Journal of Geodesy (2003) 77: 101–112
DOI 10.1007/s00190-002-0310-5

Journal of Geodesy

Fast spherical collocation: theory and examples

F. Sansò[1], C. C. Tscherning[2]

[1] DIIAR–Sezione Rilevamento, Politecnico di Milano, P.zza Leonardo da Vinci 31, 20133 Milano, Italy
e-mail: fsanso@geo.polimi.it; Tel.: +39-2-2399-6504; Fax: +39-2-2399-6530
[2] Department of Geophysics, University of Copenhagen, Juliane Maries Vej 30, 2100 Copenhagen Oe, Denmark
e-mail: cct@gfy.ku.dk; Tel.: +45-35 320582; Fax: +45-35365357

Received: 18 October 2001 / Accepted: 4 October 2002

Abstract. It has long been known that a spherical harmonic analysis of gridded (and noisy) data on a sphere (with uniform error for a fixed latitude) gives rise to simple systems of equations. This idea has been generalized for the method of least-squares collocation, when using an isotropic covariance function or reproducing kernel. The data only need to be at the same altitude and of the same kind for each latitude. This permits, for example, the combination of gravity data at the surface of the Earth and data at satellite altitude, when the orbit is circular. Suppose that data are associated with the points of a grid with N values in latitude and M values in longitude. The latitudes do not need to be spaced uniformly. Also suppose that it is required to determine the spherical harmonic coefficients to a maximal degree and order K. Then the method will require that we solve K systems of equations each having a symmetric positive definite matrix of only $N \times N$. Results of simulation studies using the method are described.

Keywords: Least-squares collocation – Gravity field – Longitude grid

1 Introduction

The gravity potential of the Earth, W, is normally split into two parts: a reference potential, U, and an anomalous potential, T, so that $W = U + T$. The reference potential may for example be a spherical harmonic expansion to degree $N = 360$ plus the centrifugal potential. Then T is a harmonic function in the space outside the Earth, when we disregard all external masses (atmosphere, planets).

T may be approximated ('modelled') using a linear combination of various types of (generally harmonic) functions: solid spherical harmonics, representers of observation functionals and point-mass potentials (see Moritz 1980)

$$\tilde{T}(P) = \sum_{i=1}^{n} a_i f_i(P)$$

In particular in the last decade, a lot of work has been successfully performed in applying the wavelet concept to the analysis of harmonc functions in spherical domains (cf. Freeden et al. 1998).

The coefficients a_i are determined by solving systems of linear equations. If we have more observations than base functions f_i the coefficients are often determined using the method of least squares (LS). The solution may, however, also be expressed in another manner using solid spherical harmonics if the base functions are harmonic. From the estimate $\tilde{T}$ we may compute coefficients T_{nm} in a spherical harmonic series. If the base, functions are themselves the spherical harmonics, then obviously a_i and T_{nm} are the same quantities

$$\tilde{T}(\vartheta, \lambda, r) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} T_{\ell m} \left(\frac{a}{r}\right)^{\ell+1} Y_{\ell m}(\vartheta, \lambda)$$

where ϑ is the co-latitude, λ the longitude, r the distance to the origin

$$Y_{\ell m}(\vartheta, \lambda) = \bar{P}_{\ell m}(\cos\vartheta) e_m(\lambda)$$

$\bar{P}_{\ell m}$ are the fully normalized Legendre functions of degree ℓ and order m and $e_m(\lambda) = \cos m(\lambda)$ when $m \geq 0$ and $e_m(\lambda) = \sin|m|(\lambda)$ when $m < 0$.

Then, reversing the summation order, we can write

$$\hat{T}_{\ell m}(\vartheta, \lambda, r) = \sum_{m=-\infty}^{\infty} \left(\sum_{\ell=|m|}^{\infty} T_{\ell m} \left(\frac{a}{r}\right)^{\ell+1} \bar{P}_{\ell m}(\cos\vartheta) \right) e_m(\lambda)$$

It is well known that we may take advantage of this re-ordering when evaluating spherical harmonic series at points with constant latitude and distance to the origin, r.

Correspondence to: F. Sansò

We will also take advantage of this in the following derivations.

If the observations form a grid in longitude when projected onto the unit sphere, further savings can be made when determining the coefficients $T_{\ell m}$. Systems of equations which have to be solved may have matrices with a nice repetitive (Toeplitz) structure. We may take advantage of this structure and reduce the numerical effort considerably (see e.g. Colombo 1979).

In the method of least-squares collocation (LSC) the setting is an infinite-dimensional Hilbert space. A solution is found by requiring a minimum norm or variance of the estimate $\tilde{T}$. This results in systems of equations as large as the number of observations (see Moritz 1980).

The increase in computer power and storage possibilities has recently permitted the determination of an estimate of $\tilde{T}$ (using simulated data; see Tscherning 2001) corresponding to a 1-degree equal-area distribution, i.e. the data did not form a regular grid in latitude and longitude.

A number of different techniques proposed since the 1980s are in fact equivalent to the collocation approach when the proper reproducing kernel Hilbert space is chosen. In particular, this is true for harmonic spherical wavelets, although due to their multi-resolution properties they are more interesting in dealing with local problems rather than with global ones as here (cf. Freeden et al. 1998).

In the following it will be shown how, in the case of a regular grid, the computational effort may be reduced dramatically for the case of LSC. Naturally there are some requirements to the data which must be fulfilled. However, we have conserved one of the important properties of LSC, namely that we can mix different data types, provided the data on each parallel are of the same kind.

In this way we are able to combine gravity gradient data at satellite altitude (300 km) with gravity anomaly data at the surface of the Earth.

Although it might seem that this condition is awkward and unlikely to be met in practice, it happens to match exactly an important situation that is going to become true for a number of geodetic satellite missions. Imagine we have a satellite on a circular orbit of radius $\bar{r}$ with some inclination I; for instance, a satellite observing second radial derivatives, as will happen to a good approximation with GOCE. Then on the geocentric sphere with radius $\bar{r}$ we can have a nice field of measurements covering the sphere, with a distribution displaying a cylindrical symmetry, but for two polar caps of radius $|I - (\pi/2)|$. It is known that this weakens considerably the determination of a number of coefficients (cf. Moritz 1980; Sneeuw et al. 2001), but the effect can be reduced if we introduce other observations on the gravity field, which are obtained from terrestrial (or aerial) measurements. The method we present here is able to manage situations like this with a relatively simple numerical apparatus.

To conclude the introduction, it is mentioned that it is always debatable, whether spherical hamonic coefficients are or are not the 'best' parameters to represent the potential T. However, as these coefficients are in themselves so important also for physical reasons, they are the functionals of T considered these days as what we have to achieve in the content of global gravity field models. In the present paper this point of view is accepted, so that we discuss mostly the estimation of $T_{\ell m}$, although in Appendix A a generalization to other functionals has been outlined.

In Sects. 2 and 3 the theory behind the new method will be developed. It is based on some simple properties of discrete Fourier series, which are demonstrated in Appendix B. We describe in Sect. 4 the considerations behind the numerical implementation of the method, and in Sect. 5 we describe the testing of the method and give examples of applications.

2 The problem

Suppose we have observations of the potential T which can be modelled as

$$Q_{oik} = B_{ik}(T) + v_{ik} \tag{1}$$

where the functionals B_{ik} have the form

$$B_{ik}(T) = \sum_{\ell=0}^{+\infty} \sum_{m=-\ell}^{\ell} T_{\ell m} b_i^{\ell m} e_m(\lambda_k) \tag{2}$$

These B_{ik} are then related to a grid of points P_{ik} on the sphere

$$\vartheta(P_{ik}) = \vartheta_i \quad i = 1, 2, \ldots, N$$
$$\lambda(P_{ik}) = \lambda_k = k\delta \quad k = 0, \ldots, M-1$$
$$\delta = \frac{2\pi}{M} \quad M = 2L$$

It is crucial in the following reasoning to assume that ϑ_i are any values but λ_k are regularly spaced. The v_{ik} are observational noise with the following characteristics:

$$E\{v_{ij} v_{jk}\} = \delta_{ij}\delta_{kn}\xi_i \quad (\xi_i > 0) \tag{3}$$

i.e. they are linearly independent and with variances depending only on ϑ_i! Here $E\{\}$ is the expectation operation with respect to noise.

