

TOWARD MULTIREOLUTION ESTIMATION AND EFFICIENT REPRESENTATION OF GRAVITATIONAL FIELDS

GREGORY BEYLKIN and ROBERT CRAMER

Department of Applied Mathematics, University of Colorado at Boulder, Boulder, CO 80309-0526, U.S.A., e-mails: beylkin@colorado.edu; bcramer@colorado.edu

(Received: 22 December 2000; accepted: 14 December 2001)

Abstract. In this paper we demonstrate performance of several local and multi-resolution gravity models derived from existing spherical harmonic models. A model of this type was provided to USAFSC in 1997, and additional variants have been developed since then. We also begin to address the problem of estimating multi-resolution models directly from gravity measurements. We hope to demonstrate that it is reasonable to expect gravity models with greater accuracy and flexibility once the spherical harmonic basis has been eliminated from the process.

Keywords: geopotential modeling, gravity estimation, spherical harmonics, multi-resolution analysis

1. Introduction

For a variety of applications it is necessary to have an accurate model of the Earth's gravitational field. Up to now such models have been constructed using bases of spherical harmonics. However, to meet future demands there is a need to improve both accuracy and efficiency of Earth gravity models beyond what is achievable with the spherical harmonic basis. The use of such basis functions imposes limitations because they are globally supported, making it impossible to increase the resolution of the model without increasing the order and degree globally. We describe below a new approach to the problems of evaluating and estimating the Earth's gravitational field. Namely, by using bases of functions which are locally supported, such as wavelets, we are able to avoid the limitations imposed by the spherical harmonic basis.

We note that our approach to estimating models of Earth's gravitational field is equally applicable to the general problem of estimating gravity fields in the vicinity of other celestial bodies, for example, asteroids.

Let us recall the form of the spherical harmonic model of the gravitational potential,



$$V(r, \phi, \theta) = \frac{\mu}{r} \left\{ 1 + \sum_{n=2}^N \left(\frac{R}{r} \right)^n Y_n(\theta, \phi) \right\}, \quad (1)$$

where μ is Earth's gravitational constant, r is the length of the radius vector from Earth's center of mass, R is the equatorial radius of the Earth, ϕ is the geocentric longitude and θ is the geocentric latitude. The spherical harmonic $Y_n(\phi, \theta)$ is defined as

$$Y_n(\phi, \theta) = \sum_{m=0}^n \bar{P}_n^m(\cos \theta) (\bar{C}_n^m \cos m\phi + \bar{S}_n^m \sin m\phi),$$

where $\{\bar{C}_n^0, \bar{C}_n^1, \dots, \bar{C}_n^n, \bar{S}_n^1, \dots, \bar{S}_n^n\}$ are normalized coefficients, and \bar{P}_n^m are normalized associated Legendre functions of degree n and order m . As it is well known, V is a solution for the Laplace equation in spherical coordinates (r, ϕ, θ) , $r > R$.

The cost of evaluating V at a point (r, ϕ, θ) via (1) grows rapidly with the degree and order. For a model with degree and order equal to N , the number of operations to evaluate V is proportional to N^2 . Thus, evaluating a model of degree and order 180 compared to evaluating a model of degree and order 18 requires roughly 100 times more operations.

We describe below several approaches for restructuring the Earth's gravitational model to improve its performance. The basic idea in all of them is simple: we localize the description of the gravitational field. A local description generally requires more memory to store the model than a spherical harmonic representation. In our approach, we achieve a good compromise by realizing a significant increase in speed with acceptable memory requirements.

In what follows we refer to both 'local' and 'multiresolution' representations. In a local representation, the elementary building blocks are basis functions with localized support (e.g. B-splines), but the resolution is uniform throughout the model and such representations typically are not adaptive. Multiresolution representations, on the other hand, also use localized basis functions (e.g. wavelets), but are adaptive, which means they can incorporate different resolutions to more efficiently accommodate different variability of the geopotential in different locations.

