocation. The final result was a gravimetric geoid on a 5' x 5' grid in the area bounded by limits $20° \leq \varphi \leq 37°$ and $-7° \leq \lambda \leq 10°$ (Benahmed Daho, 2000). These two computations show differences even up to 4 m and the standard deviation value is about 1.35 m. The high discrepancies are attributed principally to the two difference computation methodologies relative to the technique used for the geoid heights estimation and to the DTM employed for the topographic correction computation.

The statistics of the differences for both solutions before and after fitting out the systematic biases and tilts between the gravimetric geoid and the GPS/levelling data by using the appropriate four-parameter transformation are summarised in table 5 and show that no significant improvements have been reached in the GETECH solution comparative to the BGI solution. Probably, it is due to the fact that the data provided by BGI are included in the GETECH database which have been used to generate of previous free air anomalies grid.

Furthermore, and in order to estimate the real accuracy, the GPS/Levelling undulations at 242 control points are compared at adjusted ones. The figure 4 shows the histogram of these differences. We can see that the combination the gravimetric BGI_solution with GPS/levelling gives the best results. Unfortunately, and considering the number of GPS/Levelling stations used in this investigation and their distribution, the obtained accuracy can not be generalised for the all north part of the country.

	Geoid models	Mean	Min.	Max.	Stand. dev.
Before Fitting	**BGI Solution**	-1.431	-1.732	0.306	0.617
	GETECH Solution	0.608	0.455	1.531	**0.321**
After Fitting	**BGI Solution**	0.000	-0.034	0.048	**0.020**
	GETECH Solution	0.000	-0.194	0.135	0.104

Table 5. Comparison of GPS/Levelling undulations to gravimetric quasi-geoidal heights (m)

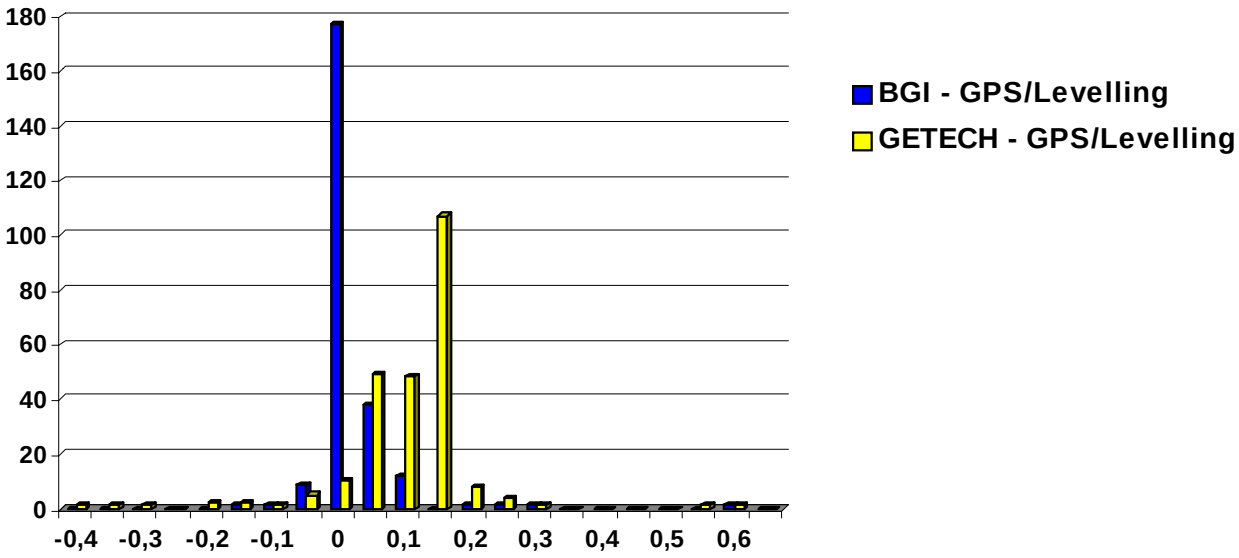

Figure 4. Histogram of the differences with GPS/Levelling geoid undulations at control points (m)

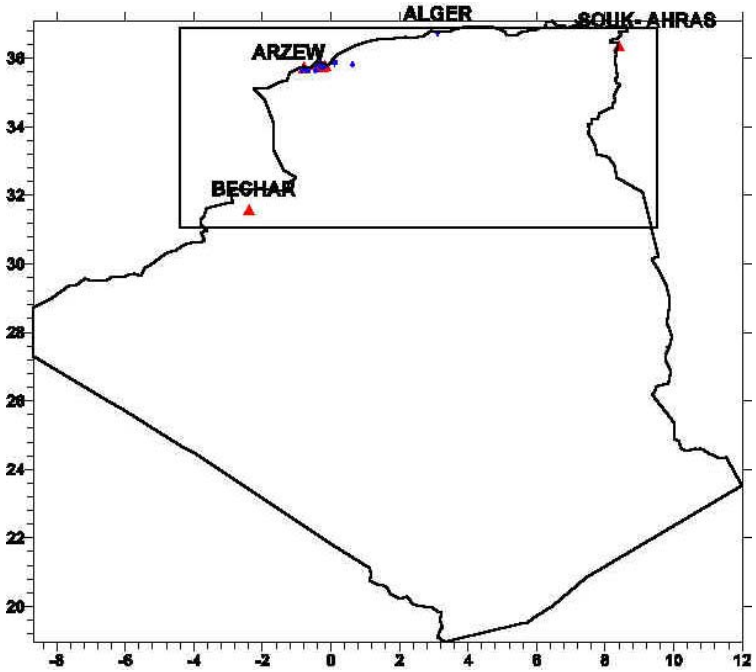

Figure 5. Geographical distribution of GPS stations
(+: Control point, Δ: Benchmark point)

Conclusion

The new Algerian quasi-geoid was computed via Fast Fourier Transformation using the Remove-Restore procedure by integrating the new gravity data supplied by GETECH through the agreement between the National Centre of Space Techniques/Geodetic Laboratory and University of Leeds/GETECH. The comparisons of the new quasi-geoid with GPS/Levelling data provide, after fitting, an RMS differences about ± 11 cm for the north part of Algeria over distances of 1 to 1000 km and prove that good fit in the test area between the new quasi-geoid and GPS/levelling data has been reached. Unfortunately, the non-availability of GPS levelling data on the set of the country with a homogeneous distribution and sufficient density didn't allow to make a more reliable assessment on the quality of the computed geoid.

Finally, the results obtained were satisfactory, so in the near future, the new solutions will be proposed and the additional comparisons should be made for a complete error assessment of the new Algerian quasi-geoid. These will include new gravity data, topographic informations, and the new data of GPS/levelling in order to reach an acceptable accuracy on all the country.

Acknowledgements

The authors wish to thank all Organisations and Persons who provided so kindly the many data and software that contributed to this work:

Prof. R. Forsberg, C. C. Tscherning and P. Knudsen for the GRAVSOFT software,

The Geophysical Exploration Technology Ltd, and the Bureau Gravimétrique International, for the gravity set on the Algerian territory, and

Prof. F. Sanso, for his help and friendly encouragement.

References

Barzaghi R., Borghi A., Sanso F., 2001. Quasi geoid estimations in South America. IGeS Bulletin N° 11 – ISSN 1128-3955.

Benahmed Daho S. A., 2000. The new gravimetric geoid in Algeria. IGeS Bulletin N°10 – ISSN 1128-3955 – [pp 85-90].

Benahmed Daho S.A, Kahlouche S., 1998: The gravimetric geoid in Algeria: First Results. Geodesy on the Move. IAG Scientific Assembly, Rio de Janeiro, September 3-9, 1997.

Benahmed Daho S.A., 2001. A tool for an adjustment of combined GPS/Levelling/Geoid networks. IAG 2001 General Assembly - Budapest - 2 - 8 September 2001

Denizar B., J. Derek F., Maria Cristina L., 1995. A preliminary gravimetric geoid for South America. IGeS Bulletin N° 4 – ISSN 1128-3955 – [pp 53-65].

Fairhead, J. D., Watts A.B., Chevalier P., El-Haddadeh B., Green C.M., Stuart G.W., Whaler K.A. and Whibdle I., 1998. African Gravity Project. Technical Report, university of Leeds Industrial Services Ltd.

Forsberg R., 1985. Gravity field terrain effect computations by FFT. Bulletin Géodésique, vol. 59, N° 4.

IAG, 1971. Geodetic Reference System 1967. Publ. Spéc. N° 3 Bull. Geod., Paris.

Merry C. L., 2001. The African Geoid Project. IAG 2001 General Assembly - Budapest - 2 - 8 September 2001

Moritz H., 1980. Advanced Physical Geodesy, H. Wichmann-Abazcus Press, Karlsruhe-Tundridge Wells.

Rapp R. H., Sanso F., 1995. Determination of the geoid present and future. International Association of Geodesy symposia, N° 106.

Rapp R. H., Y.M. Wang, and N.K. Pavlis, 1991. The Ohio State 1991 Geopotential and Sea Surface Topography Harmonic Coefficient Models. Report N° 410, Department of Geodetic Science and Surveying, The Ohio State University, Columbus.

Tscherning C.C., 1974. A FORTRAN IV program for the determination of the anomalous potential using stepwise least squares collocation. Reports of the Department of Geodetic Science, N° 212, Ohio State University.

Tscherning C.C., 1994. Geoid determination by Least Squares Collocation using GRAVSOFT. Lecture notes for the international school for the determination and use of the geoid, Milano, October.

Tsuei G.C., Arabelos D., Forsberg R., Sideris M.G., Tziavos I.N., 1994. Geoid computations in Taiwan. International Association of Geodesy symposia, N° 113.

Wenzel G., 1999. Global models of the gravity field of high and ultra-high resolution. Lecture notes for the international school for the determination and use of the geoid, Milan, February 15-19, 1999.

The performance of the space-wise approach to GOCE data analysis, when statistical homogenization is applied

F. Migliaccio
DIIAR - Politecnico di Milano - Piazza Leonardo da Vinci, 32 - 20133 Milano - Italy

M. Reguzzoni
Geophysics of Lithosphere Department - Italian National Institute of Oceanography and Applied Geophysics (OGS)
Borgo Grotta Gigante, 42/c - 34010 Sgonico (Trieste) - Italy

F. Sansò
DIIAR - Politecnico di Milano - Polo Regionale di Como - Via Valleggio, 11 - 22100 Como - Italy

C.C. Tscherning
Department of Geophysics - University of Copenhagen - Juliane Maries Vej, 30 - 2100 Copenhagen - Denmark

Abstract. The space-wise approach to GOCE data reduction exploits the spatial correlation of the observations by "projecting" them on a spherical grid at mean satellite altitude; to this aim a local collocation prediction based on a global covariance function of the potential T can be used. Since the implicit hypothesis of rotational invariance is not fulfilled by the gravitational potential, the global covariance function cannot describe the local characteristics of the field in the different interpolation areas. As a consequence the estimation error of the gridded values is not homogeneously distributed all over the reference sphere, but it is much higher over the Himalaya, the Alps, the Andes, etc., i.e. over areas where the random field presents an "unexpectedly" high variability. These errors may degrade the performance of the subsequent spherical harmonic analysis for the recovery of the potential coefficients, especially at very high degrees.

In the light of this reasoning, the whole procedure of the space-wise approach is expected to benefit by a priori making the analysed random field smoother and thereby more homogeneous. This can be done by first determining a global model which describes the main features of the potential distribution on the spherical grid and then by subtracting it from the observed data once and for all. The signal covariance function has to be corrected accordingly. In principle, the resulting field is more "stationary", also when it is regarded as a time-wise process.

This procedure has been applied on simulated GOCE data, showing that the errors of the estimated spherical harmonic coefficients slightly but systematically decrease. Moreover the improvement is much more significant in critical areas.

Keywords. GOCE mission, space-wise approach, gravity field regularization

1 An introduction to the problem

The paper deals with an attempt at assessing the effects of inhomogeneities of the gravity field on the space-wise approach as applied to the GOCE mission (ESA, 1999), as well as, to some extent, at repairing the subsequent drawbacks.

As it is known the same data set can be reduced following different approaches, like the direct method (Bruinsma et al., 2004; Ditmar et al., 2003) or the time-wise method (Pail and Plank, 2002; Schuh, 2000), which however are not in the focus of the present paper.

The space-wise approach (see Figure 1), as understood today, comprises three main steps, i.e.

- a time-wise Wiener filtering (Papoulis, 1984) of data along the satellite orbit;
- a local prediction of grids of suitable functionals of the anomalous potential (e.g. T, T_{rr}, $T_{NN}-T_{EE}$, where T_{NN}, T_{EE} are second derivatives in the northern and eastern directions);
- a harmonic analysis to estimate spherical harmonic coefficients by numerical integration (Migliaccio and Sansò, 1989) or by fast spherical collocation (Sansò and Tscherning, 2003).

The procedure has to be iterated both because when Wiener filtering is applied to raw data, information is lost in the low frequency band (where noise is

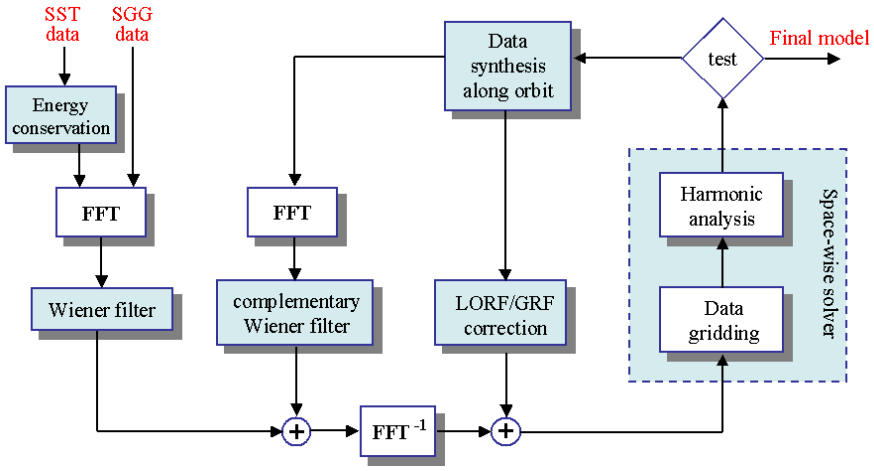

Fig. 1 Scheme of the space-wise approach to GOCE data analysis

prevailing), and in order to correct directional derivatives, due to the instrumental frame wandering; iterations provide the lost information. The procedure is known to be convergent, by numerical tests (Migliaccio et al., 2004a; Migliaccio et al., 2004b).

However, in a first phase of the mission study, when the rotations of the gradiometer were considered as controlled to the level of some arc-minutes, the grid of estimation errors after the first iteration was fairly homogeneous, showing at most some trackiness related to the strong correlation of the noise (see Figure 2).