The hypothesis that the observational functionals are of the form of Eq. (2) is not very restrictive, as we can see in the following examples. In particular, this is much less restrictive than assuming B_{ik} to be diagonal in spherical harmonics.

2.1 Examples

In the following examples we shall use the notation $a(\vartheta)$ to indicate, whenever appropriate, any smooth function of ϑ that can be suitably chosen to specify the operator.

1. B_{ik} upward continuation to point $Q_{ik} = (r_i, \vartheta_i, \lambda_k)$ (note, r has to depend only on i)

$$b_i^{\ell m} = \left(\frac{R}{r_i}\right)^{\ell+1} \bar{P}_{\ell m}(\cos\vartheta_i)$$

2. B_{ik}, radial derivative $\partial^\alpha/\partial r^\alpha$ at $Q_{ik} = (r_i, \vartheta_i, \lambda_k)$

$$b_i^{\ell m} = \frac{(-1)^\alpha(\ell+1)\cdots(\ell+2-\alpha)}{R^\alpha}\cdot\left(\frac{R}{r_i}\right)^{\ell+\alpha+1}\bar{P}_{\ell m}(\cos\vartheta_i)$$

3. B_{ik}, meridional derivative

$$B_{ik} \equiv \frac{a(\vartheta)}{r}\frac{\partial}{\partial\vartheta}\bigg|_{Q_{ik}} \text{ at } Q_{ik} \equiv (r_i, \vartheta_i, \lambda_k)$$

$$b_i^{\ell m} = \frac{1}{R}\left(\frac{R}{r_i}\right)^{\ell+2} a(\vartheta_i)\bar{P}_{\ell m}(\cos\vartheta_i)$$

or higher-order generalizations.

4. B_{ik}, longitudinal second derivative

$$B_{ik} = \frac{a(\vartheta)}{r^2}\frac{\partial^2}{\partial\lambda^2}\bigg|_{Q_{ik}} \text{ at } Q_{ik} \equiv (r_i, \vartheta_i, \lambda_k)$$

$$b_i^{\ell m} \equiv \frac{1}{R^2}\left(\frac{R}{r_i}\right)^{\ell+3}\cdot a(\vartheta_i)\bar{P}_{\ell m}(\cos\vartheta_i)(-m^2)$$

5. B_{ik}, spherical block average

$$B_{ik} = \int_{\vartheta_i-\Delta}^{\vartheta_i+\Delta} \mathrm{d}\vartheta a(\vartheta) \int_{\lambda_k-\Delta}^{\lambda_k+\Delta} \mathrm{d}\lambda$$

around the point $Q_{ik}(r_i, \vartheta_i, \lambda_k)$; considering that

$$\int_{\lambda_k-\Delta}^{\lambda_k+\Delta} e_m(\lambda)\ \mathrm{d}\lambda = 2(\sin\Delta)e_m(\lambda_k) \quad \forall\, m$$

we obtain

$$b_i^{\ell m} = \left(\frac{R}{r_i}\right)^{\ell+1} 2(\sin\Delta)\int_{\vartheta_i-\Delta}^{\vartheta_i+\Delta} \bar{P}_{\ell m}(\cos\vartheta)\ \mathrm{d}\vartheta$$

Remark. One important example does not fit into this scheme, but it can be treated in a similar way, namely the longitudinal first-order differential

$$B_{ik} = \frac{a(\vartheta)}{r}\ \frac{\partial}{\partial\lambda}\bigg|_{Q_{ik}}$$

However, we have in this case

$$\frac{\partial}{\partial\lambda}e_m(\lambda) = -me_{-m}(\lambda)$$

so that, instead of Eq. (2), the following representation holds:

$$B_{ik}(T) = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} T_{\ell-m}b_i^{\ell m}e_m(\lambda_k)$$

with

$$b_i^{\ell m} = -\frac{m}{R}\left(\frac{R}{r_i}\right)^{\ell+2} a(\vartheta_i)\bar{P}_{\ell m}(\cos\vartheta_i)$$

remember in fact that $\bar{P}_{\ell,-m} \equiv \bar{P}_{\ell m}$ by definition.

Our problem is now summarized as follows: from observations Q_{oik} we want to obtain an estimate of $T_{\ell m}$

$$T_{\ell m} = L_{\ell m}(\tilde{T})$$

by applying the LSC concept.

3 The collocation solution

Let us first remember that the collocation algorithm coincides with an ordinary stochastic (Wiener–Kolmogorov) estimation under the assumption that

$$E\{T_{\ell m}\} = 0,\ E\{T_{\ell m}T_{pq}\} \equiv \sigma_\ell^2\delta_{\ell p}\delta_{mq} \tag{4}$$

This provides us with a tool to simplify the following computations. In order to perform a collocation estimate of $T_{\ell m} = L_{\ell m}(T)$ we need the following variances and covariances:

$$E\{L_{\ell m}(T)^2\} \equiv \sigma_\ell^2 \tag{5}$$

$$\begin{aligned} E\{L_{\ell m}(T)Q_{oik}\} &\equiv E\{L_{\ell m}(T)B_{ik}(T)\} \\ &\equiv \sigma_\ell^2 b_i^{\ell m}e_m(\lambda_k) \end{aligned} \tag{6}$$

$$\begin{aligned} &E\{Q_{oik}, Q_{ojn}\} \\ &\quad\equiv E\{B_{ik}(T)B_{jn}(T)\} + \xi_i\delta_{ij}\delta_{kn} \\ &\quad= \sum_{\ell'=0}^{+\infty}\sum_{m'=-\ell'}^{\ell'}\sigma_{\ell'}^2 b_i^{\ell' m'}b_j^{\ell' m'}e_{m'}(\lambda_k)e_{m'}(\lambda_n) + \xi_i\delta_{ij}\delta_{kn} \\ &\quad\equiv \sum_{m'=-\infty}^{+\infty} e_{m'}(\lambda_k)e_{m'}(\lambda_n)\Gamma_{ij}^{m'} + \xi_i\delta_{ij}\delta_{kn} \end{aligned} \tag{7}$$

$$\Gamma_{ij}^{m'} \equiv \sum_{\ell'=|m'|}^{+\infty} b_i^{\ell' m'}b_j^{\ell' m'}\sigma_{\ell'}^2$$

Note that $\{\Gamma^{m'}\}$ is a symmetrical and positive definite matrix.

Now let us recall the following rules (cf. Appendix B):

$$\begin{aligned} &0 \le m' \le L, h \ge 0 \\ &e_{m'}(\lambda_k) \equiv e_{m'+hM}(\lambda_k) \\ &e_{m'}(\lambda_k)e_{m'}(\lambda_n) \equiv e_{m'+hM}(\lambda_k)e_{m'+hM}(\lambda_n) \end{aligned}$$

$$\begin{aligned} &0 < m' < L, h \ge 1 \\ &e_{m'}(\lambda_k) \equiv e_{hM-m'}(\lambda_k) \\ &e_{m'}(\lambda_k)e_{m'}(\lambda_n) \equiv e_{hM-m'}(\lambda_k)e_{hM-m'}(\lambda_n) \end{aligned}$$

$$\begin{aligned} &-(L-1) \le m' < 0, h \ge 0 \\ &e_{m'}(\lambda_k) \equiv e_{-hM+m'}(\lambda_k) \\ &e_{m'}(\lambda_k)e_{m'}(\lambda_n) \equiv e_{-hM+m'}(\lambda_k)e_{-hM+m'}(\lambda_n) \end{aligned}$$

$$\begin{aligned} &-(L-1) \le m' < 0, h \ge 1 \\ &e_{m'}(\lambda_k) \equiv -e_{-hM-m'}(\lambda_k) \\ &e_{m'}(\lambda_k)e_{m'}(\lambda_n) \equiv e_{-hM-m'}(\lambda_k)e_{-hM-m'}(\lambda_n) \end{aligned}$$

Accordingly we can write

$$0 \le \forall\, k, n \le M-1$$

$$\sum_{m'=-\infty}^{+\infty} e_{m'}(\lambda_k) e_{m'}(\lambda_n) \Gamma_{ij}^{m'}$$

$$\equiv \sum_{m'=-(L-1)}^{L} e_{m'}(\lambda_k) e_{m'}(\lambda_n) \tilde{\Gamma}_{ij}^{m'} \tag{8}$$

where

$$\tilde{\Gamma}_{ij}^{0} = \sum_{h=0}^{+\infty} \Gamma_{ij}^{hM}$$

$$\tilde{\Gamma}_{ij}^{L} = \sum_{h=0}^{+\infty} \Gamma_{ij}^{L+hM}$$

$$(0 < m' < L),$$

$$\tilde{\Gamma}_{ij}^{m'} = \sum_{h=0}^{+\infty} \left(\Gamma_{ij}^{m'+hM} + \Gamma_{ij}^{(h+1)M-m'} \right)$$

$$[-(L-1) \le m' < 0],$$

$$\tilde{\Gamma}_{ij}^{m'} = \sum_{h=0}^{+\infty} \left(\Gamma_{ij}^{m'-hM} + \Gamma_{ij}^{-m'-(h+1)M} \right)$$

Summarizing, we see that Eq. (7) can be written as

$$E\{Q_{oik}, Q_{ojn}\} \equiv C(Q_{oik}, Q_{ojn})$$

$$= \sum_{m'=-(L-1)}^{L} e_{m'}(\lambda_k) e_{m'}(\lambda_n) \tilde{\Gamma}_{ij}^{m'} + \xi_i \delta_{ij} \delta_{kn} \tag{9}$$

where the formula also implicitly defines $C(\cdot, \cdot)$.