The process of constructing gravitational models directly from measurements, the estimation problem, is intimately connected with the way in which the model is represented. Using spherical harmonics for the construction leads to a number of serious problems, which we now discuss.

The only way to improve resolution within the spherical harmonic representation is to increase the number of terms in the expansion. The estimation procedure which uses satellite measurements requires formation of correlation matrices, the size of which is roughly $N^2 \times N^2$, where N is the degree and order of the spherical harmonic model. These matrices are dense (full), and therefore the number of operations required to store and manipulate them grows rapidly as N is increased.

Correlation matrices are dense because coefficients of the spherical harmonic expansion are not associated with any particular spatial location. Indeed, each coefficient in the expansion contributes to the field at every spatial location. Spherical harmonics are globally supported, oscillatory functions, which depend on cancellation (constructive interference) to achieve the approximation. Changing even a single coefficient in the model has a global effect. It is difficult, if not impossible, to adjust the spatial frequency content of a spherical harmonic expansion locally. In particular, there is a difficulty in incorporating observations of the gravitational potential near the surface with those obtained from satellites, which is due in part to the different spectral contents of the data.

Another important consideration is that the linear system for the L_2 -minimization problem for model estimation has a very large condition number. Coupled with the global nature of spherical harmonic basis functions, this severely limits the overall resolution attainable in the model. We are unable to take advantage of the fact that measurements in some regions are better than in others, and the most poorly sampled region dictates the limit on the degree of resolution that can be attained globally.

Finally, computation of some satellite orbits appears to be very sensitive to the number of spherical harmonic terms retained in the model, and does not follow the common sense rule ‘more terms are better.’ It is indicative of an analogue of Gibbs’ phenomenon since the magnitude of the normalized spherical harmonic model coefficients does not decrease significantly before the model is truncated. (Gibbs’ phenomenon is a type of oscillatory error caused by abrupt truncation of a Fourier series with slowly decaying coefficients. A more detailed description is contained in Appendix B.)

Our goal is to develop multiresolution models of the Earth’s gravity that do not suffer from the difficulties of estimation and evaluation outlined above. Multiresolution models use basis functions with localized support in both space and spectral domains. This allows us to generate models where changes in *most* of the parameters (coefficients) will produce only *local* changes in the model (up to any finite but arbitrary accuracy).

This statement may at first appear unreasonable since gravitational fields always give rise to long-range interactions. To illustrate this point, we consider the analogy with electrostatic fields, which are mathematically equivalent since both satisfy the Laplace equation. The difference is, of course, that electrostatic charges can be both positive and negative, whereas gravitational charges (masses) are always positive. Electrostatic potentials consisting mainly of high-order multipoles are possible, and the effects of such fields can be ignored beyond a relatively short distance.

To form electrostatic multipoles, it is necessary to make use of cancellation between the positive and negative charges, a mechanism which does not exist in a gravitational field. However, we can produce the same effect by using a basis of wavelets with vanishing moments. Most of the functions in such a basis act as multipoles, and the use of such bases results in sparse correlation matrices

for estimation of the model parameters. Such bases also allow us to address the problem of ill-conditioning in an adaptive manner.

The particular choice of basis is a crucial issue since it will be necessary to use several different algorithms in conjunction with the representation. A number of authors (e.g. [1]) have constructed wavelet-type bases that respect the topology of the sphere. Although there certainly is merit to such constructions, we find them to be somewhat cumbersome for our purposes. Our preference is to use a more direct approach, where we employ ‘general purpose’ functions such as splines, polynomials or multiwavelets, making certain that algorithms we use can handle redundancies or slight inefficiencies that the use of such functions may entail.

2. Local and Multiresolution Gravitational Models

We describe several examples of local models. Each model comprises a set of concentric ‘shells.’ Each shell represents the field at a fixed distance from the center of the Earth, and consists of a set of coefficients that represent the geopotential (or one of its derivatives) as a two-dimensional expansion of local basis functions. To compute values of the gravity field at points between the fixed shells we use a local polynomial interpolation.

Each shell represents a scalar function. Hence, to represent the gravitational force field, it is necessary to store three sets of shells. In what follows, for simplicity, we describe representations of the Earth’s geopotential only.