On the contrary, when a stronger attitude dynamics has to be accepted in the GOCE design, the same procedure yields grids where a clear signature of geophysical features (e.g. the Andes, the Himalaya, etc.) was visible and persistent through the iterations (see Figure 3 and Figure 4).

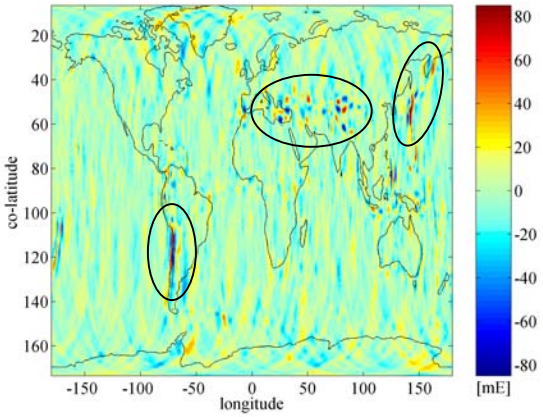

Fig. 3 Estimation error after the first iteration working with second derivatives T_{zz} along the instrumental z-axis. The critical areas are encircled

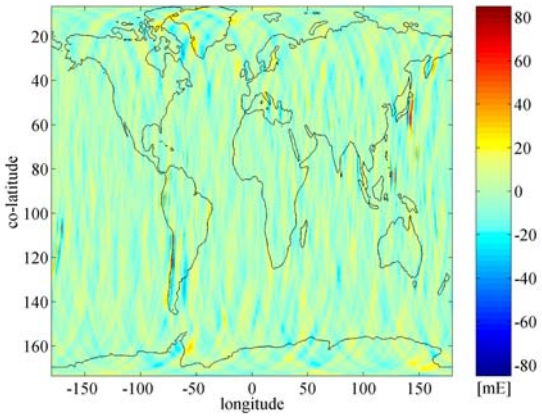

Fig. 2 Estimation error after the first iteration working with second radial derivatives T_{rr}

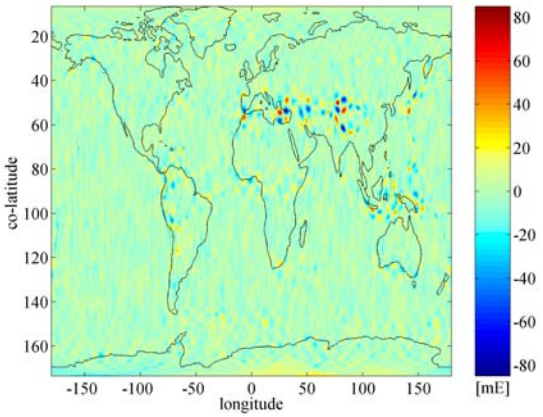

Fig. 4 Estimation error after the last iteration working with second derivatives T_{zz} along the instrumental z-axis

This of course would call for a regional gridding procedure using regionally adapted covariance functions (Arabelos and Tscherning, 2003) or for a multiscale spherical analysis (Freeden, 1999) which still remains an interesting alternative for the future; however dealing with present choices, it seemed to the authors that a kind of "homemade" multiscale analysis could be performed by:

- cutting the highest signal values (e.g. above the 2σ level, with σ computed on all the data);
- making a specific spherical harmonic expansion of a smoothed version of the "peak only" grid;
- recomputing a new data set where the effect of the peaks are removed;
- applying the space-wise approach to the regularized data set;
- adding back to the estimated spherical harmonic coefficients, those used to eliminate the peaks.

We hoped that such a remove-restore procedure, by providing a more homogeneous and smooth data set to the system, would in the end produce a more accurate estimate of T.

2 The experiment and its results

In order to test the performance of the iterative space-wise approach when the previously described regularization procedure is applied, a simulation has been performed in a typical GOCE framework.

The EGM96 model $\{T_{nm}\}$, from degree 25 to 300, is used to simulate the potential values T and the second derivatives T_{zz}, along the instrumental z-axis oscillating according to the following long-period sinusoidal laws:

$$\theta_y = A_y \sin \omega_y t \qquad A_y = 3° \qquad \frac{2\pi}{\omega_y} \cong 2\,\text{hours}$$

$$\theta_r = A_r \sin \omega_r t \qquad A_r = 1° \qquad \frac{2\pi}{\omega_r} \cong 1\,\text{hours}$$

where θ_y is the rotation around the y-axis (pitch) and θ_r around the radial direction (roll).

The observations are simulated every second along a circular orbit, at an altitude of 250 km and with an inclination of 96.5°, for a total number of data $N = 2^{23} \cong 8 \cdot 10^6$, corresponding to a mission length of about 100 days.

A coloured noise, specified by the spectrum in Figure 5, is added to the T_{zz} observations, while a white noise with a standard deviation of $0.3\text{m}^2/\text{s}^2$, corresponding to an orbital error of about 3 cm, is added to the potential T values (see Figure 6).

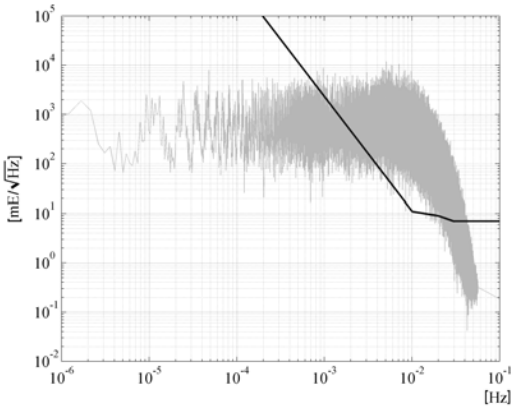

Fig. 5 Signal (in grey) and noise (in black) power spectral density of the second derivatives T_{zz}

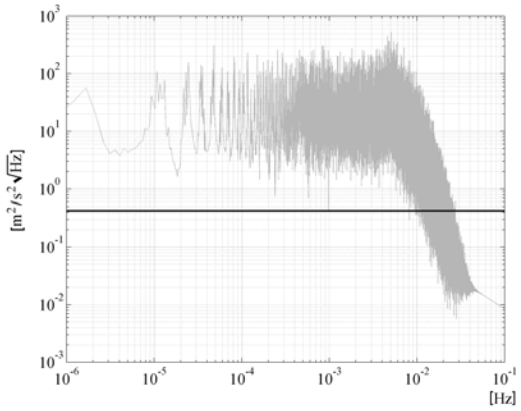

Fig. 6 Signal (in grey) and noise (in black) power spectral density of the potential T

A multiple-input multiple-output (MIMO) Wiener filter is designed by using the empirical spectra of T_{rr} and T as well as the empirical cross-spectrum between T_{rr} and T, all computed from the EGM96 model.

After filtering, the data are locally interpolated on a spherical grid at satellite altitude, with a cell size of 0.72°×0.72°; in fact this allows for an almost exact integer partitioning of the latitude interval covered by the observations, excluding the polar gaps. The local gridding is implemented by collocation, cell by cell, on a moving window of double size with respect to the grid cell.

At this point, we set up the homogenization procedure, selecting the T_{rr} gridded data which are outside the global 2σ value (in our simulation σ is equal to 143.7 mE). Note that the thresholding is only based on the T_{rr} values and not on the potential T, which is much smoother and thus more homogeneous.

Figure 7 shows the T_{rr} grid as obtained by the local gridding procedure, while Figure 8 represents the "peak only" grid, referred back to the zero level; in other words, we have considered a field only in the peak areas, with values $(T_{rr} - 2\sigma)$ where $T_{rr} > 2\sigma$ and $(T_{rr} + 2\sigma)$ where $T_{rr} < -2\sigma$.

After a moving-average smoothing to avoid Gibbs effects, we get a spherical harmonic global model $\{\delta T_{nm}\}$ (by numerical integration), which reasonably describes the "peak only" grid.

This model has to be subtracted from the original one, in order to obtain a more homogeneous data set, i.e.

$$T_{nm}^{new} = T_{nm} - \delta T_{nm}$$

$$(T_{rr}, T)^{new} = (T_{rr}, T) - (\delta T_{rr}, \delta T)$$

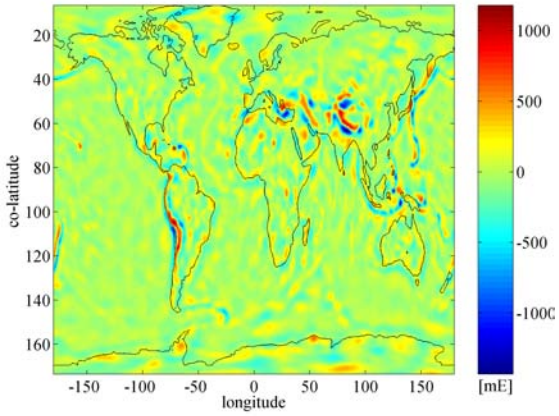

Fig. 7 T_{rr} estimated grid after the first step of the iterative space-wise approach

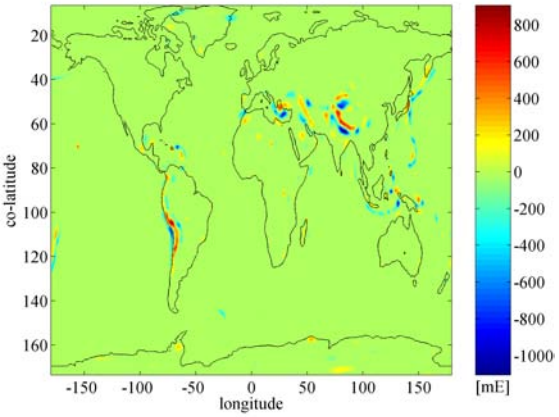

Fig. 8 Grid of the highest T_{rr} values, above the 2σ threshold; the values are referred back to the zero level

Note that the signal degree variances of the new model, and hence the corresponding global covariance functions, remain practically the same, as shown in Figure 9.

Starting from the new regularized data, we first apply the space-wise approach and then, at the end of the iterative procedure, add back the "peak only" model.

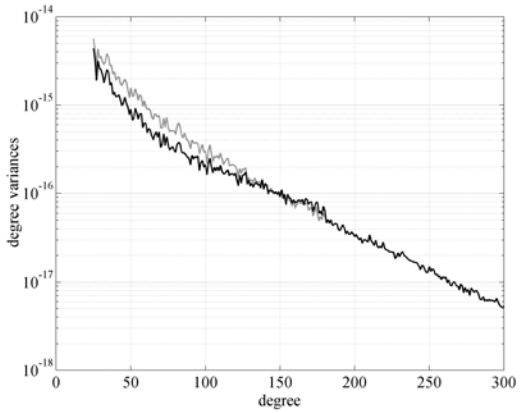

Fig. 9 Degree variances of the EGM96 model (in grey) and of the new one (in black) after regularization

Let us come to the results of this technique, comparing the cases with and without statistical homogenization.

The error degree variances of the new solution, shown in Figure 10, are globally systematically better. The GOCE requirement of solving the gravity field up to degree 200 can be safely reached.

The amount of the improvement in the harmonic coefficients estimation can be evaluated by the following relative differences

$$\varepsilon_n = \frac{\delta\sigma_n - \delta\sigma_n^{new}}{\delta\sigma_n}$$

where $\delta\sigma_n$, $\delta\sigma_n^{new}$ are the error degree standard deviations after the last step of the space-wise iterative procedure, without and with regularization respectively. This index is displayed in Figure 11, degree by degree, clearly showing a systematic improvement of the order of about 5%. Let us also note that, apart from one value, $\delta\sigma_n > \delta\sigma_n^{new}$.

Nevertheless the real advantage of this technique can be better perceived locally. To this aim we concentrate on a critical area, e.g. in the Himalaya ($78° < \lambda < 92°$, $60° < \theta < 68°$), where the grid estimation errors are very high.

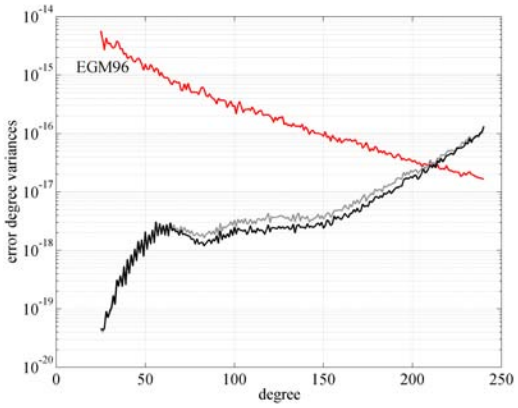

Fig. 10 Error degree variances of the estimated model with regularization, after the first (in grey) and last (in black) iteration. The reference model is EGM96

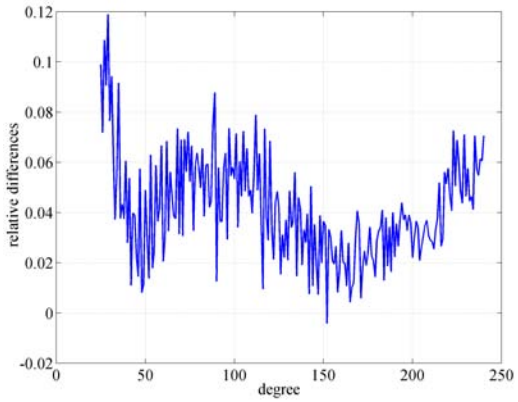

Fig. 11 Relative differences of the error degree standard deviations without and with regularization, after the last step of the iterative procedure

The final errors, in terms of geoid undulation *N*, are shown in Figure 12 and Figure 13, without and with the statistical regularization, respectively. The same errors for the gravity anomalies Δg are displayed in Figure 14 and Figure 15. These errors have been computed as differences with respect to the "true" EGM96 model, up to degree 200. It is clear that the new estimated model produces a much more homogeneous error distribution, reducing the "local peak" down to 30%. Some statistical indicators are reported in Table 1 and Table 2. The values are rather large considering the global goal of the GOCE mission of 1 cm in the geoid and 1-2 mgal in the gravity anomalies at about 1 degree wavelengths. However, note that we have simulated only 3 months of data and we have not used all the components provided by the GOCE gradiometer. Their additional use may improve the solution considerably (Tscherning, 2003).