For the sake of readibility, from now on we will choose Krarup's notation (see Krarup 1969) where the covariance of any two functionals of $T, A(T), B(T)$ is derived from one and the same covariance, namely

$$C_{TT}(P, Q) = E\{T(P) T(Q)\}$$

and written as

$$C(A, B) \equiv E\{A(T), B(T)\} \equiv A_P\{B_Q C_{TT}(P, Q)\}$$

Now we can consider the solution of our problem. We put

$$\hat{T}_{\ell m} = \sum_{(ik)} \Lambda_{ik}^{\ell m} Q_{oik} \tag{10}$$

and we want to determine $\Lambda_{ik}^{\ell m}$ and the estimation error of $\hat{T}_{\ell m}$.

Note: during the whole reasoning (ℓ, m) in Eq. (10) is fixed, i.e. we estimate one coefficient at a time.

The matrix with coefficients $\Lambda_{ik}^{\ell m}$ is determined by solving the system

$$\sum_{(jn)} C(Q_{oik}, Q_{ojn}) \Lambda_{jn}^{\ell m} = C(L_{\ell m}, Q_{oik}) \tag{11}$$

or

$$\sum_{m'=-(L-1)}^{L} e_{m'}(\lambda_k) \sum_{(j)} \tilde{\Gamma}_{ij}^{m'} \sum_{(n)} e_{m'}(\lambda_n) \Lambda_{jn}^{\ell m} + \xi_i \Lambda_{ik}^{\ell m} \equiv \sigma_\ell^2 b_i^{\ell m} e_m(\lambda_k) \tag{12}$$

Remark. In the right-hand side, m is any integer between $-\infty$ and $+\infty$; however, in the interval $-(L-1) \le t \le L$ there is only one $t = t(m)$ such that

$$m = t + hM \quad (\text{i.e. } t = [m \text{ Mod } M] - [L-1]) \tag{13}$$

and also then

$$e_m(\lambda_k) \equiv e_t(\lambda_k)\ ,\ \forall\, k \tag{14}$$

Accordingly, we can substitute $e_t(\lambda_k)$ in the right-hand side of Eq. (12), with t given by Eq. (13).

Let us multiply Eq. (12) by $e_r(\lambda_k)$, with $-(L-1) \le r \le L$, and sum over k. Because of the orthogonality relation of Eq. (B12) (cf. Appendix B), we have

$$\frac{M}{2} \epsilon_r \sum_{(j)} \tilde{\Gamma}_{ij}^{r} \left(\sum_{(n)} e_r(\lambda_n) \Lambda_{jn}^{\ell m} \right) + \xi_i \sum_{(k)} \Lambda_{ik}^{\ell m} e_r(\lambda_k) \equiv \sigma_\ell^2 b_i^{\ell m} \frac{M \epsilon_r}{2} \delta_{rt} \tag{15}$$

where

$$\epsilon_r = (1 + \delta_{r0})(1 + \delta_{rL})$$

It should be noted that by multiplying Eq. (15) by $1/\varepsilon_r e_r(\lambda_n)$ and summing over r, by exploiting Eq. (B13) in Appendix B, we go back to Eq. (12), thus proving that the two equations are equivalent.

This means that the matrix

$$\Delta_{ir}^{\ell m} = \sum_{(k)} \Lambda_{ik}^{\ell m} e_r(\lambda_k) \tag{16}$$

satisfies the equation

$$\sum_{(j)} \left(\tilde{\Gamma}^{r} + \frac{2}{M \epsilon_r} \Xi \right)_{ij} \Delta_{ir}^{\ell m} = \sigma_\ell^2 \delta_{rt} b_i^{\ell m} \tag{17}$$

where

$$\Xi \equiv \text{Diag}\ \{\xi\} \tag{18}$$

Then, considering that Γ is positive definite unless the functionals are not independent, and in any case Ξ is always positive

$$\Delta_{ir}^{\ell m} \equiv \sum_{(k)} \Lambda_{ik}^{\ell m} e_r(\lambda_k)$$

$$\equiv \sum_{(j)} \left(\tilde{\Gamma}^{r} + \frac{2}{M \epsilon_r} \Xi \right)_{ij}^{-1} b_j^{\ell m} \sigma_\ell^2 \delta_{rt} \tag{19}$$

Using Eq. (B13) (cf. Appendix B) in Eq. (19) yields

$$\Lambda_{in}^{\ell m} \equiv \sum_{(j)} \left(\tilde{\Gamma}^r + \frac{2}{M\epsilon_t} \Xi \right)_{ij}^{-1} b_j^{\ell m} \sigma_\ell^2 \frac{2}{M\epsilon_t} e_t(\lambda_n) \tag{20}$$

$$[\vartheta = \vartheta_i, \quad \lambda = \lambda_n \ , \ t = t(m)]$$

i.e. the sought solution. Again it is trivial, using Appendix B, to prove that the reverse is also true.

Finally, the estimation error of $T_{\ell m}$ is given by

$$\mathscr{E}_{\ell m}^2 \equiv E\left\{ (\widehat{T}_{\ell m} - T_{\ell m})^2 \right\} \equiv \sigma_\ell^2 - \sum_{(in)} \Lambda_{in}^{\ell m} C(L_{\ell m}, Q_{oin})$$

$$\equiv \sigma_\ell^2 - \sum_{(ij)} \sigma_\ell^2 b_i^{\ell m} \left(\tilde{\Gamma}^r + \frac{2}{M\epsilon_t} \Xi \right)_{ij}^{-1} b_j^{\ell m} \sigma_\ell^2$$

$$\cdot \sum_{(n)} \frac{2}{M\epsilon_t} e_t(\lambda_n) e_m(\lambda_n)$$

On the other hand, recalling Eq. (B12) and (13), we have

$$\sum_{n=0}^{M-1} e_t(\lambda_n) e_m(\lambda_n) = \sum_{t=0}^{M-1} e_t(\lambda_n)^2 = \frac{M\epsilon_t}{2}$$

so that the final formula is obtained as

$$\mathscr{E}_{\ell m}^2 = \sigma_\ell^2 - \sigma_\ell^4 \sum_{(ij)} b_i^{\ell m} \left(\tilde{\Gamma}^t + \frac{2}{M\epsilon_t} \Xi \right)_{ij}^{-1} b_j^{\ell m} \tag{21}$$

Note that in Eq. (21) t has to be considered as function of m according to Eq. (13).

4 Numerical procedures

In this section we aim to re-organize in an advantageous and fast manner the solution procedure explained in Sect. 3.