We note that in our models we retain spherical harmonic terms up to degree and order 2 to represent the low-frequency part of the field (more terms could be retained if desired). We made this choice because there is a significant reduction in magnitude of the spherical harmonic coefficients in going from degree 2 to degree 3, but the magnitude decreases only slowly thereafter. The coefficient \bar{C}_2^0 is of magnitude 10^{-3} , whereas for degrees greater than 2, all coefficients are of magnitude 10^{-5} or less.

2.1. DOUBLY-PERIODIC SPLINE MODEL

We begin by describing the simplest of the local models. This model was designed solely for evaluation of the gravitational field, and is not suitable for estimation due to the pole problem, that is over-sampling in the polar regions relative to the sampling rate in equatorial regions.

In this model, the surface of the sphere is mapped to a unit square in the plane. Thus, each concentric shell takes the form of a square. Variables (x, y) on the unit square are related to the spherical coordinates (ϕ, θ) on the surface of the sphere by

$$\phi = 2\pi x, \quad \theta = 2\pi y \quad \text{where} \quad 0 \leq x, \quad y \leq 1. \quad (2)$$

Note that the square covers the sphere twice—we discuss this more fully below.

We introduce a grid on each square which is equispaced in the variables x and y . However, the projection of this grid back onto the spherical surface suffers a severe distortion due to stretching near the poles, since adjacent grid points become closer together the nearer they are to the poles. This gives rise to the pole problem mentioned above.

Using locally supported B-splines, we construct a representation of the geopotential on each shell, which interpolates the geopotential at each point of the grid.

Let r_i be a fixed distance from the center of the Earth which corresponds to the location of one of the concentric shells. The function S_i that describes the geopotential on the shell is defined by

$$S_i(x, y) = V(r_i, \phi, \theta). \quad (3)$$

The left-hand side has the form

$$S_i(x, y) = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} s_{i,j,k} \beta(Nx - j) \beta(Ny - k), \quad (4)$$

where β denotes the central B-spline of sufficiently high order. Grid spacing on the square is N^{-1} , which is also the length of the largest subinterval on which the spline is a polynomial. Function S_i is an approximation of the geopotential V , hence the two sides of Equation (3) agree to within prescribed accuracy. Choice of the integer N and the order of the B-splines depends upon accuracy and memory requirements. Both are adjustable parameters in the model.

The cost of evaluating the B-spline series (4) is roughly equal to the square of the order of the B-spline. If the order of B-splines is M , then roughly M^2 multiplications are required to evaluate the right-hand side of (4). Note that this cost is independent of the parameter N , which instead governs memory storage requirements. It is obvious from (4) that N^2 coefficients must be stored in computer memory for each shell. To achieve the prescribed accuracy, the choice of the values of N and M are inversely proportional. The larger N is, the smaller M is, and vice versa. Thus, the choice of parameters is a trade-off between computational speed and memory storage requirements.

2.1.1. Spacing between the concentric shells

Examination of Equation (1) reveals that, as r increases, the higher-order terms in the series become insignificant. This means that, for large r , the field can be approximated by low order polynomials and, thus, shells that represent the geopotential far above the Earth need not be so closely spaced as shells near the surface. Consequently, in our models, we increase the distance between consecutive shells as r is increased.

We introduce the dimensionless variable z , defined by

$$z = 1 - \frac{R}{r}. \quad (5)$$

Thus, $z = 0$ corresponds to $r = R$, the surface of Earth, and $z = 1$ corresponds to $r \rightarrow \infty$. The locations of the shells are thus determined by a partition $0 \leq z_0 < z_1 < \dots < z_K \leq 1$ of the interval $[0, 1]$. A typical choice is $z_k = (k/K)^2$, for $k = 0, 1, \dots, K$.

More generally, we may choose any monotone-increasing function f , such that $f(0) = 0$, $f(1) = 1$, and $f(z) < z$ for $0 < z < 1$, then define $z_k = f(k/K)$. Here we do not address the question of the optimal choice for f .