Table 1. Maximum and r.m.s. error of the geoid undulation, in the Himalaya region, without and with the regularization

max error of N		r.m.s. error of N	
without reg.	with reg.	without reg.	with reg.
1.071 m	0.722 m	0.336 m	0.229 m
relative difference		relative difference	
32.57%		31.80%	

Table 2. Maximum and r.m.s. error of the gravity anomalies, in the Himalaya region, without and with the regularization

max error of Δg		r.m.s. error of Δg	
without reg.	with reg.	without reg.	with reg.
23.28 mgal	16.53 mgal	7.74 mgal	5.52 mgal
relative difference		relative difference	
29.01%		28.75%	

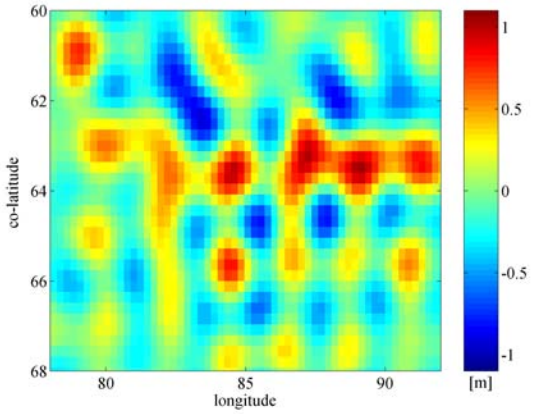

Fig. 12 Estimation errors of the geoid undulation *N* in the Himalaya region without applying the regularization

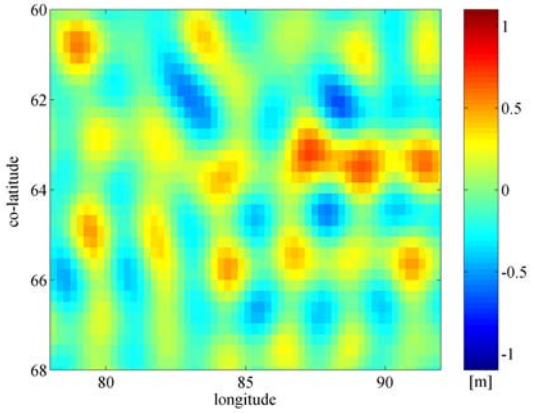

Fig. 13 Estimation errors of the geoid undulation *N* in the Himalaya region after applying the regularization

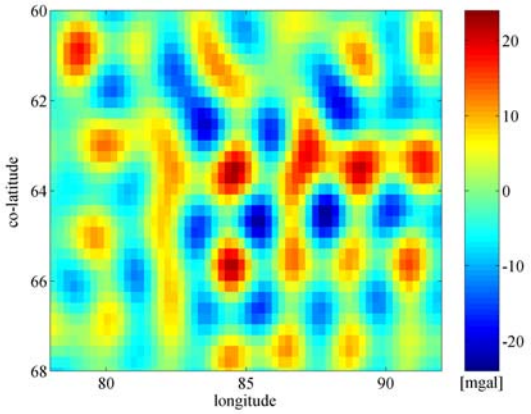

Fig. 14 Estimation errors of the gravity anomalies Δg in the Himalaya region without applying the regularization

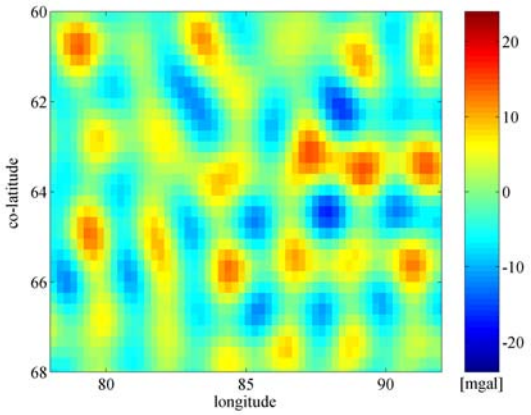

Fig. 15 Estimation errors of the gravity anomalies Δg in the Himalaya region after applying the regularization

3 Conclusions

Although the use of this remove-restore approach to smooth the high level inhomogeneities before applying collocation has not a very large impact on the accuracy of harmonic coefficients estimation, we think that:

- the improvement is systematic and therefore worthwhile;
- the local effect of this correction in critical areas is very significant, thus justifying the effort of its implementation.

The reason why such a strong signature of the errors was not visible in previous simulations of the space-wise approach, stems in our opinion from the fact that, by introducing strong rotational dynamics of the satellite, we make the hypothesis of stationarity, on which the Wiener filtering is based, much less realistic (Albertella et al., 2004).

As a result the areas of strongest signal introduce an error with a signature directly proportional to the signal strength. Would this hypothesis be confirmed, one should also think of applying a similar procedure to the time-wise approach.

Acknowledgements

This work has been prepared under ESA contract no. 18308/04/NL/MM (GOCE High-level Processing Facility).

References

Albertella, A., F. Migliaccio, M. Reguzzoni, F. Sansò (2004). Wiener filters and collocation in satellite gradiometry. In: *Proc. of 5th Hotine-Marussi Symposium on Mathematical Geodesy*, Matera, Italy, 17-21 June 2002, pp. 32-38.

Arabelos, D., C.C. Tscherning (2003). Globally covering a-priori regional gravity covariance models. *Advances in Geosciences*, Vol. 1, pp. 143-147.

Bruinsma, S., J.C. Marty, G. Balmino (2004). Numerical simulation of the gravity field recovery from GOCE mission data. In: *Proc. of 2nd International GOCE User Workshop*. Frascati, Italy, 8-10 March 2004.

Ditmar, P., R. Klees, F. Kostenko (2003). Fast and accurate computation of spherical harmonic coefficients from satellite gravity gradiometry data. *Journal of Geodesy*, 76, pp. 690-705.

ESA (1999). Gravity Field and Steady-State Ocean Circulation Mission. ESA SP-1233 (1). ESA Publication Division, c/o ESTEC, Noordwijk, The Netherlands.

Freeden, W. (1999). Multiscale modelling of spaceborne geodata. Teubner Verlag, Stuttgart.

Migliaccio, F., F. Sansò (1989). Data Processing for the Aristoteles Mission. Proceedings of the Italian Workshop on the European Solid-Earth Mission Aristoteles, Trevi, Italy, 30-31 May 1989, pp 91-123.

Migliaccio, F., M. Reguzzoni, F. Sansò (2004a). Space-wise approach to satellite gravity field determination in the presence of coloured noise. *Journal of Geodesy*. Accepted for publication.

Migliaccio, F., M. Reguzzoni, F. Sansò, C.C. Tscherning (2004b). An enhanced space-wise simulation for GOCE data reduction. In: *Proc. of 2nd International GOCE User Workshop*. Frascati, Italy, 8-10 March 2004.

Pail, R., G. Plank (2002). Assessment of three numerical solution strategies for gravity field recovery from GOCE satellite gravity gradiometry implemented on a parallel platform. *Journal of Geodesy*, 76, pp. 462-474.

Papoulis, A. (1984). Signal analysis. McGraw Hill, New York.

Sansò, F., C.C. Tscherning (2003). Fast spherical collocation: theory and examples. *Journal of Geodesy*, 77, pp.101-112.

Schuh, W.D. (2000). Scientific data processing algorithms. In: *From Eötvös to Milligal*, ESA Project, Final Report.

Tscherning, C.C. (2003). Testing frame transformation, gridding and filtering of GOCE gradiometer data by Least-Squares Collocation using simulated data. In: *Proc. of the IUGG General Assembly*, Sapporo, July 2003.

The effect of an unknown data bias in least-squares adjustment: some concerns for the estimation of geodetic parameters

C. Kotsakis
Department of Geodesy and Surveying, Aristotle University of Thessaloniki
University Box 440, Thessaloniki, GR-54124, Greece, Email: kotsaki@topo.auth.gr

Abstract

Least-squares (LS) estimation is a standard tool for the optimal processing of geodetic data. In the framework of global gravity field modelling, for example, such methods are extensively applied for the determination of geoid solutions from CHAMP and GRACE data via the estimation of a large set of spherical harmonic coefficients. Frequently, in geodetic applications additional nuisance parameters need to be included in the least-squares adjustment procedure to account for external biases and disturbances that have affected the available measurements. The objective of this paper is to expose a trade-off which exists in the LS inversion of linear models that are augmented by additional parameters in the presence of unknown systematic errors in the input data. Specifically, it is shown that if a linear model of full rank is extended by a scalar parameter to account for a common bias in the data, the LS estimation accuracy of *all* the other model parameters automatically worsens. Some simple numerical examples are also given to demonstrate the significance of this accuracy degradation in the geodetic practice.

1. Introduction

Least-squares (LS) estimation is a standard tool for the optimal processing of geodetic data. Its use is commonly associated with a linear(-ized) model of observation equations

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v} \tag{1}$$

$$E\{\mathbf{v}\} = \mathbf{0} \ , \ E\{\mathbf{v}\mathbf{v}^{\mathrm{T}}\} = \mathbf{C} \tag{2}$$

where $\mathbf{y}$ is a vector of observations, $\mathbf{v}$ is a vector of zero-mean random errors with a covariance (CV) matrix $\mathbf{C}$, $\mathbf{A}$ is a matrix of known coefficients, and $\mathbf{x}$ is a parameter vector that needs to be estimated from the given data. Despite its simplistic linear character and its inherent restriction for additive noise, the above model is frequently employed in every field of geodetic research (e.g., *Dermanis and Rummel* 2000).

A problem that is often encountered in geodetic data analysis with the aforementioned general model is the presence of external disturbances (biases) in the available measurements. In cases where the effect of such disturbances can be determined *a-priori* with sufficiently high accuracy (i.e. with an uncertainty that is significantly lower than the data noise level), the original measurements $\mathbf{y}$ should be replaced with a ‘corrected’ set that is compatible with the theoretical

model of Eqs. (1) and (2). If, on the other hand, the magnitude of the external biases is completely unknown (or, at least, poorly known), then the model of Eq. (1) should be augmented by additional parametric terms that describe the systematic effects in the input data. An integrated LS estimation procedure can then lead to optimal estimates for the original model parameters **x** and the additional bias-related parameters that will be included in the data adjustment process.

Bias modeling and elimination has been a topic of continuous research interest in geodesy, by both theoreticians and practitioners. An extensive discussion, related mostly to the mathematical details of bias treatment in geodetic data analysis is given in *Kukuča* (1987). Some interesting aspects from the theory of nuisance parameter elimination in LS adjustment models can be found in *Schaffrin and Grafarend* (1986). A comparison of different approaches for dealing with systematic effects that arise from the integration of heterogeneous geodetic data sets is given in *Schaffrin and Baki-Iz* (2001). The problem of bias modeling and elimination has also been treated in the framework of various particular fields of geodetic research, including gravimetric data processing (e.g. *Kubáčková and Kubáček* 1993, *Harnisch* 1993), altimetric data processing (e.g. *Tscherning and Knudsen* 1986, *Gaspar et al.* 1994), GPS data processing (e.g. *Satirapod et al.* 2003, *Jia et al.* 2000) and global gravity field modelling (e.g. *Lerch* 1991).

Motivated by the fact that most geodetic observations are carried under complex physical conditions which do not correspond exactly to the standard mathematical models that we often employ for their analysis, the purpose of this paper is to expose an important trade-off which exists in the LS inversion of linear models that are augmented by additional (nuisance) parameters in the presence of unknown systematic errors in the input data. In particular, it is shown that if a linear model of full rank is extended by a single scalar parameter to account for a common bias in the data, then the LS estimation accuracy of *all* the other model parameters automatically worsens. Two different cases are considered in this study, depending whether *some* or *all* of the observations are affected by a common unknown bias. Some numerical examples are also given to demonstrate the significance of this accuracy degradation in the geodetic practice.

2. LS parameter estimation with different types of models

In this section, we examine the results obtained from the LS inversion of three different models for the linear(-ized) system of observation equations. Specifically, we study the cases where *none*, *all*, or *some* of the input measurements are affected by a common unknown bias.

2.1 Linear model with non-biased data

This is the archetypical case whose formulation and description have already been given in the previous section; see Eqs. (1) and (2). For the sake of economy in our discussion, we shall assume that the design matrix **A** has full column rank and that the CV matrix **C** is fully known. In such a case, the LS inversion of the system of observation equations in Eq. (1) leads to the well-known optimal solution

$$\hat{\mathbf{x}} = (\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{y} \tag{3}$$

The quality of the above unbiased estimate is generally described by its associated CV matrix

$$\mathbf{C}_{\hat{\mathbf{x}}} = E\{(\hat{\mathbf{x}}-\mathbf{x})(\hat{\mathbf{x}}-\mathbf{x})^{\mathrm{T}}\} = (\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1} \tag{4}$$

whose diagonal elements give a measure of the statistical accuracy for the estimated model parameters $\hat{\mathbf{x}}$.

2.2 Linear model with uniformly biased data – Case I

In the presence of a common unknown bias in the measurements **y**, the functional model of Eq. (1) should be modified as follows

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \beta\mathbf{s} + \mathbf{v} \tag{5}$$

where β is a scalar parameter that describes the systematic offset in the input data, and **s** corresponds to a vector with all its values equal to one, $\mathbf{s} = [1 \ \ldots \ 1]^{\mathrm{T}}$. The data noise **v** follows the same stochastic behaviour as in the case of the simple model with non-biased data.

Assuming that the partitioned matrix [**A** **s**] has full column rank, the LS inversion of the augmented system in Eq. (5) leads to the following optimal estimates for the original model parameters **x** and the bias parameter β

$$\hat{\beta} = k\,\mathbf{s}^{\mathrm{T}}\mathbf{C}^{-1}(\mathbf{y} - \mathbf{A}\mathbf{x}^{b}) \tag{6}$$

$$\hat{\mathbf{x}} = \mathbf{x}^{b} - \hat{\beta}\,\mathbf{q} \tag{7}$$

where the auxiliary quantities $\mathbf{x}^{b}$, $\mathbf{q}$ and k are defined by the expressions

$$\mathbf{x}^{b} = (\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{y} \tag{8}$$

$$\mathbf{q} = (\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{s} \tag{9}$$

$$k = \left(\mathbf{s}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{s} - \mathbf{q}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})\mathbf{q}\right)^{-1} \tag{10}$$

Note that the term $\mathbf{x}^{b}$ corresponds to the LS estimate that we would obtain if we ignored the bias presence in the input data. Applying the (co-)variance propagation law to the unbiased estimators of Eqs. (6) and (7), we obtain the CV matrix of the estimated parameter vector $\hat{\mathbf{x}}$

$$\mathbf{C}_{\hat{\mathbf{x}}} = E\{(\hat{\mathbf{x}}-\mathbf{x})(\hat{\mathbf{x}}-\mathbf{x})^{\mathrm{T}}\} = (\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1} + k\,\mathbf{q}\mathbf{q}^{\mathrm{T}} \tag{11}$$

and also the variance of the estimated bias parameter $\hat{\beta}$

$$\sigma_{\hat{\beta}}^2 = E\{(\hat{\beta}-\beta)^2\} = k \tag{12}$$

The last two formulae can easily be derived by using standard rules of matrix calculus.