Summarizing and re-organizing all the formulas of Sect. 3, we find that a 'solution' can be sequentially computed through

$$\Delta_{im}^{\ell m} = \sum_{j=1}^{N} \left(\tilde{\Gamma}^t + \frac{2}{M\epsilon_t} \Xi \right)_{ij}^{-1} b_i^{\ell m} \sigma_\ell^2 \tag{22}$$

$$\hat{Q}_{oim} = \frac{1}{M} \sum_{k=0}^{M-1} e_m(\lambda_k) Q_{oik} \tag{23}$$

$$\hat{T}_{\ell m} = \frac{2}{\epsilon_m} \sum_{i=1}^{N} \Delta_{im}^{\ell m} \hat{Q}_{oim} \tag{24}$$

$$\mathscr{E}_{\ell m}^2 = \sigma_\ell^2 \left[1 - \sum_{i=1}^{N} \Delta_{im}^{\ell m} b_i^{\ell m} \right] \tag{25}$$

In particular Eq. (22) derives from Eq. (19) for $t = r = m$ and Eqs. (24) and (25) justify that $\Delta_{im}^{\ell m}$ are the only coefficients needed, i.e. one N-vector for every ℓ, m; moreover, Eq. (24) derives from a combination of Eqs. (10) and (19). Therefore the main numerical effort is in computing Eq. (22), i.e. in solving the linear systems

$$\sum_{j=1}^{N} \left(\tilde{\Gamma}_{ij}^m + \frac{2}{M\epsilon_m} \xi_i \delta_{ij} \right) \Delta_{jm}^{\ell m} = \sigma_\ell^2 b_i^{\ell m} \tag{26}$$

once more we underline that each system of Eq. (26) has dimension $N \times N$, namely 360×360 in the cases of highest resolution foreseeable today, when treating global data sets.

Note: remember that N is the number of parallels along which data are given, so that $N = 360$ corresponds to a 0.5° resolution.

The only problem we have to face here is the computation of the matrices $\tilde{\Gamma}^m \equiv \{\tilde{\Gamma}_{ij}^m\}$, considering that to follow the definition of Eqs. (7) and (9) is a fairly inefficient approach.

A more viable solution is to go back to the matrix

$$C_{ik,jn} \equiv E\{B_{ik}(T) B_{jn}(T)\}$$

$$\equiv B_{P_{ik}} \{ B_{P_{jn}} C_{TT}(P_{ik}, P_{jn}) \} \cdot C_{TT}(P_{ik}, P_{jn}) \tag{27}$$

$$= E\{T(P_{ik}) T(P_{jn})\}$$

$$\equiv \sum_{\ell=0}^{+\infty} \frac{\sigma_\ell^2}{2\ell + 1} P_\ell(\cos \psi_{P_{ik}, P_{jn}}) \equiv C(\psi_{P_{ik}, P_{jn}}) \tag{28}$$

where P_{ik}, P_{jn} are thought of as grid points on the unit sphere, with spherical distance $\psi_{P_{ik}, P_{jn}}$, and the functionals B_{ik} are given by Eq. (2).

Typically we have analytical models for C_{TT}, so that Eq. (28) allows a relatively simple direct computation of the matrix $C_{ik,jn}$ by applying analytically the observation functionals B_{ik}.

Then, recalling Eq. (8), we see that

$$C_{ik,jn} \equiv \sum_{m=-(L-1)}^{L} e_m(\lambda_k) e_m(\lambda_n) \tilde{\Gamma}_{ij}^m \tag{29}$$

and it is from this formula that we want to determine $\tilde{\Gamma}_{ij}^m$.

A formal solution of Eq. (29) is [cf. Eq. (B12) in Appendix B]

$$[-(L-1) \le m \le L]$$

$$\tilde{\Gamma}_{ij}^m = \left(\frac{2}{M\epsilon_m} \right)^2 \sum_{k,n=0}^{M-1} C_{ik,jn} e_m(\lambda_k) e_m(\lambda_n) \tag{30}$$

which looks like a double DFT (discrete Fourier transform) along parallels i and j.

As they are written the numbers $\tilde{\Gamma}_{ij}^m$ are a very large set, namely $N^2 \cdot M$, which for the reference case of $0.5^\circ \times 0.5^\circ$ resolution amounts to more than $93 \cdot 10^6$ numbers. On the other hand, if we take advantage of a fast Fourier algorithm we have to compute at the same time $\tilde{\Gamma}_{ij}^m$ for ij fixed and all the values of m, so that we are likely to be obliged to store the whole data set together. Fortunately a number of symmetries can be exploited to save memory and computer time.

First of all $\tilde{\Gamma}_{ij}^{m}$ are symmetric in the exchange $i \leftrightarrow j$. In addition, in all the examples reported in Sect. 2 we have either

$$b_i^{\ell,-m} \equiv b_i^{\ell,m} \tag{31}$$

or

$$b_i^{\ell,-m} \equiv -b_i^{\ell,m} \tag{32}$$

In this case we immediately see from Eq. (7) that

$$\tilde{\Gamma}_{ij}^{m} \equiv \tilde{\Gamma}_{ij}^{-m}$$

which, from Eq. (9), implies

$$\tilde{\Gamma}_{ij}^{m} \equiv \tilde{\Gamma}_{ij}^{-m} \tag{33}$$

Accordingly, under the hypothesis of Eq. (31) or Eq. (32), we find that

$$C_{ik,jn} = F_{i,j,k-n} \equiv \sum_{m=0}^{L} \tilde{\Gamma}_{ij}^{m} \cos m(\lambda_k - \lambda_n) \tag{34}$$

where

$$\lambda_k - \lambda_n = (k-n)\frac{2\pi}{M} \tag{35}$$

which is straightforwardly inverted by

$$(0 \le m \le L, i \le j)$$

$$\tilde{\Gamma}_{ij}^{m} \equiv \frac{2}{\epsilon_m M} \sum_{k=0}^{M-1} C_{ik,jo} \cos m\lambda_k \tag{36}$$

Already at this point we have divided the total number $\tilde{\Gamma}_{ij}^{m}$ by a factor of about 4; further savings are possible when the grid is also symmetrical in the index i, with respect to the equator. In this case we can exploit the well-defined parity of each Legendre function $\bar{P}_{\ell m}$.

In conclusion, by observing Eq. (36) we can say that we have to compute the covariance matrix C, by fixing the point $\lambda_n = 0$ on a parallel $\vartheta = \vartheta_j$ and computing all the covariances of the observational functionals $B_{jo}(T)$ and $B_{ik}(T)$ for all $\vartheta_i \le \vartheta_j$, and for all the $k = 0, \ldots, N-1$ values; each string of data is then cosine transformed to provide $\tilde{\Gamma}_{ij}^{m}$. Let us stress also that, since $\cos\lambda_{M-k} = \cos\lambda_k$ and $C_{iM-k,jo} \equiv C_{ik,jo}$, the covariances $C_{ik,jo}$ need to be computed only up to $\lambda_k = \pi$.

The final result is that problems like these, satisfying our hypotheses of Eqs. (2) and (31) or (32), can be treated, as we will see in the next section, with very realistic settings, on a simple modern PC.

5 Examples

A FORTRAN program 'sphgrid.f' has been written following the principles described in Sect. 4. However, not all the functionals discussed in Sect. 2 can be used in the present version of the program – only the height anomaly, the radial derivatives of first and second order, and a gravity anomaly functional, where the derivative is taken with respect to the radius vector, r. Mean gravity anomalies are treated as the sum of values (here $n \times n$), not at the same radial distance, but at distances corresponding to the same ellipsoidal height. This corresponds approximately to how mean gravity anomalies are computed in practice.

The program has been tested for small grids, by verifying that the 'full' collocation solution and the 'fast' solution were identical.

Covariance functions must be given by providing a set of degree variances. This means that the variance will be too large at the poles and too small at the equator, see Sect. 5.3. (This issue will be treated in a forthcoming paper.)

The program may use spherical approximation or no approximation, (see Sect. 5.1) and may combine data at different altitudes and of different kinds (see Sect. 5.2). Parallels may occur several times, with different observations such as ground gravity anomalies and satellite-tracked second-order derivatives.

The program may handle very large data sets. Solutions to degree 720 have been computed using simulated data. In Sect. 5.3 we describe the computation of a degree 360 solution using real data. The EGM96 (Lemoine et al. 1998) gravity data have been used.

5.1 Errors due to spherical approximation

Spherical approximation is in general use in gravity-field-related computations. The error is supposed to be of the order of the flattening, i.e. 1/300 (see e.g. Moritz 1980).

Spherical approximation has two different meanings in geodetic literature. One is just to substitute the normal potential with the spherical potential μ/r in coefficients of observational functionals; the effect of this approximation has been studied for a long time and it is known to be of the order of flattening. The other kind of approximation (see e.g. Moritz 1980) is to make a coordinate mapping which transforms the Earth ellipsoid into the mean Earth sphere; as a result the expression of the Laplacian in these new 'mean spherical' coordinates is changed. The perturbations are known to be continuous in the flattening parameter and they are assumed to be of the same order of magnitude. In this experiment we show that this is not true.

Pertubations of the EGM96 coefficients were calculated so that they had a standard deviation equal to the EGM96 error degree variances (see Fig. 1). A covariance function consistent with this was defined using the error degree variances of EGM96 up to degree 180 as degree variances, and the GPM98 (Wenzel 1998) degree variances from degree 181 to degree 720. Gravity anomalies (between the perturbed and the original EGM96 model) in a regular grid with a spacing of 1 degree in geodetic latitude and longitude were generated at 300 km altitude over the Earth ellipsoid. No noise was added to the data, but in the calculations a noise standard deviation of 0.03 mGal was used.

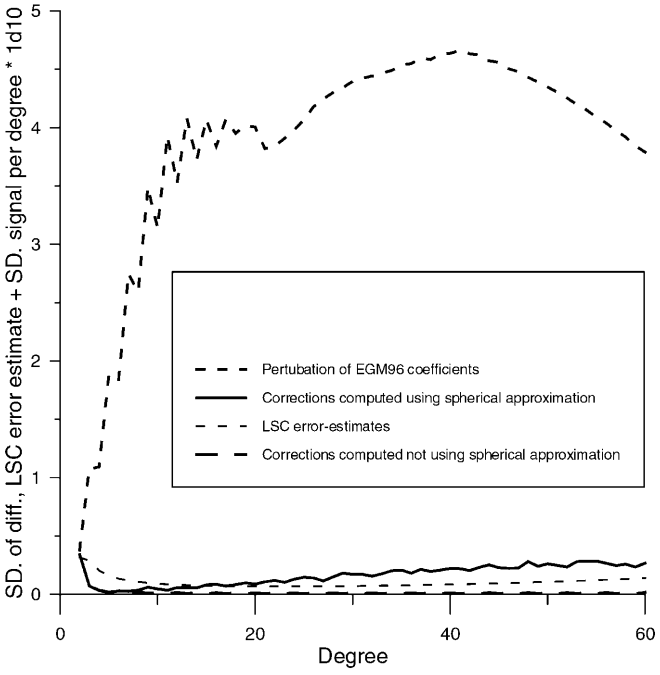

Fig. 1. Standard deviation of EGM96 perturbed coefficients and corrections per degree computed using or not using spherical approximation

Since actual data are now on exactly circular parallels in space, although not belonging to the same sphere, we can perform our collocation solution without further approximation. This has been computed with a 'spherical approximation' solution obtained by first mapping the actual observational configuration on a sphere at 300 km height over the mean Earth sphere.

Figure 1 shows the resulting standard deviations of the differences between the predicted coefficients and the 'true' ones per degree. The relative error obtained using 'spherical approximation' in the above sense is not of the order of the flattening, but between 0.1 and 0.05. The error obtained by not using any approximation is 100 times smaller. The formal error estimate is nearly the same in both cases, since it depends mainly on the covariance function. The error obtained when not using spherical approximation is somewhat smaller than the formal error estimate due to the fact that no noise was added to the data. However, for spherical approximation the error was much larger than the estimated error. As a conclusion we could say that further investigations on this type of error are needed.

5.2 Data combinations

In order to illustrate the potential of the method, it has been used to study the influence of the polar gaps on a spherical harmonic solution based on anomalous vertical gravity gradient datas (cf. Sect. 2). Data were simulated as pertubations of EGM96, but in fact not needed, since what we wanted to study were the error estimates of the spherical harmonic coefficients. Five of the many experiments conducted are described here. A grid with 2-degree spacing in longitude was used in all five cases: anomalous gravity gradient data on parallels spaced 2 degrees apart from latitude −82 degrees to 82 degrees, altitude 300 km. Noise standard deviation equal to 0.005 EU was assigned to the data. The same kind of data was added at the parallels with latitude 84, 86 and 88 degrees, and at parallels with latitude −84, −86 and −88 degrees. Anomalous gravity was added to the data set used in Eq. (1) at the parallels between 83 and 89 with 1-degree spacing. The altitude used was 5 km and the data noise was 0.2 mGal. Data with the same spacing was added in Antarctica.

In Fig. 2 the effect of the polar gap on the coefficients of degree 28 is clearly seen, and also the effect of covering the gaps with gradient data. Degree 28 was chosen because its behaviour was representative of that of all the other degrees. Experiments with gravity anomalies added in different altitudes and with different noise assigned were conducted. By 'iteration' it was found that rather high-resolution gravity anomalies (5 km altitude corresponding to 0.5 degree means), spaced with the double density as compared to the gravity gradients, were needed to obtain approximately the same effect as when gravity gradient values were added (see Fig. 3). The experiments correspond quite well to the situation we will have using GOCE (European Space Agency 1999) data, but after we have improved the spherical harmonics using CHAMP (GFZ 2001). This does not correspond to the present situation, because an isotropic covariance function was used in the calculations. The use of a uniform noise variance for the gradient data corresponds very well to what we expect from the GOCE mission.

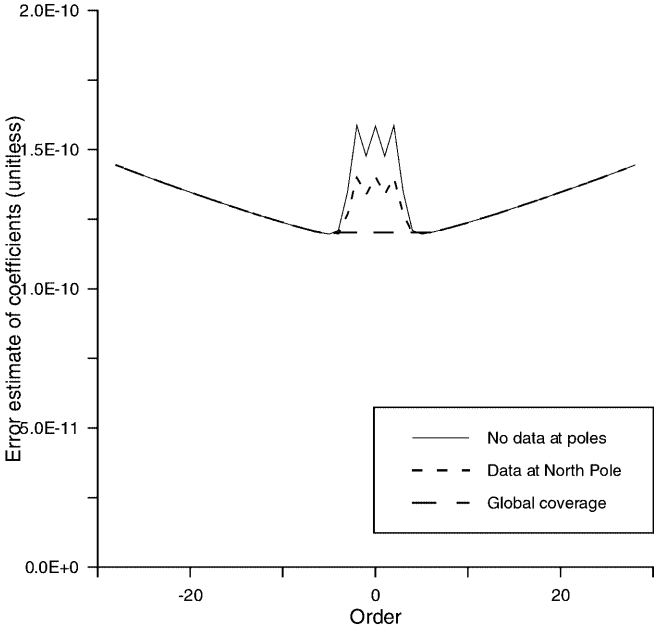

Fig. 2. Error estimates for degree 28. Only vertical gravity gradient data used. Longitude equidistance 2 degrees

5.3 Solution using EGM96 0.5-degree equi-angular mean free-air anomalies

EGM96 has been computed using satellite orbit perturbations and ground gravity. The gravity data were 0.5-degree mean values, downward continued to the ellipsoid. Error estimates had been assigned so that the gravity data would not deteriorate the satellite data, and the gravity data were used to improve on a satellite-only solution to degree 72. Consequently in this study we subtracted the contribution of EGM96 to degree 72, and used as degree variances in the covariance function the error degree variances of EGM96 to degree 72. This gives a strong a priori constraint on the coefficients up to degree 72, the effect of which can be seen in Figs. 4 and 5.

For degree 73–720 we used degree variances calculated from Wenzel's GPM98 coefficients (see Figs. 6 and 7), where the square roots of the degree variances are shown. It should be noted that a consequence of using this kind of isotropic model is that the signal variance will vary with latitude. The factor R/r, where R is the Earth mean radius, will be small at the poles and large at the equator. (In degree-variance models which have an infinite number of coefficients, a Bjerhammar sphere with a radius smaller than the semi-minor axis of the Earth must be used. Such models have not been included in the program in order to keep the program simple.)

A number of numerical experiments were conducted, treating the mean values as point values or as mean values and assigning different error estimates to the data. The radial derivative was used, while the data are associated with a derivative close to the normal to the ellipsoid. In the results of the numerical experiments shown here, we used the error estimates of EGM96 to calculate mean values of error estimates for each parallel, which were then assigned to all data on the same parallel. In Fig. 4 and 5 the results are shown treating the data as point values.

We see that rather large differences with respect to the EGM96 coefficients exist.