2.3 Linear model with uniformly biased data – Case II

A useful generalization of the results given in the previous section is obtained if we assume that only *a part* of the data is affected by a common unknown bias. In this case, the standard model of Eq. (1) should be modified in the following way

$$\underbrace{\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix}\mathbf{A}_1\\ \mathbf{A}_2\end{bmatrix}}_{\mathbf{A}}\mathbf{x} + \beta\begin{bmatrix}\mathbf{s}\\ \mathbf{0}\end{bmatrix} + \underbrace{\begin{bmatrix}\mathbf{v}_1\\ \mathbf{v}_2\end{bmatrix}}_{\mathbf{v}} \tag{13}$$

The total data vector $\mathbf{y}$ is now partitioned into two subsets, $\mathbf{y}_1$ and $\mathbf{y}_2$, whereas the submatrices $\mathbf{A}_1$ and $\mathbf{A}_2$ provide the corresponding partition for the design matrix $\mathbf{A}$ of the initial model in Eq. (1). Only the first group of measurements is affected by a common unknown bias, which is denoted again by the scalar parameter β. The symbol $\mathbf{s}$ denotes a vector of ones, whereas $\mathbf{0}$ is the zero vector. Finally, the subvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ contain the random errors of the two corresponding data subsets, with the following stochastic characteristics

$$E\{\mathbf{v}_1\} = E\{\mathbf{v}_2\} = \mathbf{0} \tag{14}$$

$$E\{\mathbf{v}\mathbf{v}^{\mathrm{T}}\} = \mathbf{C} = E\{\begin{bmatrix}\mathbf{v}_1\mathbf{v}_1^{\mathrm{T}} & \mathbf{v}_1\mathbf{v}_2^{\mathrm{T}}\\ \mathbf{v}_2\mathbf{v}_1^{\mathrm{T}} & \mathbf{v}_2\mathbf{v}_2^{\mathrm{T}}\end{bmatrix}\} = \begin{bmatrix}\mathbf{C}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{C}_2\end{bmatrix} \tag{15}$$

For simplicity, we assume zero cross-correlation in the data noise between the two groups of measurements. Assuming also that the augmented model of Eq. (13) has full rank, its LS inversion yields the following optimal solution

$$\hat{\beta} = \lambda\,\mathbf{s}^{\mathrm{T}}\mathbf{C}_1^{-1}(\mathbf{y}_1 - \mathbf{A}_1\mathbf{x}^b) \tag{16}$$

$$\hat{\mathbf{x}} = \mathbf{x}^b - \hat{\beta}\,\mathbf{p} \tag{17}$$

where the auxiliary quantities $\mathbf{p}$ and λ are defined by the equations

$$\mathbf{p} = (\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1}\mathbf{A}_1^{\mathrm{T}}\mathbf{C}_1^{-1}\mathbf{s} \tag{18}$$

$$\lambda = \left(\mathbf{s}^{\mathrm{T}}\mathbf{C}_1^{-1}\mathbf{s} - \mathbf{p}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})\mathbf{p}\right)^{-1} \tag{19}$$

The auxiliary term $\mathbf{x}^b$ has already been defined in Eq. (8). Applying the (co-)variance propagation law to the unbiased estimators of Eqs. (16) and (17), we get the CV matrix for the estimated parameter vector $\hat{\mathbf{x}}$

$$\mathbf{C}_{\hat{\mathbf{x}}} = E\left\{(\hat{\mathbf{x}} - \mathbf{x})(\hat{\mathbf{x}} - \mathbf{x})^{\mathrm{T}}\right\} = (\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1} + \lambda\,\mathbf{p}\mathbf{p}^{\mathrm{T}} \tag{20}$$

and also the variance of the estimated bias parameter $\hat{\beta}$

$$\sigma_{\hat{\beta}}^2 = E\left\{(\hat{\beta} - \beta)^2\right\} = \lambda \tag{21}$$

All the previous formulae can easily be verified by using standard rules of matrix calculus. A brief discussion on the estimation performance that is achieved by employing each of the three previous models follows in the next section.

3. Remarks

By comparing the results from Eqs. (4), (11) and (20), it is seen that the accuracy of the estimated parameters $\hat{\mathbf{x}}$ becomes worse when an overparameterized model is used for the LS adjustment of the input data. Indeed, in the cases where biased measurements are processed with an augmented linear model, the CV matrix of $\hat{\mathbf{x}}$ "increases" by the amount $k\mathbf{q}\mathbf{q}^{\mathrm{T}}$ (or $\lambda\mathbf{p}\mathbf{p}^{\mathrm{T}}$), with respect to the result obtained from the LS adjustment with non-biased data. Since the scalar quantities k and λ are positive, the diagonal elements of $\mathbf{C}_{\hat{\mathbf{x}}}$ in Eqs. (11) and (20) will always be larger than the corresponding elements of $\mathbf{C}_{\hat{\mathbf{x}}}$ in Eq. (4).

It can thus be concluded that the presence of an additional bias parameter in a linear model causes an overall degradation in the LS estimation accuracy for the rest of the model parameters. Note that the noise level is assumed to be the same in the cases of non-biased and biased data.

Furthermore, it is interesting to point out that the extent of the accuracy reduction is completely independent of the magnitude of the bias that has affected the input data. That is evident from the fact that the terms $k\mathbf{q}\mathbf{q}^{\mathrm{T}}$ and $\lambda\mathbf{p}\mathbf{p}^{\mathrm{T}}$ do not depend on β. Hence, any constant data bias (regardless of how small or large its true value really is) that is taken into account by adding an extra parameter within the LS adjustment framework, will cause the same degradation on the estimation accuracy for the rest of the model parameters.

4. Numerical examples

Two simple examples are given in this section to demonstrate the practical significance of the theoretical results that were previously discussed.

4.1 Gravity network adjustment

The first example refers to the LS adjustment of a simulated gravity network, as shown in Figure 1. The observables consist of eight gravity differences that are specified explicitly in Table 1, along with their corresponding accuracy. Note that all observations are considered uncorrelated in this example.

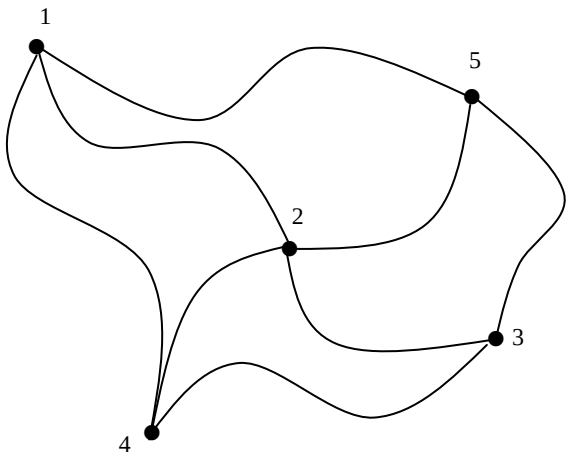

Figure 1. A simulated gravity network.

Table 1. The simulated observables, along with their corresponding accuracy, for the gravity network of Figure 1.

i	**Observables**	**Measurement accuracy** σ_i (in μgal)
1	$\Delta g_{12} = g_2 - g_1$	11
2	$\Delta g_{23} = g_3 - g_2$	10
3	$\Delta g_{34} = g_4 - g_3$	15
4	$\Delta g_{14} = g_4 - g_1$	16
5	$\Delta g_{51} = g_1 - g_5$	17
6	$\Delta g_{52} = g_2 - g_5$	16
7	$\Delta g_{53} = g_3 - g_5$	15
8	$\Delta g_{24} = g_4 - g_2$	10

Assuming a known gravity value for *point 1*, which is held fixed for the purpose of datum definition, the unknown parameter vector **x** in the network adjustment includes the absolute gravity values g_i for the four other network points. Using the information provided in Table 1, the CV matrix $\mathbf{C}_{\hat{\mathbf{x}}}$ of the estimated parameters, for the case of non-biased data, can be computed from Eq. (4). The result is

$$\mathbf{C}_{\hat{\mathbf{x}}} = \begin{bmatrix} 74.225 & 63.862 & 55.921 & 48.589 \\ 63.862 & 116.945 & 63.138 & 65.136 \\ 55.921 & 63.138 & 100.259 & 42.254 \\ 48.589 & 65.136 & 42.254 & 125.248 \end{bmatrix} [\mu gal^2] \tag{22}$$

If a common unknown bias is assumed to exist in all observations, then the extended model of Eq. (5) must be used for the LS network adjustment. In this case, the CV matrix $\mathbf{C}_{\hat{\mathbf{x}}}$ for the estimated gravity values is obtained from Eq. (11), which yields the result

$$\mathbf{C}_{\hat{\mathbf{x}}} = \begin{bmatrix} 104.602 & 126.932 & 147.301 & 22.803 \\ 126.932 & 247.795 & 252.794 & 11.617 \\ 147.301 & 252.794 & 375.151 & -35.318 \\ 22.803 & 11.617 & -35.318 & 147.138 \end{bmatrix} [\mu gal^2] \tag{23}$$

whereas the standard deviation of the estimated bias parameter $\hat{\beta}$ according to Eq. (12) is

$$\sigma_{\hat{\beta}} = 10.81 \ \mu gal \tag{24}$$

Table 2. Standard deviation for the estimated gravity values in the network of Figure 1.

Model parameters (x)	**Estimation accuracy obtained from the *simple* LS adjustment model**	**Estimation accuracy obtained from the *extended* LS adjustment model**
g_2	8.62 μgal	10.23 μgal
g_3	10.81 μgal	15.74 μgal
g_4	10.01 μgal	19.37 μgal
g_5	11.19 μgal	12.13 μgal

By comparing the diagonal elements from the CV matrices in Eqs. (22) and (23), it can be deduced that an average degradation of about 4.2 μgal in the accuracy of the estimated gravity values occurs, when the augmented model is used. The actual decrease in the estimation accuracy for each individual gravity value is analytically shown in Table 2.

4.2 Gravity determination from repeated free-fall experiments

For our second example, we consider the determination of absolute gravity through repeated observations of a free-falling mass *m* at a certain point *P* (see Figure 2).

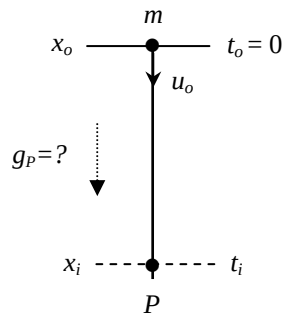

Figure 2. Free-fall experiment for absolute gravity determination.

Assuming a homogeneous gravity field within the region of the experiment, we have the following well-known equation

$$x_i = x_o + u_o t_i + \frac{1}{2} g t_i^2 \tag{25}$$

which describes the Newtonian motion of the test mass. The quantities x_o and u_o denote the known initial position and velocity, x_i is the observed vertical position of the test mass at the time instant t_i, and g corresponds to the gravity acceleration at the observation point.

The model of Eq. (25) provides the basic observation equation that we can use to determine the gravity value g via the LS adjustment of multiple repeated measurements $\{ x_i \}$ of the vertical position of the test mass. For simplicity, we assume that (i) the time values t_i are known without any error, and (ii) the initial conditions are such that $x_o = 0$ and $u_o = 0$.

Our simulated experiment consists of twenty-five different drops of the test mass. The recorded times t_i are shown in Table 3, along with the corresponding accuracy for each of the observed positions x_i. Using the information from Table 3, we can easily construct the design matrix **A** and the data noise CV matrix **C** for the underlying adjustment problem. The standard deviation of the estimated gravity value is then computed by the general formula of Eq. (4), which yields the result

$$\sigma_{\hat{g}} = 0.281 mgal \tag{26}$$

Table 3. Accuracy of the simulated observables x_i and the corresponding time intervals t_i for the free-falling mass experiment.

i	t_i (in sec)	σ_{x_i} (in mm)
1	0.516	0.001
2	0.474	0.002
3	0.487	0.002
4	0.502	0.002
5	0.467	0.002
6	0.491	0.002
7	0.487	0.001
8	0.501	0.001
9	0.512	0.001
10	0.493	0.003
11	0.479	0.003
12	0.492	0.003
13	0.514	0.002
14	0.505	0.002
15	0.496	0.002
16	0.493	0.003
17	0.495	0.003
18	0.510	0.003
19	0.501	0.002
20	0.486	0.002
21	0.483	0.003
22	0.495	0.003
23	0.510	0.002
24	0.511	0.002
25	0.495	0.002

In the case where a common unknown bias exists in all available observations x_i, the following extended model must be employed for the optimal determination of the gravity acceleration g

$$x_i = x_o + u_o t_i + \frac{1}{2} g t_i^2 + \beta \tag{27}$$

where β corresponds to the common bias in the measurements. The standard deviation of the LS estimated gravity acceleration is now computed from the general formula of Eq. (11), which gives the result

$$\sigma_{\hat{g}} = 5.473\, mgal \tag{28}$$

whereas the standard deviation of the bias estimate, according to Eq. (12), is

$$\sigma_{\hat{\beta}} = 6.83 \times 10^{-3}\ mm \tag{29}$$

From the numerical results given in Eqs. (26) and (28), it is evident that the standard deviation of the estimated gravity acceleration $\hat{g}$ increases by almost a factor of 20, when a single bias parameter is included in the LS adjustment model for the repeated observations $\{ x_i \}$!

5. Conclusions

The elimination of a common unknown bias from a set of geodetic measurements can be achieved by extending the classic linear(-ized) model $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}$ with an additional scalar parameter that describes the systematic offset in the data values. In this way, optimal unbiased estimates for the model parameters $\mathbf{x}$ can be obtained when employing the standard LS estimation criterion with a biased data set.

Nevertheless, as it was explained in this paper, the statistical accuracy of the estimated model parameters $\hat{\mathbf{x}}$ will always be worse when an additional bias parameter is included in the LS adjustment process. Geodesists should be aware of this important trade-off, since in many applications the theoretical models often need to be augmented with additional nuisance parameters in order to account for external disturbances in the observations.

References

Dermanis A, Rummel R (2000) Data analysis methods in geodesy. In A. Dermanis, A. Gruen and F. Sanso (eds) *Geomatic Methods for the Analysis of Data in Earth Sciences*. Lecture Notes in Earth Sciences Series, 95: 17-92, Springer Verlag, Berlin Heidelberg.

Gaspar P, Ogor F, Le Traon P-Y, Zanife O-Z (1994) Estimating the sea state bias of the TOPEX and POSEIDON altimeters from crossover differences. *J Geoph Res*, 99(C12): 24,981-24,994.

Harnisch G (1993) Systematic errors affecting the accuracy of high precision gravity measurements. *IAG Symposia*, vol. 112, Springer-Verlag, Berlin Heidelberg, pp. 200-204.

Jia M, Tsakiri M, Stewart M (2000) Mitigating multipath errors using semi-parametric models for high precision static positioning. *IAG Symposia*, vol. 121, Springer-Verlag, Berlin Heidelberg, pp. 393-398.

Kubáčková L, Kubáček L (1993) A group of gravimeters, stochastical problems and their solution. *IAG Symposia*, vol. 112, Springer-Verlag, Berlin Heidelberg, pp. 275-278.

Kukuča J (1987) Systematic influences and their elimination. In: Kubáčková L, Kubáček L and Kukuča J: *Probability and statistics in geodesy and geophysics*. Elsevier, Amsterdam, pp. 257-289.

Lerch FJ (1991) Optimum data weighting and error calibration for estimation of gravitational parameters. *Bull Geod*, 65: 44-52.

Satirapod C, Wang J, Rizos C (2003) Comparing different global positioning system data processing techniques for modeling residual systematic errors. *J Surv Eng*, 129(4): 129-135.

Schaffrin B, Grafarend E (1986) Generating classes of equivalent linear models by nuisance parameter elimination. *Manuscr Geod*, 11: 262-271.

Schaffrin B, Baki-Iz H (2001) Integrating heterogeneous data sets with partial inconsistencies. *IAG Symposia*, vol. 123, Springer-Verlag, Berlin Heidelberg, pp. 49-54.

Tscherning CC, Knudsen P (1986) Determination of bias parameters for satellite altimeter by least-squares collocation. *Proceedings of the 1st Hotine-Marussi Symposium on Mathematical Geodesy*, Rome, June 3-6, 1985, pp. 833-851.

Quick-look outlier detection for GOCE gravity gradients

Johannes Bouman
SRON National Institute for Space Research
Sorbonnelaan 2, 3584 CA Utrecht, The Netherlands

December 20, 2004

Abstract

GOCE will be the first satellite ever to measure the second order derivatives of the Earth's gravitational potential in space. With these measurements it is possible to derive a high accuracy and resolution gravitational field if systematic errors and/or outliers have been removed to the extent possible from the data. It is necessary to detect as many outliers as possible in the data pre-processing because undetected outliers may lead to erroneous results when the data are further processed, for example in the recovery of a gravity field model. Outliers in the GOCE gravity gradients, as they are likely to occur in the real observations, will be searched for and detected in the processing step preceding gravity field analysis.

As the diagonal gravity gradients are the main gradient observables for GOCE, three methods are discussed to detect outliers in these gradients. The first is the tracelessness condition, that is, the sum of the diagonal gradients has to be zero. The second method compares GOCE gravity gradients with model or filtered gradients. Finally, along track interpolation of gravity gradient anomalies is discussed. Since the difference between an interpolated value and a measured value is large when outliers are present, along track interpolation is known to be suitable for outlier detection. The advantages and disadvantages of each method are discussed and it is shown that the final outlier detection algorithm, which is a combination of the three methods, is able to detect almost all outliers while the number of falsly detected outliers remains small.

Key words. Outliers · GOCE mission · Gradiometry · Statistical tests

1 Introduction

Outlier detection is one of the important tasks in the GOCE data pre-processing. In this paper, the focus will be on the gravity gradients (GG). The outliers themselves may point to possible instrument problems, while undetected outliers may lead to erroneous results when the data are further processed, for example in the external calibration of the gravity gradients, in the gravity field analysis or in the error assessment. It is therefore important to detect as many outliers as possible in the data pre-processing. A restriction in the context of quick-look data processing is the time required to detect outliers. The outlier detection algorithm that is implemented should have a short run-time, while the intervention by an operator should be minimal, that is, the s/w has to be fully automated.

First, the outlier detection is described in the pre-processing context. Second, several outlier detection methods are discussed and compared. Finally, the outlier detection methods are tested in a simulation study and the overall algorithm is discussed. Alternatives are discussed in, for example (Albertella et al. 2000; Bouman et al. 2004a; Kern et al. 2004; Tscherning 1991).

2 Pre-processing

Since the main goal of the GOCE mission (expected launch in August 2006) is to provide unique models of the Earth's static gravity field (ESA 1999), the GOCE gravity gradients need to be corrected for temporal gravity field variations such as tides. Furthermore, even after in-flight calibration the observations will be contaminated with stochastic and systematic errors. Systematic errors include GG scale factor errors and biases (Cesare 2002) which one tries to correct for in the external calibration step (see e.g. Bouman et al. 2004b). Also outliers in the GOCE gravity gradients need to be searched for and detected in the Level 2 pre-processing step. The steps for quick-look pre-processing are:

1. corrections for temporal gravity field variations;
2. outlier detection and correction;
3. external calibration and error assessment;
4. iteration of steps 2 and 3.

3 Outlier detection

We will consider time series of gravity gradients

$$V_{\eta\eta}(\tau_i),\ i = 1 : m \tag{1}$$

with $\tau_i - \tau_{i-1} = 1$ s, $\eta\eta = xx, yy$ or zz and m the number of observations.

First, data snooping is discussed when a model of condition equations is used, which will be the starting point for our outlier detection. Then we will discuss the tracelessness condition which will be the baseline method for outlier detection. The comparison of measured gradient with model gradients (gradient anomalies) and the interpolation of the gradient anomalies are discussed as well.

3.1 Data snooping

Let's assume that the $m \times 1$ vector $\underline{y}$ contains the diagonal gravity gradients which errors are normally distributed with known error variance matrix Q_y:

$$\underline{y} \sim N(E\{\underline{y}\}, Q_y) \tag{2}$$

with E the expectation operator. All single observations will be tested for outliers. The hypothesis

$$H_0 :\ B^T E\{\underline{y}\} = 0 \tag{3}$$

will be tested against

$$H_A :\ B^T E\{\underline{y}\} = c_t \nabla,\ \nabla \neq 0 \tag{4}$$

where B^T is the condition equation matrix, $c_t = B^T c_y$ and c_y is a unit vector with 1 at row i if the i-th observations is to be tested, and ∇ is an outlier with unknown size. In the condition equation (3), the matrix B^T has b rows, the number of conditions, and it has m columns, the number of observations. It can be shown that H_0 will be rejected if, see (Teunissen 2000):

$$w < -k_\alpha \text{ or } w > k_\alpha \tag{5}$$

with

$$\underline{w} = \frac{c_t^T Q_t^{-1} \underline{t}}{\sqrt{c_t^T Q_t^{-1} c_t}} \tag{6}$$

where $\underline{t} = B^T \underline{y}$ is the vector of misclosures, $Q_t = B^T Q_y B$, and k_α is the critical value which depends on the significance level α. The random variable $\underline{w}$ is the w-teststatistic and has a standard normal distribution under H_0. A disadvantage of data snooping may be that it is computationally intensive. In general Q_t is a full matrix and its inverse has to be computed.

One can make two types of errors in hypothesis testing (Teunissen 2000). *Type I error*: rejection of H_0 when H_0 is true, that is, an outlier is detected, but there is no outlier; *Type II error*: acceptance of H_0 when H_0 is false, that is, an outlier goes undetected.

3.2 Tracelessness condition

The sum of the diagonal gravity gradients has to be zero which is called Laplace's equation or the tracelessness condition (Heiskanen and Moritz 1967). For observed gradients one has

$$E\{V_{xx} + V_{yy} + V_{zz}\} = 0 \tag{7}$$

and

$$Q_y = \begin{pmatrix} Q_{V_{xx}} & 0 & 0 \\ 0 & Q_{V_{yy}} & 0 \\ 0 & 0 & Q_{V_{zz}} \end{pmatrix} \tag{8}$$

under the assumption that there is no error correlation between the different gradients. One problem is that the gravity gradients suffer from systematic errors before external calibration of which biases and scale factor errors are the most important. For GG the effect of a scale factor error is the largest at a frequency of 0 Hz and also the bias is manifest at this frequency. Therefore, the median of the sum of the diagonal gradients over the time interval considered is subtracted in (7). Since the mean is more sensitive to outliers than the median, the former is not used.

The Q_t-matrix, needed in the test (6), is

$$Q_t = Q_{V_{xx}} + Q_{V_{yy}} + Q_{V_{zz}}. \tag{9}$$

If the along-track error correlation is neglected, then the $Q_{V_{\eta\eta}}$ are diagonal and the w-teststatistic becomes

$$w(i) = \frac{V_{xx}(i) + V_{yy}(i) + V_{zz}(i) - \text{median}}{\sqrt{\sigma^2_{V_{xx}}(i) + \sigma^2_{V_{yy}}(i) + \sigma^2_{V_{zz}}(i)}} \tag{10}$$

with $i = 1 : m$ and m the number of observation points. It is well known that the GOCE GG along-track error correlation is high. Nevertheless, it may be that for outlier detection one can neglect this correlation. Moreover, in the simulation study the $Q_{V_{\eta\eta}}$ matrices are taken as scaled unit matrices. If the along-track error correlation would not have been neglected, the Q_t matrix would have become full, which makes the computation of its inverse much more problematic. One could of course work with distinct patches of, for example, 50 observations neglecting the correlation between patches. However, the choice of the patch size is arbitrary and the observations within one patch would be treated differently, which leads to an incoherent test method.

The advantage of the tracelessness condition is that the signal-to-noise ratio (SNR) is very small. In fact, the SNR can not be smaller, which means that outliers that are well above the noise can be detected easily. A disadvantage is that with this method one can not discriminate between outliers in V_{xx}, V_{yy} and V_{zz}.

3.3 Gravity gradient anomalies

The idea is to predict gravity gradients in the GOCE orbit points from a global gravity field model and to compare them with the GOCE GG. The condition equation is

$$E\{V_{\eta\eta} - U_{\eta\eta}\} = 0 \tag{11}$$

with error matrix

$$Q_y = \begin{pmatrix} Q_{V_{\eta\eta}} & 0 \\ 0 & Q_{U_{\eta\eta}} \end{pmatrix} \tag{12}$$

where the error of the model gradients is described by $Q_{U_{\eta\eta}}$. Among others, this error depends on the accuracy of the global model, the omission error, attitude errors and the accuracy of the orbit.

In this case, the Q_t-matrix is

$$Q_t = Q_{V_{\eta\eta}} + Q_{U_{\eta\eta}}. \tag{13}$$

If it is assumed that Q_t is diagonal (no along-track error correlation), then the w-teststatistic becomes

$$w(i) = \frac{V_{\eta\eta}(i) - U_{\eta\eta}(i) - \text{median}}{\sqrt{\sigma^2_{V_{\eta\eta}}(i) + \sigma^2_{U_{\eta\eta}}(i)}} \tag{14}$$

for $i = 1 : m$. Also here the median is subtracted. Furthermore, in the simulation study the Q matrices are taken as scaled unit matrices.

One problem of condition (11) may be that the omission error is large, for example when a GRACE-only model is used, or that the commission error is large, for example when OSU91A is used. Thus the total σ may be large and the test has little 'power'. However, all GG are tested separately and point wise and it is therefore an unambiguous test.

3.4 Overhauser spline interpolation

As an alternative to (11), consider the interpolation of anomalies along tracks. The interpolated gravity gradient anomaly can be compared with the 'measured' anomaly, $T_{\eta\eta} := V_{\eta\eta} - U_{\eta\eta} -$ median, and

$$E\{T^{int}_{\eta\eta} - T_{\eta\eta}\} = 0, \quad \eta\eta = xx, yy, zz \tag{15}$$

with $T^{int}_{\eta\eta}$ the interpolated value. The error of the interpolated anomalies depends on the interpolation method, the orbit accuracy and on the errors of the observed GG. To a lesser extent also the accuracy of the model gradients $U_{\eta\eta}$ has an effect on the error. There are many possible interpolation methods but only Overhauser splines are considered (Overhauser 1968). The computations are simple and fast, while the interpolation errors are small, see (Bouman and Koop 2003). For equidistant data along track the condition equations take the form

$$y_i = \frac{2}{3}(y_{i-1} + y_{i+1}) - \frac{1}{6}(y_{i-2} + y_{i+2}). \tag{16}$$

If there are m observations, then i may take the values $3 : m - 2$. (The weights given in (Bouman and Koop 2003) are incorrect. The interpolation errors reported there become even smaller using the correct weights.)

Under the assumption that the error matrices are scaled unit matrices, the w-teststatistic may be approximated with

$$w(i) = \frac{1}{\sqrt{70}\sigma_{T_{\eta\eta}}} \sum_{k=i-2}^{i+2} a_k T_{\eta\eta}(k) \tag{17}$$

Table 1: *Noise, outlier and gravity gradient anomaly ($V_{\eta\eta} - U_{\eta\eta}$) properties, values in [mE].*

Small set	(86,351 pts)	V_{xx}	V_{yy}	V_{zz}
noise	mean	1443.7	-805.2	2248.9
	rms	2.2	4.4	5.7
outliers	mean	0.5	0.3	0.0
	rms	58.9	27.9	52.7
anomalies	mean	0.0	1.5	-1.5
	rms	36.4	35.3	58.9
Large set	(5,097,835 pts)	V_{xx}	V_{yy}	V_{zz}
noise	mean	0.0	0.0	0.0
	rms	10.1	2.7	10.0
outliers	mean	0.0	0.0	0.0
	rms	78.5	78.5	78.5
anomalies	mean	0.0	-0.4	0.4
	rms	37.2	35.3	60.0

Table 2: *Type I error for case 1 in % (no outliers, critical value is ± 2);* T – *tracelessness condition,* M – *model gradients,* S – *spline interpolation,* TMS – T + M *or* T + S.

	Small data set			Large data set		
Method	V_{xx}	V_{yy}	V_{zz}	V_{xx}	V_{yy}	V_{zz}
T	0.0	0.0	0.0	4.7	4.7	4.7
M	6.2	6.0	5.9	6.3	6.4	6.2
S	0	0	0	0	0	0
TMS	0	0	0	0.3	0.3	0.3

for $i = 3 : m - 2$, with weights $a_{i-2} = a_{i+2} = -1$, $a_{i-1} = a_{i+1} = 4$ and $a_i = -6$. ($\sum a_k^2 = 70$)

The advantage of this method is that each GG can be tested separately. A disadvantage is that several consecutive points are combined, which may hinder the identification of points with outliers (masking).

4 Simulation study

Two data sets with different characteristics were studied. One is a small data set with a length of 1 day which contains various types of outliers. The second data set has a length of 59 days and contains single and bulk outliers. These gradients allow for gravity field analysis (GFA).

The first data set used in this study consists of the diagonal gravity gradients V_{xx}, V_{yy} and V_{zz} which were simulated using EGM96 (Lemoine et al. 1998) for a 1 day orbit with a sampling rate of 1 s. Simulated, correlated noise was added to the signals, the data statistics are given in Table 1. The model gradients were generated using OSU91A (Rapp et al. 1991). The second data set used in this study also consists of the diagonal gravity gradients which were simulated using OSU91A for a 59 day orbit with a sampling rate of 1 s (over five million data points). Simulated, correlated noise was added to the signals. In this case, model gradients were generated using EGM96.

Errors other than outliers and simulated noise were not considered in this study, that is, orbit errors, omission errors, etc. are all zero.

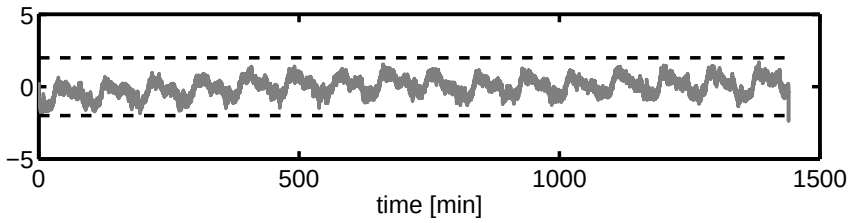

Figure 1: *Test values for tracelessness condition, case 1, small data set. The dashed lines denote the critical value* ± 2.