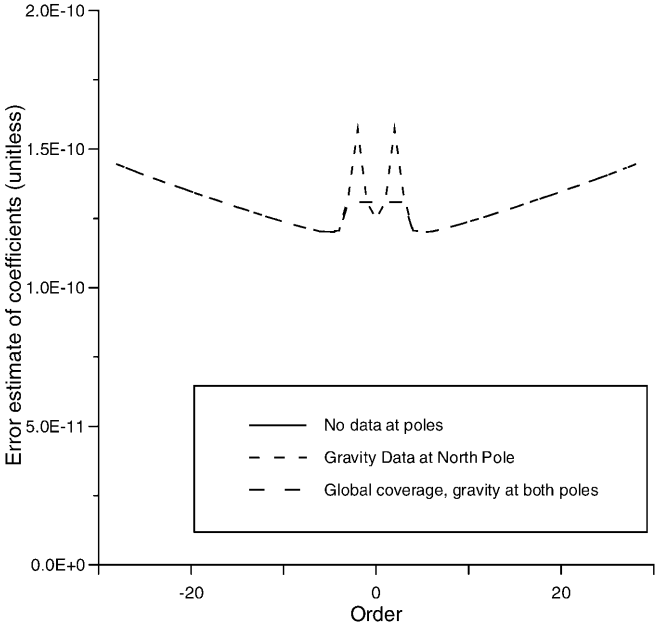

Fig. 3. Error estimates for degree 28. Vertical gravity gradient data and gravity anomalies at 5 km. Longitude equidistance 2 degrees for gradients, 1 degree for gravity

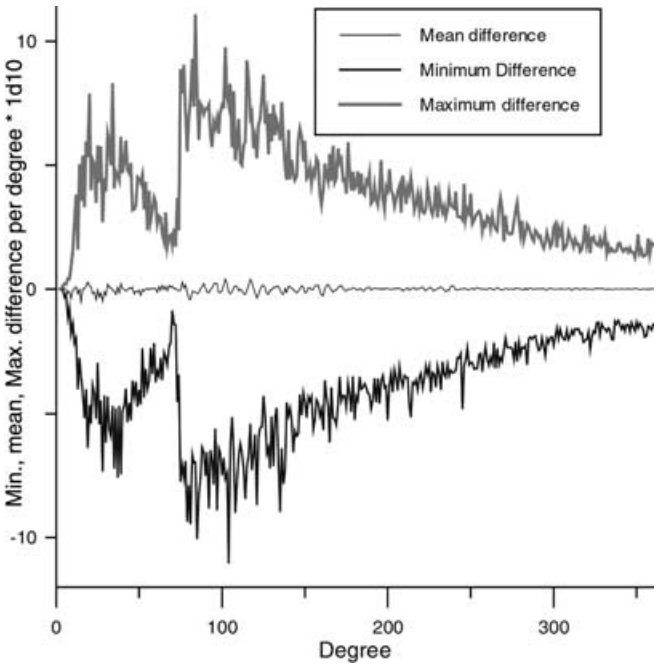

Fig. 4. Mean, max and min differences between EGM96 and predicted coefficients from data spaced at 0.5 degrees. EGM96 original 0.5-degree mean values treated as point gravity anomalies

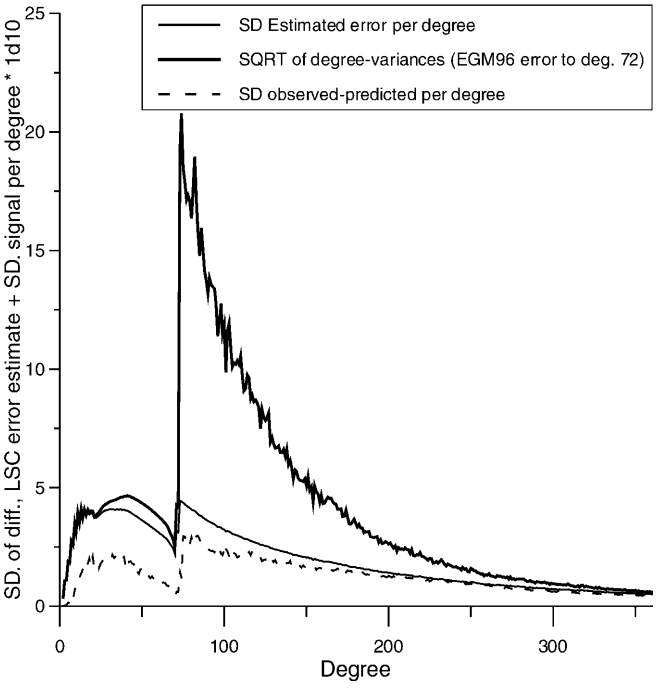

Fig. 5. Standard deviation of degree variances, of error estimates and of differences observed–predicted, 0.5-degree spacing. EGM96 gravity anomalies used as point data

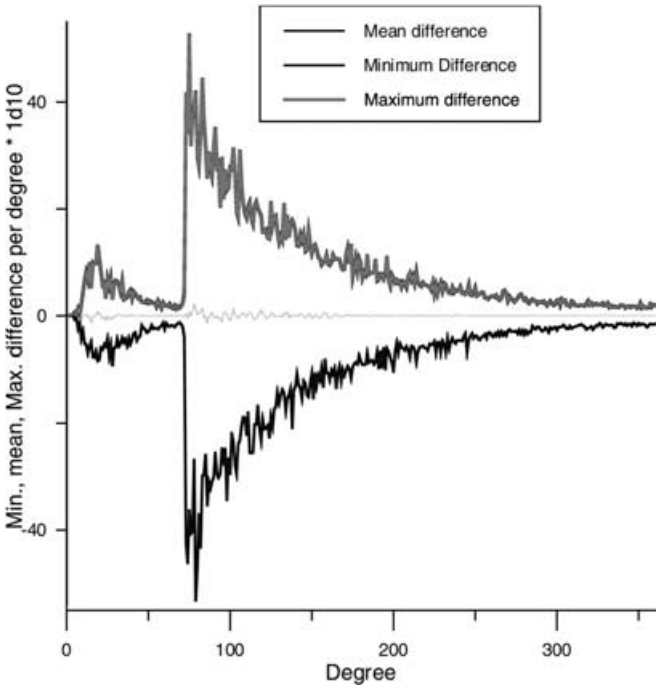

Fig. 6. Mean, max and min differences between EGM96 and coefficients predicted from EGM96 0.5-degree mean values spaced at 0.5 degrees. EGM96 to degree 72 subtracted

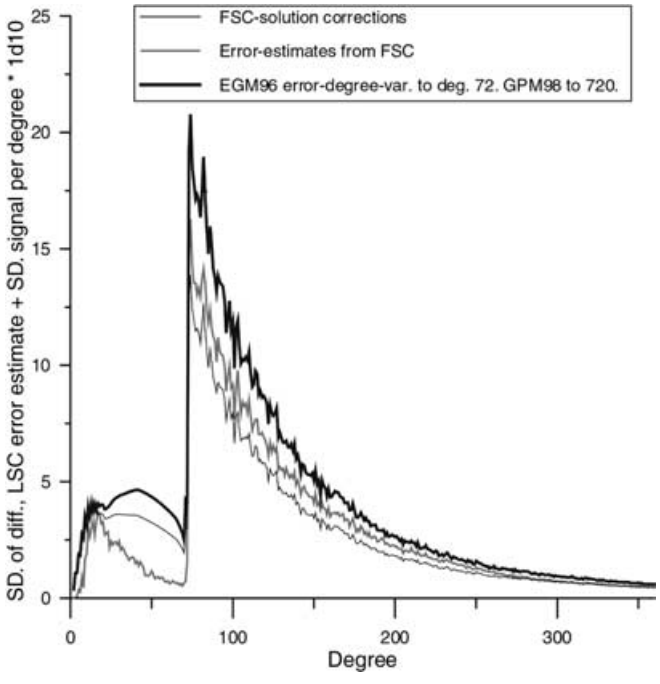

Fig. 7. Standard deviation of degree variances, of collocation error estimates and of differences observed–predicted, 0.5-degree spacing, EGM96 gravity anomalies used. Coefficients computed from EGM96 input 0.5-degree mean gravity minus EGM96 gravity to degree 72

The 'corrections' to the coefficients were added to the EGM96 coefficients, and the coefficients were used to calculate mean gravity anomalies which were compared with the original EGM96 values. Using the unperturbed EGM96 coefficients gave a considerable better fit to the data compared to that using the 'corrected' coefficients.

Here we have to keep in mind that a collocation solution will give an exact fit to noise-free data, and a fit within the 'noise band' for noisy data. Hence had we computed estimates for all coefficients up to degree 720, a perfect fit would have been obtained – especially if we had assigned a small noise to the data. We want to study these findings in depth, because they could help to explain some of the long-wavelength errors found in geoid calculations where EGM96 is used in a remove–restore procedure. It may be that methods which are used to determine coefficients, and a priori project the solution to a space of low dimension, will 'press' too much of the information into these coefficients.

6 Conclusion

The results of Colombo (1979) have been generalized so that data on grids with parallels in varying altitude and with different data types can be processed very efficiently using LSC. No spherical approximation is necessary. However, the use of an isotropic covariance implies that signal variances generally will be too low at the equator and too large at the poles. Solutions to this problem are being considered.

The fast spherical collocation method has been used to demonstrate the problems associated with the use of spherical approximation. It has been used to simulate results where satellite gravity gradient data were combined with ground gravity. Furthermore, the method has been used to compute a set of spherical harmonic coefficients and associated error estimates to degree and order 360 using the same ground gravity data as used when determining EGM96. The differences with respect to EGM96 need to be further analysed in order to be understood.

Acknowledgements. The authors are extremely grateful to the reviewers; due to their precise and open-minded work, a number of corrections and improvements have been made to the paper.

Appendix A

A generalization

When the observational functionals B_{ik} still satisfy the condition of Eq. (2), we can apply the *same* reasoning as developed in the present paper to the estimation of any linear functional $A(T)$

$$A(T) = \sum_{p,q} T_{pq} a^{pq} \tag{A1}$$

such that

$$\sum_{p,q} \sigma_p^2 (a^{pq})^2 < +\infty$$

We now put

$$A(\hat{T}) \equiv \sum_{(ik)} \Lambda^A_{ik} Q_{oik} \tag{A2}$$

and Λ^A_{ik} is determined by an equation similar to Eq. (12), but for the known term which has to be as in the following:

$$\begin{aligned}
&\sum_{(in)} C(Q_{oik}, Q_{ojn}\Lambda^A_{jn} \\
&\quad = C\{A(T), Q_{oik}\} \equiv C\{A(T), B_{ik}(T)\} \\
&\quad \equiv \sum_{\ell',m'} \sigma^2_{\ell'} a^{\ell' m'} b_i^{\ell' m'} e_{m'}(\lambda_k) \\
&\quad \equiv \sum_{m'=-\infty}^{+\infty} e_{m'}(\lambda_k) \sum_{\ell'=|m|}^{+\infty} \sigma^2_{\ell'} a^{\ell' m'} b^{\ell' m'}_{b_i}
\end{aligned} \tag{A3}$$

Now, if we put

$$\eta_i^{m'} \equiv \sum_{\ell'=|m'|}^{+\infty} \sigma^2_{\ell'} a^{\ell' m'} b_i^{\ell' m'} \tag{A4}$$

and, similar to Eq. (9)

$$\begin{aligned}
&\tilde{\eta}_i^o = \sum_{h=0}^{+\infty} \eta_i^{hM} \\
&\tilde{\eta}_i^L = \sum_{h=0}^{+\infty} \eta_i^{L+hM} \\
&(u < m' < L),\ \tilde{\eta}_i^{m'} = \sum_{h=0}^{+\infty} \eta_i^{m'+hM} + \eta_i^{(h+1)M-m'} \\
&[-(L-1) \le m' < 0], \tilde{\eta}_i^{m'} = \sum_{h=0}^{+\infty} \eta_i^{m'-hM} - \eta_i^{-m'-(h+1)M}
\end{aligned} \tag{A5}$$

we find

$$\begin{aligned}
C\{A(T), Q_{oik}\} &\equiv \sum_{m'=-\infty}^{+\infty} e_{m'}(\lambda_k)\eta_i^{m'} \\
&\equiv \sum_{m'=-(L-1)}^{L} e_{m'}(\lambda_k)\tilde{\eta}_i^{m'}
\end{aligned} \tag{A6}$$

Remark. Beware that in Eq. (A5) we have the sign $(-)$ in the expression of $\tilde{\eta}_i^{m'}$, when $m' < 0$, because, contrary to Eq. (9), we have here only one $e_{m'}(\lambda_k)$.

So, multiplying Eq. (39) by $e_r(\lambda_k), [-(L-1) \le r \le L]$, and summing over k, we obtain

$$\sum_{(j)} \tilde{\Gamma}^r_{ij} \left(\sum_{(n)} e_r(\lambda_n)\Lambda^A_{jn} \right) + \frac{2}{M\epsilon_r} \xi_i \sum_{(k)} \Lambda^A_{ik} e_r(\lambda_k) \equiv \tilde{\eta}^r_i \tag{A7}$$

The solution of Eq. (43) is

$$\sum_{n=1}^{M-1} \Lambda^A_{in} e_r(\lambda_k) = \sum_i \left(\tilde{\Gamma}^r + \frac{2}{M\epsilon_r}\Xi \right)^{-1}_{ij} \tilde{\eta}^r_j \tag{A8}$$

and then

$$\Lambda^A_{in} = \sum_{r=-(L-1)}^{L} \frac{2e_r(\lambda_k)}{M\epsilon_r} \left(\sum_{(j)} \left(\tilde{\Gamma}^r + \frac{2}{M\epsilon_r}\Xi \right)^{-1}_{ij} \tilde{\eta}^r_j \right) \tag{A9}$$

Finally, the estimation error is

$$\begin{aligned}
\mathscr{E}^2(A) &= E\{A(T)^2\} - \sum_{in} \Lambda^A_{in} E\{A(T), Q_{oin}\} \\
&\equiv \sum_{pq} \sigma^2_p (a^{pq})^2 - \sum_{in} \Lambda^A_{in} \sum_{m'=-(L-1)}^{L} e_{m'}(\lambda_n)\tilde{\eta}^{n1}_j \\
&\equiv \sum_{pq} \sigma^2_p (a^{pq})^2 - \sum_{r=-(L-1)}^{L} \sum_{ij} \tilde{\eta}^r_i \left(\tilde{\Gamma}^r + \frac{2}{M\epsilon_r}\Xi \right)^{(-1)} \tilde{\eta}^r_j
\end{aligned}$$

Appendix B

A syllabus of real discrete Fourier transform and Nyquist frequency

The subject of this appendix is well known and described in all textbooks on signal analysis. However, we find it useful to report it in this form, where the transformation from the complex basis to the real one, particularly with the labelling we use in geodesy, is made explicitly.

Readers are warned to pay attention to the range of integers explicitly indicated into the formulas.

Assume we have a real periodical and continuous $f(\lambda),\ 0 \le \lambda < 2\pi$

$$f(\lambda) = a_0 + \sum_{m=1}^{+\infty} (a_m \cos m\lambda + b_m \sin m\lambda) \tag{B1}$$

sampled at the M points

$$\begin{aligned}
&\lambda_k = k\delta = k\frac{2\pi}{M} \\
&k = 0, 1, \ldots, M-1 \\
&M = 2L \quad \text{(hypothesis!)}
\end{aligned}$$

As we know, we obviously have

$$\begin{aligned}
&\cos(m + hM)k\delta \equiv \cos mk\delta \\
&\qquad\qquad \forall\, h(\text{integer}) \\
&\sin(m + hM)k\delta \equiv \sin mk\delta
\end{aligned} \tag{B2}$$

But there is also another identification

$$\cos k\delta \equiv \cos(M-1)k\delta \quad \forall\, k = 0, \ldots, M-1$$

because

$$\cos(M-1)k\delta = \cos Mk\delta \cos k\delta + \sin Mk\delta \sin k\delta$$

Moreover

$$\sin k\delta \equiv -\sin(M-1)k\delta \qquad \forall\, k$$

because

$$\sin(M-1)k\delta \equiv \sin Mk\delta \cos k\delta - \cos Mk\delta \sin k\delta$$

In the same way

$$\cos 2k\delta \equiv \cos(M-2)k\delta \quad \forall k$$
$$\sin 2k\delta \equiv -\sin(M-2)k\delta \quad \forall k$$

and, up to L

$$\cos Lk\delta = \cos k\pi = (-1)^k \quad \forall k$$
$$\sin Lk\delta = \sin k\pi \equiv 0 \quad \forall k$$