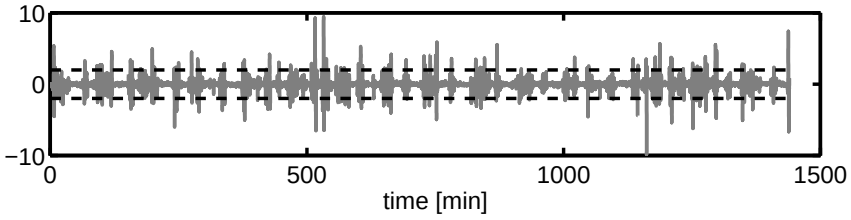

Figure 2: *Test values for* $V_{xx} - U_{xx}$*, case 1, small data set.* V_{xx} *are simulated GOCE gradients, whereas* U_{xx} *are model gradients. Test values for* V_{yy} *and* V_{zz} *are similar to* V_{xx}*. The dashed lines denote the critical value* ± 2.

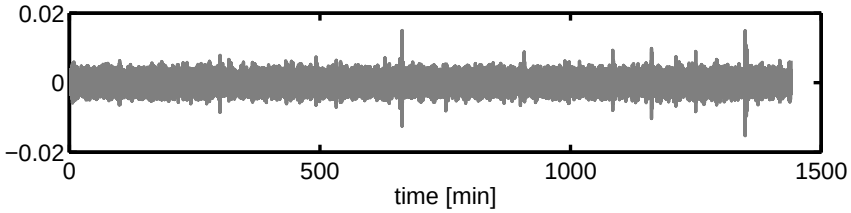

Figure 3: *Test values for* T_{xx} *spline interpolation, case 1, small data set. Test values for* T_{yy} *and* T_{zz} *are similar to* T_{xx}.

4.1 Case 1: no outliers

A first test was done that used the noisy gradients without any outliers (case 1). Fig. 1 - 3 show the w-test values for the tracelessness condition, gradient anomalies and spline interpolation respectively (small data set), while the type I error is summarised in Table 2. Given the critical value of $k = 2$, approximately 4.6% of the observations should be rejected although they are correct. For the tracelessness condition, however, the type I error is 0% for the small data set, while it is according to the expected value for the large data set. The former may be due to the error correlation between the different simulated diagonal gradients. In the small data set these errors are heavily correlated, which is neglected, whereas there is no error correlation between different gradients for the large data set. The type I error is also 0% for the spline interpolation for both data sets. Both the small and the large data sets have errors with a high spatial correlation along tracks. These long wavelength errors, however, cancel in the spline interpolation as it is a local interpolation method. This could explain the small type I error.

The model gradients have a large type I error as it is dominated by the model error, that is, the difference between EGM96 and OSU91A. The type I error is probably larger than expected because we have used a simple scaled unit matrix as error covariance matrix. Despite the some-

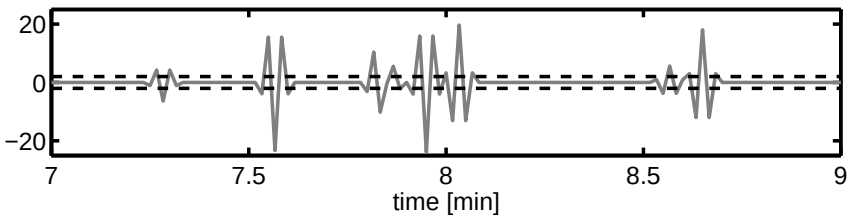

Figure 4: *Test values for T_{xx} spline interpolation, case 2a. Panel zooms in on $t = 7 - 9$ min. The dashed lines denote the critical value ± 2.*

what larger type I error, model gradients may be usefull. First of all, it may be that at the time GOCE flies a more accurate gravity field model is available, which would reduce the type I error. Secondly, with the gradient anomalies one can test the individual gradients point wise. In contrast, with the trace condition one tests point wise the sum of three gradients, while with spline interpolation individual gradients are tested on an interval. Thus the three methods are complementary.

A combination of the three methods gives good results which is shown in Table 2. In this case combined means that if an outlier is detected by the tracelessness condition and if it is confirmed by a 2nd method, an outlier is flagged. The type I error is close to zero for all gradients.

4.2 Case 2: outliers on V_{xx}, V_{yy} and V_{zz}

To the small data set outliers with the following characteristics were added (case 2a): A total of 3891 randomly distributed single outliers for V_{xx} with an absolute size varying between 0.07 E and 0.1 E; An offset of 0.5 E during one minute ($\tau = 20 - 79$ s) for V_{yy} and a bulk of outliers during six minutes ($\tau = 50000 - 50359$ s) with an absolute size varying between 0.07 E and 0.1 E; A total of 994 'twangs', randomly distributed, for V_{zz}, that is, an outlier at $\tau = \tau_i$ is followed by an outlier with opposite sign and of the same size at $\tau = \tau_{i+1}$. In total $2 * 994 = 1988$ outliers with absolute size between 0.07 E and 0.1 E. For the large data set 83153 outliers were added to all three gradients with an absolute size varying between 0.05 E and 1.8051 E (case 2b). The outliers were randomly distributed single outliers as well as bulk outliers, see Table 1 for data statistics.

The tracelessness condition detects almost all outliers, see Table 3 and 4. If we would use the tracelessness condition only and no other method then one cannot discriminate which diagonal GG contains an outlier and all three GG would be flagged if one of them does contain an outlier, which leads to large type I errors.

Most of the outliers are detected when gradient anomalies are used, but the type II error can be relatively large. Only 3/4 of the V_{zz} outliers of the small data set are detected for example, which is caused by the larger GOCE GG error and the larger difference between the 'true' GG and the model GG. The type I error is of course at the level of case 1, see Table 3 and 4. As an alternative to model gradients, one could also filter the GOCE GG with outliers and use these as 'model gradients'. A 2nd order low-pass Butterworth filter has been used with a cut-off frequency of 0.2 Hz. Except for the gradients with an offset, the method with filtered gradients detects most of the outliers. The type I error can be large because low-pass filtering not only reduces the size of the outliers, but redistributes their power over neighbouring points as well. An offset tends to cancel and is therefore hard to detect with filtered gradients.

With spline interpolation the major part of the V_{xx} outliers (case 2a) is detected, but many valid observations are flagged as outliers, see Table 3. The type I error is large because one outlier may affect five consecutive w-test values, see Fig. 4. As an alternative to this outlier detection in one step, one could use an iterative procedure, that is, reject the global maximum, replace the

Table 3: *Detected outliers for case 2a in % (outliers on all three diagonal gradients, small data set, critical value is* ± 2*);* T – *tracelessness condition,* M – *model gradients,* F – *filtered gradients,* S – *spline interpolation,* TMS – T + M *or* T + S, TFS – T + F *or* T + S.

	V_{xx}		V_{yy}		V_{zz}	
Method	correct	type I	correct	type I	correct	type I
T	99.9	2.6	99.8	6.7	99.9	4.8
M	93.6	5.9	92.6	6.0	76.9	5.8
F	99.8	23.5	84.5	0.0	100	2.2
S	98.9	11.6	77.6	0.0	99.5	2.0
TMS	99.8	0.5	98.6	0.4	99.7	0.4
TFS	99.9	0.7	87.1	0	99.9	0.1

Table 4: *Detected outliers for case 2b in % (outliers on all three diagonal gradients, large data set, critical value is* ± 2*);* T – *tracelessness condition,* M – *model gradients,* F – *filtered gradients,* S – *spline interpolation,* TMS – T + M *or* T + S, TFS – T + F *or* T + S.

	V_{xx}		V_{yy}		V_{zz}	
Method	correct	type I	correct	type I	correct	type I
T	99.8	7.7	99.9	7.7	99.9	7.7
M	96.8	5.9	97.5	6.2	92.3	5.9
F	99.0	5.6	99.7	10.5	98.9	5.6
S	86.8	2.2	98.6	4.7	86.8	2.2
TMS	97.7	0.6	99.8	0.8	94.6	0.6
TFS	99.6	0.4	99.9	0.8	99.6	0.4

associated observation with the interpolated value, find next global maximum, etc. The iterative procedure was implemented and tested, and the type I error decreased significantly. However, also the number of detected outliers decreased, which has a negative effect on the combination solution as well. Therefore, the iterative procedure is abandoned. The spline interpolation method detects most of the bulk outliers, but is unable to detect the offset, see Table 3. An offset cancels using spline interpolation and cannot be detected directly. The type I error is small in this case because the w-test 'side lobes' drown in the bulk outliers. The results of the large data set (case 2b) suggest that the larger the GG noise the smaller the number of detected outliers. The GG V_{xx} and V_{zz} have a higher noise level than V_{yy}, while the number of detected outliers is the largest for the latter.

The last two rows of Table 3 and 4 show that a combination of three methods yields excellent results. An outlier is flagged if the tracelessness condition is confirmed by one of two methods. In general the method with model gradients detects less outliers than the method with filtered gradients. However, the latter is less suited to detect an offset. For a critical value of 2, about 99% of the outliers are detected, while the type I error is small, below 1%. The expected type I error is 4.6% which is much larger. This is likely due to the combination of the three methods.

Finally, Fig. 5 shows the gravity field anomaly differences between OSU91A and a quick-look GFA solution up to degree and order 250. As input the TFS cleaned GG were used (case 2b). The rms difference, excluding polar caps of 10°, is 9.8 mGal, which is somewhat larger than the difference when gradients without outliers are used in the GFA (6.7 mGal). When the TFS cleaned GG are used with an additional outlier search in the GFA, 100% of the outliers are detected and the rms gravity anomaly difference is only 6.9 mGal, see also (Bouman et al. 2004a).

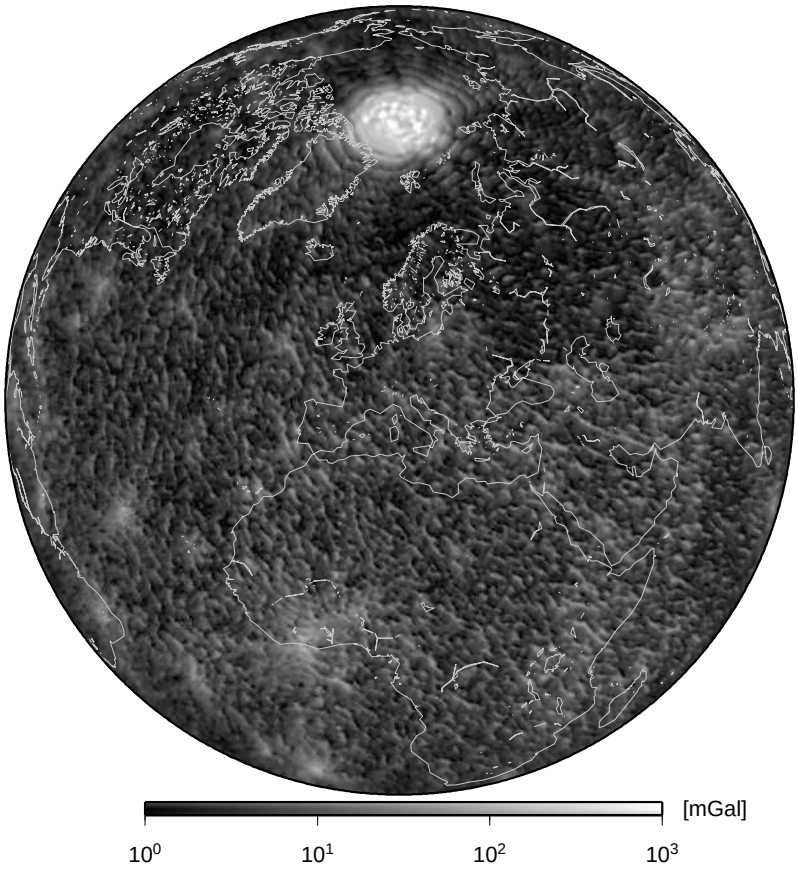

Figure 5: *Gravity anomaly differences* OSU*91*A – QL-GFA *(log scale), pre-processing outlier detection using* TFS.

5 Conclusions and outlook

The tracelessness condition is the baseline method for the quick-look outlier detection algorithm studied here. If an outlier is detected by this method and if it is confirmed by either model (or filtered) gradients and/or by spline interpolation, then an outlier is flagged. Although the individual methods have their disadvantages, their combination yields high outlier detection rates and only a small number of falsly detected outliers. The w-test, which was used, explicitely accounts for the GG errors. However, to obtain a managable solution, the error correlations were neglected and it was even assumed that the error matrices are scaled unit matrices. Despite the heavy error correlation, outliers can be very well detected. It therefore seems that the simplifications do little harm.

To further improve the performance, the spline interpolation may be replaced by, for example, least-squares prediction. The outlier detection results may also improve by taking the spatial correlation between the observables into account using least-squares collocation (LSC), see (Tscherning 1991). It may be that the turn-around time of the current GRAVSOFT implemenation of LSC (Tscherning 1974) is acceptable for operational use. This needs, however, to be studied. Future simulation studies should also include the V_{xy}, V_{xz} and V_{yz} gradients. In addition, more realistic GOCE error characteristics should be used, and other errors, such as orbit errors or attitude quaternion errors, should be accounted for.

6 Acknowledgements

This study was performed in the framework of the ESA-project GOCE High-level Processing Facility (GOCE HPF, Main Contract No. 18308/04/NL/MM). Michael Kern computed the QL-GFA solutions and provided the GFA figure. All this is gratefully acknowledged.