So we can write Eq. (B1) for $\lambda = \lambda_k = k\delta$ as

$k = 0, 1, \ldots, M-1$

$$\begin{aligned} f(\lambda_k) = &\left(\sum_{h=0}^{+\infty} a_{hM}\right) + \sum_{h=0}^{+\infty} a_{L+hM} \cos L\lambda_k \\ &+ \sum_{m=1}^{L-1}\left(\sum_{h=0}^{+\infty} a_{m+hM} + \sum_{h=1}^{+\infty} a_{hM-m}\right)\cos m\lambda_k \\ &+ \sum_{m=1}^{L-1}\left(\sum_{h=0}^{+\infty} b_{m+hM} - \sum_{h=1}^{+\infty} b_{hM-m}\right)\sin m\lambda_k \end{aligned} \tag{B3}$$

In other words, when $f(\lambda)$, [Eq. (B1)] is sampled at λ_k we can write

$$f(\lambda) = \tilde{a}_0 + \sum_{m=1}^{L} \tilde{a}_m \cos m\lambda_k + \sum_{m=1}^{L-1} \tilde{b}_m \sin m\lambda_k \tag{B4}$$

where the lumped coefficients $\tilde{a}, \tilde{b}$ are given by

$$\begin{aligned} \tilde{a}_0 &= \sum_{h=0}^{+\infty} a_{hM}, \quad \tilde{a}_L = \sum_{h=0}^{+\infty} a_{L+hM} \\ \tilde{a}_m &= \sum_{h=0}^{+\infty} a_{m+hM} + \sum_{h=1}^{+\infty} a_{hM-m}, \quad (0 < m < L) \end{aligned} \tag{B5}$$

$$\tilde{b}_m = \sum_{h=0}^{+\infty} b_{m+hM} - \sum_{h=1}^{+\infty} b_{jM-m}, \quad (0 < m \le L) \tag{B6}$$

L is the Nyquist frequency and, if there is no signal power above L, we have

$$\tilde{a}_m = a_m, \quad \tilde{b}_m = b_m$$

Remark. We can count

$$\left.\begin{array}{l}\text{Number } (\tilde{a}) = L+1 \\ \text{Number } (\tilde{b}) = L-1\end{array}\right\rangle \text{Number (coefficient)} = 2L = M$$

i.e. in Eq. (B4) we have exactly M samples from 0 to $(M-1)$ and M coefficients.

Discrete orthogonality and completeness: finally, from the well-known (and easy-to-prove) relation

$$\sum_{k=0}^{M-1} e^{i(m-n)k\delta} = M\delta_{m,n}, \quad 0 \le m, n \le M-1 \tag{B7}$$

we have also

$$\begin{aligned} &(0 \le m, n \le M-1) \\ &\sum_{k=0}^{M-1} \cos(m-n)k\delta = \delta_{m,n} M \\ &\sum_{k=0}^{M-1} \sin(m-n)k\delta \equiv 0 \end{aligned} \tag{B8}$$

Accordingly we obtain

$$(0 \le m, n < L)$$

$$\begin{aligned} &\sum_{k=0}^{M-1} \cos mk\delta \cos nk\delta \\ &\quad = \frac{1}{2}\sum_{k=0}^{M-1} \cos(m+n)k\delta + \frac{1}{2}\sum_{k=0}^{M-1} \cos(m-n)k\delta \\ &\quad = \frac{1+\delta m_0}{2} M\delta_{mn} = \begin{cases} M & m=0, n=0 \\ \frac{M}{2}\delta_{mn} & \text{all other cases} \end{cases} \end{aligned}$$

because $m+n < 2L = M$, although indeed we can have $m = n = 0$. Moreover

$$\sum_{k=0}^{M-1} (\cos Lk\delta)^2 = M$$

because $\cos Lk\delta = (-1)^k$.

This can be summarized as

$$\begin{aligned} &(0 \le m, n \le L) \\ &\sum_{k=0}^{M-1} \cos mk\delta \cos nk\delta \\ &\quad \equiv \frac{(1+\delta_{m0})(1+\delta_{mL})}{2} M\delta_{mn} \equiv \epsilon_m \frac{M}{2}\delta_{mn} \end{aligned} \tag{B9}$$

Moreover we have, for $1 \le m, n \le L-1$,

$$\begin{aligned} &\sum_{k=0}^{M-1} \sin mk\delta \sin nk\delta = \frac{1}{2}\sum_{k=0}^{M-1} \cos(m-n)k\delta \\ &\quad - \frac{1}{2}\sum_{k=0}^{M} \cos(m+n)k\delta = \frac{M}{2}\delta_{mn} \end{aligned} \tag{B10}$$

because it is always $0 < m+n < 2L = M$.

Similarly, we prove that

$$\begin{aligned} &\sum_{k=0}^{M-1} \cos mk\delta \sin nk\delta \\ &\quad = \frac{1}{2}\sum \sin(n+m)k\delta \\ &\quad - \frac{1}{2}\sum \sin(n-m)k\delta \equiv 0 \end{aligned} \tag{B11}$$

So finally, if we use the notation

$$e_m(\lambda) = \begin{cases} \cos m\lambda & m \geq 0 \\ -\sin m\lambda & m < 0 \end{cases}$$

and we take $m = -(L-1) \ldots 0 \ldots L$ [$2L$ coefficients, note that $e_{-L}(\lambda_k) \equiv 0$], we can say that the fundamental discrete orthogonality relation holds

$$[-(L-1) \leq (m,n) \leq L]$$
$$\sum_{k=0}^{M-1} e_m(\lambda_k) e_n(\lambda_k) = \frac{M}{2} \epsilon_m \delta_{mn} \tag{B12}$$
$$\epsilon_m = \begin{cases} 2 & m = 0, m = L \\ 1 & 0 < m < L \end{cases}$$

Let us observe that despite the different set of integers on which m and k run, $-(L-1) \leq m \leq L$, $0 \leq k \leq M-1$, since both are arrays of length $2L = M$, the matrix

$$E \equiv \sqrt{\frac{2}{M\varepsilon_m}} \{e_m(\lambda_k)\}$$

is square, so that Eq. (B12) can be written

$$E \cdot E^{+} = I$$

On the other hand, we know that this relation implies

$$E^{+} \cdot E = I$$

too, which in indexes reads

$$[0 \leq (k,j) \leq M-1]$$
$$\sum_{m=-(L-1)}^{L} \frac{1}{\epsilon_m} e_m(\lambda_k) e_m(\lambda_j) = \frac{M}{2} \delta_{kj} \tag{B13}$$

This is in a sense the reciprocal of Eq. (B12), also known as the completeness relation.

References

Colombo O (1979) Optimal estimation from data regularly sampled on a sphere with applications in geodesy. Rep. 291, Department of Geodetic Science, The Ohio State University, Columbus

European Space Agency (1999) Gravity field and steady-state ocean circulation mission. ESA SP-1233(1), ESA Publication Division, c/o ESTEC, Noordwijk

Freeden W, Gervens T, Schreiner M (1999) Constructive approximation on the sphere. Clarendon Press, Oxford

GFZ (GeoForschungsZentrum) (2001) Announcement of opportunity for Champ. GFZ, Potsdam

Krarup T (1969) A contribution to the mathematical foundation of physical geodesy. Geodaetisk Institut Rep no. 44, Köbenhavn

Lemoine FG, Kenyon SC, Factor JK, Trimmer RG, Pavlis NK, Chinn DS, Cox CM, Klosko SM, Luthcke SB, Torrence MH, Wang YM, Williamson RG, Pavlis EC, Rapp RH, Olson TR (1998) The development of the joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) geopotential model EGM96. NASA/TP-1998-206861, Goddard Space Flight Center, Greenbelt, MD

Moritz H (1980) Advanced physical geodesy. Herbert Wichmann, Karlsruhe

Sneeuw N, Dorobantu R, Gerlach C, Müller J, Oberndorfer H, Rummel R, Koop R, Visser P, Hoyng P, Selig A, Smit M (2001) Simulation of the GOCE gravity field mission. In: Benciolini B (ed) Proc. IV Hotine–Marussi Symposium on Mathematical Geodesy, Trento, 14–17, September 1998. IAG Symposia, 122. Springer, Berlin Heidelberg, New York

Tscherning CC (2001) Computation of spherical harmonic coefficients and their error estimates using least squares collocation. J. Geod. 75: 14–18

Wenzel HG (1998) Ultra hochaufloesende Kugelfunktionsmodelle GMP98A und GMP98B des Erdschwerefeldes. Proc Geodaetische Woche, Kaiserslautern