References

Albertella A, Migliaccio F, Sansò F, Tscherning CC (2000) Scientific data production quality assessment using local space-wise pre-processing. In H. Sünkel, editor, From Eötvös to mGal, Final Report. ESA/ESTEC contract no. 13392/98/NL/GD. European Space Agency, Noordwijk

Bouman J, Kern M, Koop R, Pail R, Haagmans R, Preimesberger T (2004a) Comparison of outlier detection algorithms for GOCE gravity gradients. Paper presented at the IAG Porto meeting (GGSM2004)

Bouman J, Koop R (2003) Error assessment of GOCE SGG data using along track interpolation. Advances in Geosciences, 1:27-32

Bouman J, Koop R, Tscherning CC, Visser P (2004b) Calibration of GOCE SGG data using high-low SST, terrestrial gravity data and global gravity field models. Journal of Geodesy, 78, DOI 10.1007/s00190-004-0382-5

Cesare S (2002) Performance requirements and budgets for the gradiometric mission. Issue 2 GO-TN-AI-0027, Preliminary Design Review, Alenia, Turin

ESA (1999) Gravity field and steady-state ocean circulation mission. Reports for mission selection. The four candidate Earth explorer core missions. ESA SP-1233(1). European Space Agency, Noordwijk

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

Kern M, Preimesberger T, Allesch M, Pail R, Bouman J, Koop R (2004) Outlier detection algorithms and their performance in GOCE gravity field processing. Accepted for publication in Journal of Geodesy

Lemoine F, Kenyon S, Factor J, Trimmer R, Pavlis N, Chinn D, Cox C, Klosko S, Luthcke S, Torrence M, Wang Y, Williamson R, Pavlis E, Rapp R, Olson T (1998) The development of the joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) geopotential model EGM96. TP 1998-206861, NASA Goddard Space Flight Center, Greenbelt

Overhauser A (1968) Analytic definition of curves and surfaces by parabolic blending. Techn. Report No. SL68-40, Scientific Research Staff Publication, Ford Motor Company, Detroit

Rapp R, Wang Y, Pavlis N (1991) The Ohio State 1991 geopotential and sea surface topography harmonic coefficient models. Rep 410, Department of Geodetic Science and Surveying, The Ohio State University, Columbus

Teunissen P (2000) Testing theory. Series on Mathematical Geodesy and Positioning 2. Delft University Press, Delft

Tscherning CC (1974) A FORTAN IV program for the determination of the anomalous potential using stepwise least squares collocation. Rep 225, Department of Geodetic Science and Surveying, The Ohio State University, Columbus

Tscherning CC (1991) The use of optimal estimation for gross-error detection in databases of spatially correlated data. Bulletin d'Information, no. 68, 79-89, BGI

Status of the European Gravity and Geoid Project EGGP

H. Denker [(1)], J.-P. Barriot, R. Barzaghi, R. Forsberg, J. Ihde, A. Kenyeres, U. Marti, I.N. Tziavos

[(1)] Institut für Erdmessung, Universität Hannover, Schneiderberg 50, D-30167 Hannover, Germany
E-mail: denker@ife.uni-hannover.de; Fax: +49-511-7624006

Abstract. The European Gravity and Geoid Project (EGGP) is a project within IAG Commission 2, reporting to Sub-commission 2.4. The main goal of the project is to compute an improved European geoid and quasigeoid model. This appears to be possible now because significant new and improved data sets have become available since the last computation in 1997 (EGG97). These improvements include better global geopotential models from the CHAMP and GRACE missions, better digital elevation models (DEMs) in some regions (e.g., new national DEMs, SRTM3, GTOPO30), updated gravity data sets for selected regions, updated ship and altimetric gravity data including improved merging procedures, the use of GPS/levelling data, as well as improved modelling and computation techniques.

An overview is given on the project structure, the computation strategy, the available data sets, the expected accuracies, the time schedule, and the work done so far. The primary input data sets are high-resolution gravity and terrain data supplemented by a state-of-the-art global geopotential model. The general computation strategy is the remove-restore procedure. The initial computations are based on the spectral combination approach with integral formulas evaluated by 1D FFT. First results based on an updated terrestrial gravity data set and GRACE geopotential models show significant improvements (up to 60 %) as compared to GPS/-levelling. Moreover, also the tilts, existing in previous solutions, have been reduced to typically below 0.1 ppm.

Keywords. Geoid, quasigeoid, gravity field modelling, GPS/levelling, EGGP, CHAMP, GRACE

1 Introduction

The latest high-resolution European geoid and quasigeoid models (EGG97) were computed at the Institut für Erdmessung (IfE), University of Hannover, acting since 1990 as the computing center of the International Association of Geodesy (IAG) Sub-commission for the European Geoid (predecessor of the EGGP), for details cf. Denker and Torge (1998). EGG97 is based on high-resolution gravity and terrain data in connection with the global geopotential model EGM96. The evaluation of EGG97 by GPS/levelling data revealed the existence of long wavelength errors at the level of 0.1 to 1 ppm, while the agreement over distances up to about 100 km is at the level of 0.01 m in many areas with a good quality and coverage of the input data (Denker and Torge, 1998; Denker, 1998).

Since the development of EGG97, significant new or improved data sets have become available, including strongly improved global geopotential models from CHAMP and GRACE, new national and global terrain data sets, new or updated gravity data sets, improved altimetric results, as well as new GPS/levelling campaigns. Furthermore, also the combination of ship and altimetric data has been refined, and new gravity field modelling techniques, e.g., wavelet techniques, fast collocation, etc., have become operational.

Considering all these advancements, a complete re-computation of the European geoid/quasigeoid is appropriate and promises significantly improved accuracies, especially at long wavelengths. Therefore, after the IUGG General Assembly in Sapporo in 2003, it was decided to support this task in the form of an IAG Commission 2 Project, named CP2.1 and entitled “European Gravity and Geoid Project EGGP”. The project is reporting to Sub-commission 2.4 and has strong connections to the International Gravity Field Service (IGFS), with its centers Bureau Gravimétrique International (BGI), International Geoid Service (IGeS), National Geospatial-Intelligence Agency (NGA), and GeoForschungsZentrum Potsdam (GFZ), as well as to several other IAG bodies, e.g., EUREF. The EGGP is running within the 4-year period from 2003 to 2007 until the next IUGG General Assembly. The project is organised by a steering committee (H. Denker (Chair), J.-P. Barriot, R. Barzaghi, R. Forsberg, J. Ihde, A. Kenyeres, U. Marti, I.N. Tziavos) and has about 50 national delegates (project members) from most of

the countries in Europe. The EGGP terms of reference can be found in EGGP (2003).

This contribution gives an overview on the general computation strategy and on the progress in the collection of gravity data, terrain data, and global geopotential models from the new space missions CHAMP and GRACE. First updated geoid/quasigeoid solutions are presented based on the new global geopotential models from CHAMP and GRACE. Moreover, results from an improved terrestrial gravity data set, including reprocessed ship gravity and new altimetric anomalies, are presented.

2 Computation Strategy

The basic computation strategy is based on the remove-restore technique, considering high-resolution terrestrial gravity and terrain data in combination with a state-of-the-art global geopotential model (probably based on the GRACE mission). Terrain reductions will be applied to smooth the data and to avoid aliasing effects. At present, the residual terrain model (RTM) technique according to Forsberg and Tscherning (1981) is favoured. Bathymetry and density data may be considered in special test areas. Moreover, GPS/levelling data will be used for control purposes, and may also be used for a combined solution (e.g., Denker et al., 2000), depending on the quality and availability of data. All data sets will be referred to uniform horizontal, vertical, and gravity reference systems. The collection of the relevant data sets is pursued by the steering committee and the members of the project.

A significant problem with high-resolution gravity and terrain data is the confidentiality of data, which must be assured to most of the data owners. For this reason, it was decided to have only one data and computation center at the Institut für Erdmessung (IfE), University of Hannover. In addition, a second confidential gravity data center is setup at Bureau Gravimétrique International (BGI) to use the expertise of BGI in the validation and cleaning of large gravity data sets. The inclusion of data in the confidential BGI project database requires separate agreements between the data owners and BGI, and there will be no connection to the BGI public database.

The primary gravity field quantity to be computed will be the height anomaly or the quasigeoid undulation, with the advantage that only gravity field observations at the Earth's surface and in its exterior enter into the calculation, avoiding assumptions about the Earth's interior gravity field. A geoid model is then derived by introducing a density hypothesis, which should be identical to the one used for the computation of corresponding orthometric heights.

Initially, the gravity modelling at IfE will be based on the spectral combination technique with integral formulas (e.g., Wenzel, 1982). In this method, the combination of terrestrial gravity data and a global geopotential model is done by means of spectral weights, which depend on the accuracy of the input data sets. Due to the high accuracy of the global models at long wavelengths, the terrestrial data mainly contribute the shorter wavelength components. Lateron, IfE may also test other modelling techniques, e.g., least squares collocation or wavelets. Moreover, it is planned to use the fast collocation approach developed by the Milan group (e.g., Sansò and Tscherning, 2003). Regarding the time frame, it is planned to have the final geoid/quasigeoid models in 2007 and preliminary solutions in 2005 and 2006.

The final goal is to strive for an accuracy of 0.01 m for the computed geoid/quasigeoid models (for distances up to some 100 km). Obviously, this is only possible in areas with a good coverage and quality of the input gravity and terrain data. The data requirements can be derived from theoretical and numerical studies including spectral analysis. With respect to the gravity data, a spacing of 2 to 5 km and an accuracy at the level of 1 mgal (white noise) is sufficient (Denker, 1988; Forsberg, 1993; Grote, 1996), but on the other hand even small systematic gravity errors affecting large regions may integrate up to significant geoid errors. For the elevation models, a resolution of roughly 100 m to 1000 m is adequate for alpine to low relief, respectively, with an accuracy at the level of some ten meters.

3 Recent Progress in Data Collection

3.1 Gravity Data

Since the start of the project, significant improvements of the gravity database have been made, including new data sets for several countries, e.g., Belgium, Luxemburg, Germany, Slovenia, Switzerland, and Netherlands. Moreover, positive responses, indicating a data update in the near future, were received from Austria, the Baltic States, Croatia, France, Greece, Poland, Serbia, Russia, the Scandinavian countries, etc. In addition to this, also the public domain data set from the Arctic Gravity Project became available (Forsberg and Kenyon,

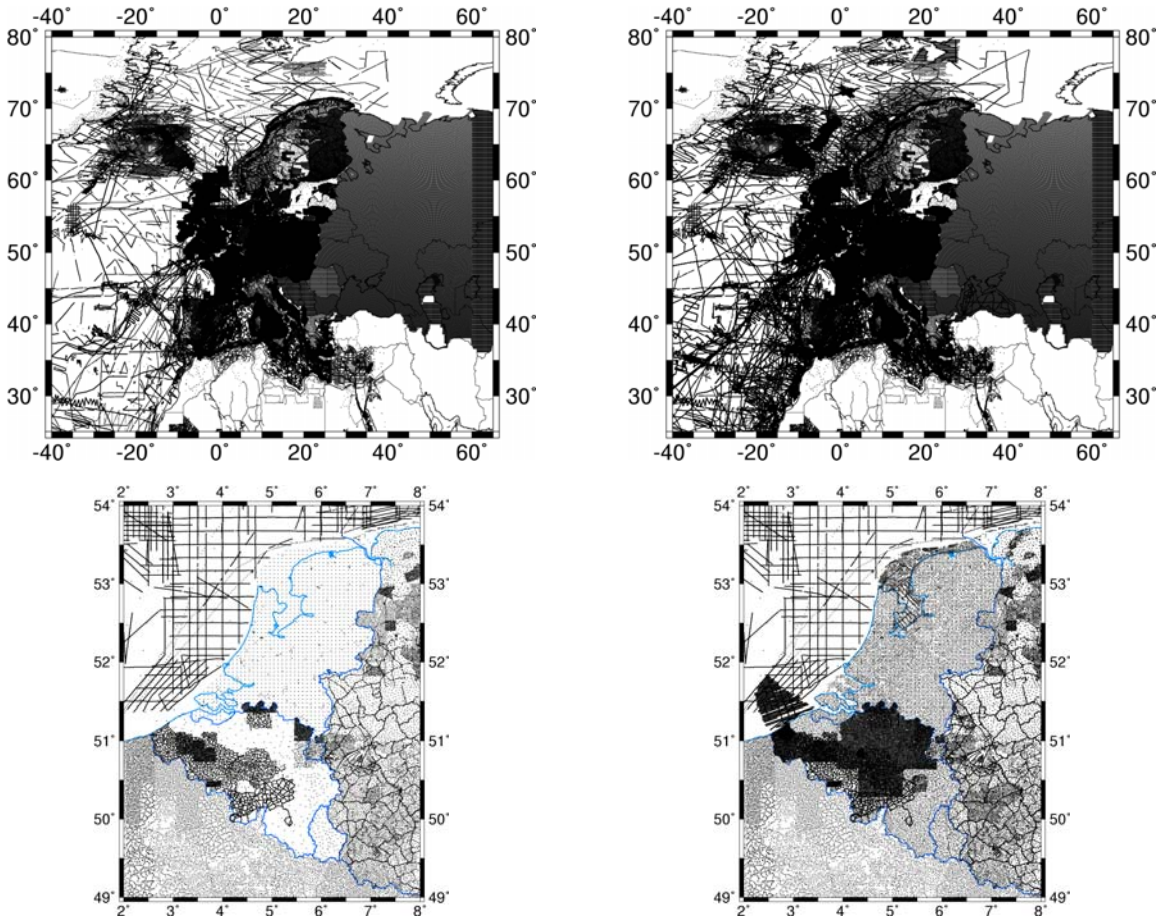

Fig. 1. Locations of terrestrial gravity data for entire Europe (top) and a sub-area (bottom). The left part shows the status for EGG97 and the right part shows the status of July 2004.

2004). As one example to document the progress in the collection of gravity data, Fig. 1 (bottom) depicts the old status (EGG97, Denker and Torge, 1998) and the new status in July 2004 for an area covering Belgium, Luxemburg, Netherlands, as well as parts of France and Germany.

In addition, the older gravity data sets were revised regarding the underlying reference systems, the target systems being ETRS (European Terrestrial Reference System), UELN (United European Levelling Network) and absolute gravity. Within the EGGP, only data which can be related without any doubts to the target reference systems will be included in the primary data base.

Significant progress was also made in the collection and reprocessing of ship gravity data (e.g., at IfE and other institutions). The ship gravity data, collected from several institutions for the European Seas, were crossover adjusted using a bias per track error model in order to reduce instrumental and navigational errors, incorrect ties to harbour stations, etc. (for details see Denker and Roland, 2003). An “original” and an “edited” data set were considered, where the edited data set excluded data affected by ship turns, data in the Red Sea, data from short tracks (< 3 points), and tracks with large crossover differences. Table 1 shows the crossover statistics for both data sets before and after the adjustment. The table clearly shows that the editing of some very bad observations resulted already in an improvement of the crossover differences by a factor of two, while the crossover adjustment again reduced the crossovers by a factor of two. Before the adjustment, the RMS crossover difference is 15.5 mgal for the original and 8.4 mgal for the edited data set; the corresponding values after the adjustment are 7.0 mgal and 4.7 mgal, respectively.

Table 1. Statistics of crossover differences from ship gravity observations. Units are mgal.

data set adjustment	original before	original after	edited before	edited after
#	89,328	89,328	78,929	78,929
Mean	0.20	-0.02	0.04	-0.01
RMS	15.48	7.01	8.37	4.69
Min	-258.43	-204.98	-109.91	-48.56
Max	+259.54	+198.37	+128.40	+49.16

The improvement of the ship data by editing and crossover adjustment was also illustrated by comparisons with altimetric anomalies from the KMS02 model (Andersen et al., 2003), giving a RMS difference of 18.0 mgal for the original data set and 10.2 mgal for the edited data set, both before the adjustment. The crossover adjustment further reduced the RMS difference to 7.8 mgal for the edited data set, proving the effectiveness of the entire procedure. In sub-areas, e.g., around Iceland, the RMS difference between the ship and KMS02 data is only 4.2 mgal.

Fig. 1 (top) depicts the locations of gravity data for entire Europe; the left part shows the status for EGG97, and the right part shows the status for July 2004 including the reprocessed ship gravity data.

3.2 Digital Elevation Models (DEMs)

For the EGG97 model, digital elevation models (DEMs) were available with a resolution of about 200 m for most countries in Central and Western Europe, while coarser grids with a resolution of 0.5 km to 10 km had to be used for the remaining parts of Europe. For EGG97, only in Germany a DEM with a very high resolution of 1" × 1" (approx. 30 m) was available. Meanwhile, also Switzerland has released a 1" × 1" DEM, and Austria has indicated the release of a corresponding model. However, especially in Eastern Europe and some other areas, fill-ins from global public domain databases have to be used, either because high-resolution DEMs do not exist or are not released for confidentiality reasons. For this purpose, the SRTM3 model with a resolution of 3" × 3" (JPL, 2004) and the public domain global model GTOPO30 with a resolution of 30" × 30" (LP DAAC, 2004) can be used. The SRTM3 model has been released recently from the analysis of the Shuttle Radar Topography Mission as a preliminary and "research-grade" model, covering the latitudes between 60°N and 54°S. On the other hand, the GTOPO30 model has global coverage and was derived already in 1996 from several raster and vector sources of topographic information (LP DAAC, 2004).

The SRTM3 and GTOPO30 DEMs were evaluated at IfE by comparisons with national DEMs for Germany, based on 1" × 1" data (Denker, 2004a). The comparisons of the national and SRTM3 models revealed that one of the national models contained less accurate fill-ins in some areas outside of Germany. After excluding these areas, the differences between the best national model and the SRTM3 DEM show a standard deviation of 7.9 m with maximum differences up to about 300 m. The largest differences are located in opencast mining areas and result from the different epochs of the data. Histograms of the differences show a clear deviation from the normal distribution with a long tail towards too high SRTM3 elevations. Moreover, the presently available SRTM3 "research-grade" models contain numerous data voids (undefined elevations), which cause significant problems. The filling of these data gaps by interpolation must be handled with care, especially for larger gaps in mountainous areas (Denker, 2004a).

The evaluation of the GTOPO30 model by national and SRTM3 DEMs demonstrated that in Germany the longitudes of GTOPO30 should be increased by 30" (one block). The longitude shift reduced the standard deviation of the differences to the national and SRTM3 models by roughly 75 %, yielding final values of about 6.8 m and 11.5 m for the national and SRTM3 models, respectively.

Thus, the national DEMs, augmented by the SRTM3 and GTOPO30 data will allow the creation of DEMs for entire Europe with a resolution of at least 30" × 30", which is a significant improvement compared to the previous EGG97 computation.

3.3 Global Geopotential Models

The CHAMP and GRACE missions have led to significant improvements in the modelling of long wavelength gravity signals. This is documented, e.g., by the accumulated formal geoid error, which does not exceed 0.01 m for spherical harmonic degrees up to about 25 for the CHAMP models (e.g., Reigber et al., 2004a) and 75 for the GRACE models (e.g., Reigber et al., 2004b). On the other hand, the limit of 0.01 m is already exceeded at degree 8 for the EGM96 model. Correspondingly, the limit of 0.05 m is exceeded at about degree 20 for EGM96, 40 for the CHAMP models, and 90 for the GRACE models.

The new geopotential models from the CHAMP and GRACE missions, in combination with terrestrial gravity data of good quality (±1 mgal) and coverage, allow the computation of significantly improved continental-scale geoid and quasigeoid models. Error estimates based on the degree variance approach result in standard deviations of about 0.02 m to 0.03 m for solutions based on the GRACE models, with the largest contribution (about 0.02 m) coming from the degree range 90 to 360. The corresponding values for geoid solutions based on the CHAMP models and EGM96 are about 0.04 m and 0.06 m, respectively.

4 First Results

Updated European geoid/quasigeoid models were computed based on the new CHAMP and GRACE geopotential models. The computations were done using the EGG97 terrestrial gravity data set as well as an updated data set (section 3.1). The computations were done using the remove-restore technique in connection with the least squares spectral combination method. The spectral weights were derived from the error estimates of the global models and the terrestrial data. Terrain reductions were done using the RTM method. The computation area is 25°N – 77°N and 35°W – 67.4°E. The grid spacing is 1' × 1.5', yielding 3,120 × 4,096 grid points. The GRS80 constants, the zero degree undulation terms, and the zero-tide system were used throughout all computations (for details see Denker, 2004b).

All computed quasigeoid models were evaluated by GPS/levelling data from the European EUVN data set (Ihde et al., 2000) and by national campaigns. Fig. 2 shows the differences (after subtracting a common bias) between 166 stations of the EUVN GPS/levelling data set (only stations with UELN normal heights were used) and the EGG97 gravimetric quasigeoid based on EGM96 (left part), as well as a new solution (right part) based on the EIGEN-GRACE02S geopotential model (Reigber et al., 2004b); the terrestrial gravity data are identical in both solutions (status of EGG97). Fig. 2 shows clearly that the long wavelength discrepancy over Central Europe almost disappears for the GRACE solution; the largest discrepancies remain at coastal stations, especially around the Mediterranean Sea where the gravity data quality is weak. The RMS difference is 0.262 m for EGG97 and reduces to 0.230 m for the EIGEN-GRACE02S solution (12 % improvement). Correspondingly, a solution based on the CHAMP EIGEN-3 model (Reigber et al., 2004a) gives a RMS difference of 0.238 m (9 % improvement). When using the updated terrestrial gravity data set from 2004 in combination with the EIGEN-GRACE02S model, the RMS difference reduces to 0.203 m (23 % improvement compared to EGG97). Furthermore, when transforming the GPS results (according to Poutanen et al., 1996) from the non-tidal to the zero-tide system, which is used for the quasigeoid solutions, another slight reduction of the RMS difference to 0.197 m can be observed (25 % total improvement versus EGG97). Tilt parameters were also computed, but not considered any further as they were less than 0.1 ppm in all cases.

Additional comparisons of EGG97 and the new quasigeoid solution based on EIGEN-GRACE02S

Table 2. RMS differences from comparisons of GPS/levelling with EGG97 and a new quasigeoid based on EIGEN-GRACE02S. A constant bias is subtracted. Units are m.

Country	# pts.	EGG97/ EGM96	EGG04/ EIGEN-GRACE02S	Improvement
Belgium	31	0.061	0.046	25 %
France	965	0.128	0.084	34 %
Germany	678	0.107	0.041	62 %
Hungary	299	0.089	0.057	36 %
Netherlands	84	0.035	0.031	11 %
Switzerland	147	0.084	0.063	24 %
EUVN	166	0.262	0.230	12 %

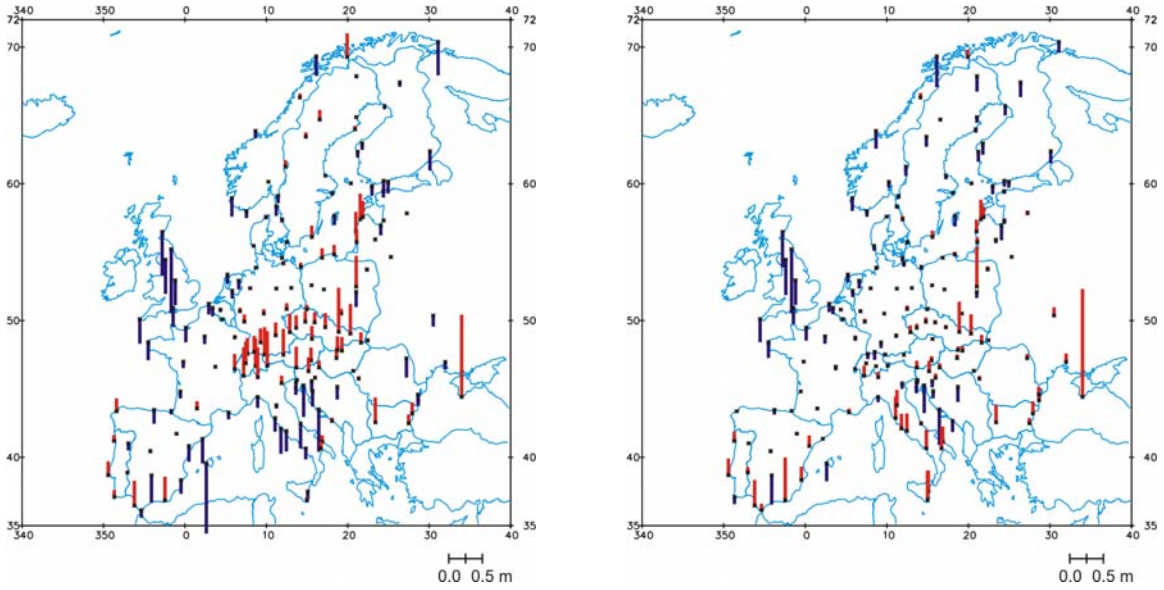

Fig. 2. Comparison of EGG97 quasigeoid solution (left) and a new solution based on the EIGEN-GRACE02S global model (right) with GPS/levelling data from the EUVN campaign. A constant bias is subtracted.

with several national GPS/levelling data sets are shown in Table 2. Again, both solutions use identical terrestrial gravity data (status EGG97). The table provides the RMS differences for both solutions after subtracting a common bias. The table clearly shows that in all cases the use of the new GRACE model improves the geoid/quasigeoid modelling significantly. The maximum improvement is more than 60 % for the German data set. A more detailed analysis shows that the tilts, existing in EGG97, are reduced to typically below 0.1 ppm, i.e. by one order of magnitude in some cases.

Furthermore, with the updated solutions based on the GRACE models, accurate determinations of W_0 (reference geopotential of the vertical datum) and vertical datum unifications become possible. When considering the 2004 terrestrial gravity data set, the EIGEN-GRACE02S model, the EUVN GPS/levelling data, and a transformation of the GPS heights to the zero-tide system, an estimate of W_0 (Europe) of 62,636,857.02 ±0.15 m^2s^{-2} is obtained. The corresponding value from the German GPS/-levelling data is 62,636,856.91 ±0.02 m^2s^{-2}. Both values are in good agreement with the value 62,636,857.25 m^2s^{-2} published for Europe by Burša et al. (2002). However, the European values deviate by about 1.0 m^2s^{-2} from the global best estimates (e.g., Burša et al., 2002).

5 Conclusions

Significant progress was made within the framework of the European Gravity and Geoid Project EGGP regarding the collection and homogenization of high-resolution gravity and terrain data. Several new data sets became available, and especially the new geopotential models from the CHAMP and GRACE missions improved the geoid/quasigeoid modelling very much. In the GPS/levelling comparisons, the RMS differences reduced up to about 60 % when using the GRACE models and up to 30 % for the solutions based on CHAMP, as compared to the previous EGG97 model relying on EGM96. In addition, the tilts, existing in EGG97, were also reduced to typically below 0.1 ppm. Due to the support with data by numerous people and agencies, further improvements are to be expected in the future.

References

Andersen, O.B., P. Knudsen, S. Kenyon, and R. Trimmer (2003). KMS2002 Global Marine Gravity Field, Bathymetry and Mean Sea Surface. Poster, IUGG2003, Sapporo, Japan, June30-July11, 2003.

Burša, M., et al. (2002). World height system specified by geopotential at tide gauge stations. IAG Symposia 124: 291-296, Springer Verlag.

Denker, H. (1988). Hochauflösende regionale Schwerefeldmodellierung mit gravimetrischen und topographischen Daten. Wiss. Arb. Fachr. Verm.wesen, Univ. Hannover, Nr. 156.

Denker, H. (1998). Evaluation and Improvement of the EGG97 Quasigeoid Model for Europe by GPS and Leveling Data. Reports of the Finnish Geodetic Institute, 98:4, 53-61, Masala, 1998.

Denker, H. (2004a). Evaluation of SRTM3 and GTOPO30 terrain data in Germany. Proceed. GGSM2004, Porto.

Denker, H. (2004b). Improved European geoid models based on CHAMP and GRACE results. Proceed. GGSM2004, Porto.

Denker, H., M. Roland (2003). Compilation and evaluation of a consistent marine gravity data set surrounding Europe. Proceed., IUGG2003, Sapporo, Japan, June30-July11, 2003.

Denker, H., W. Torge (1998). The European gravimetric quasigeoid EGG97 – An IAG supported continental enterprise. IAG Symposia 119:249-254, Springer Verlag.

Denker, H., W. Torge, G. Wenzel, J. Ihde, U. Schirmer (2000). Investigation of Different Methods for the Combination of Gravity and GPS/Levelling Data. IAG Symposia 121:137-142, Springer-Verlag.

EGGP (2003). Commission Project 2.1, European Gravity and Geoid, Terms of Reference and Objectives, http://www.ceegs.ohio-state.edu/iag-commission2/cp2.1.htm.

Forsberg, R. (1993). Modelling the fine-structure of the geoid: methods, data requirements and some results. Surveys in Geophys. 14: 403-418.

Forsberg, R., S. Kenyon (2004). Gravity and geoid in the Arctic Region – The northern GOCE polar gap filled. Proceed. 2nd Internat. GOCE Workshop, Esrin, March 8-10, 2004.

Forsberg, R., C.C. Tscherning (1981). The use of height data in gravity field approximation by collocation. Journal of Geophys. Research 86: 7843-7854.

Grote, T. (1996). Regionale Quasigeoidmodellierung aus heterogenen Daten. Wiss. Arb. Fachr. Verm.wesen, Univ. Hannover, Nr. 212.

Ihde, J. et al. (2000). The height solution of the European Vertical Reference Network (EUVN). Veröff. Bayer. Komm. für die Internat. Erdmessung, Astronom. Geod. Arb., Nr. 61: 132-145, München.

JPL (2004). SRTM – The Mission to Map the World. Jet Propulsion Laboratory, California Inst. of Techn., http://www2.jpl.nasa.gov/srtm/index.html.

LP DAAC (2004). Global 30 Arc-Second Elevation Data Set GTOPO30. Land Process Distributed Active Archive Center, http://edcdaac.usgs.gov/gtopo30/gtopo30.asp.

Poutanen, M., M. Vermeer, J. Mäkinen (1996). The permanent tide in GPS positioning. J. of Geodesy 70: 499-504.

Reigber, Ch., et al. (2004a). Earth gravity field and seasonal variability from CHAMP. Earth Observation with CHAMP: 25-30, Springer Verlag.

Reigber, Ch., et al. (2004b). An Earth gravity field model complete to degree and order 150 from GRACE: EIGEN-GRACE02S. J. of Geodynamics, accepted June 25, 2004.

Sansò, F., C.C. Tscherning (2003). Fast spherical collocation: theory and examples. Journal of Geodesy 77:101-112.

Wenzel, H.-G. (1982). Geoid computation by least squares spectral combination using integral kernels. Proceed. IAG General Meet., 438-453, Tokyo.