2.1.2. Interpolating between the concentric shells

Let us describe how we use (4) to obtain the values of the geopotential and its derivatives at locations between the tabulated shells. Suppose we want to evaluate at a point (r', ϕ', θ') above the Earth's surface, such that $r_i < r' < r_{i+1}$, where r_i and r_{i+1} correspond to tabulated shells. If we were to use a linear interpolation, we compute the value of the potential on the shells at r_i and r_{i+1} at the appropriate locations (ϕ', θ') , then construct the linear interpolating polynomial that joins these two points. This polynomial in r is then evaluated at $r = r'$.

In practice we use polynomials of higher degree, which requires evaluation on several shells. For example, to use a cubic interpolating polynomial, we evaluate on shells at $r_{i-1}, r_i, r_{i+1}, r_{i+2}$ at the location (ϕ', θ') , and construct the cubic polynomial that passes through these points.

The total cost of evaluation per point for the geopotential spline model is thus $C \cdot P \cdot M^2$, where C is a constant, M is the order of the B-spline and P is the order of the polynomial used to interpolate between shells. Note that P is also an adjustable parameter in the model, and its value can be made smaller to increase the computational speed, if we are willing to increase the number of shells which are tabulated and stored in the computer memory.

Current implementations typically use $M = 12$ and $P = 6$, which corresponds to B-splines of degree 11 and interpolating polynomial of degree 5.

2.1.3. Computing the B-spline expansion coefficients

Let us outline the steps in computing the coefficients for the B-spline expansion (4). For fast and accurate computation of the coefficients it is convenient to extend the potential function V so that it is 2π -periodic in both ϕ and θ . Since V is already periodic in the geocentric longitude ϕ , we only need to extend the geocentric latitude θ from $[0, \pi]$ to $[0, 2\pi]$ in a consistent manner. We define the periodic extension V_p by

$$V_p(r, \phi, \theta) = \begin{cases} V(r, \phi, \theta), & \text{if } 0 \leq \theta < \pi, \\ V(r, \phi + \pi, 2\pi - \theta), & \text{if } \pi \leq \theta < 2\pi. \end{cases}$$

Next we form the expansion

$$S_i(x, y) = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} A_{i,j,k} L(Nx - j)L(Ny - k), \quad (6)$$

with coefficients defined by

$$A_{i,j,k} = V_p(r_i, \phi_j, \theta_k), \quad \text{where } \phi_j = \frac{2\pi j}{N} \quad \text{and} \quad \theta_k = \frac{2\pi k}{N}.$$

The function $L(x)$ is the interpolating spline, which satisfies $L(n) = 1$ if $n = 0$ and $L(n) = 0$ if $n \neq 0$ for integer n . This property implies that the expansion (6) interpolates V_p at the points (r_i, ϕ_j, θ_k) .

Changing the basis from interpolating splines in (6) to B-splines in (4) is fairly straightforward since the interpolating spline is itself a linear combination of B-splines so that the left-hand sides of (4) and (6) are identical. Note that, since the function $S_i(x, y)$ in (4) and (6) is 1-periodic in both x and y , the change of basis can be carried out in the Fourier domain. We apply an FFT to the $N \times N$ matrix of coefficients $\{A_{i,j,k}\}$, multiply each element of the result by a factor, then apply the inverse FFT to obtain the B-spline coefficients $\{s_{i,j,k}\}$ in (4).

Once the coefficients $\{s_{i,j,k}\}$ have been obtained, it is not necessary to store all of them, since the doubly-periodic representation covers the sphere twice. Instead of N^2 , we store only $(N^2/2) + N(M - 1)$ coefficients per shell, where M is the order of the B-spline.

2.1.4. Multiresolution decomposition

Once the representation (4) for the geopotential has been obtained, it is easy to build its multiresolution version. Using the multiresolution decomposition, we are able to remove the high frequency content from our B-spline representation of the geopotential (or its derivatives), and obtain a smoothed, or averaged, version which is supported on a coarser grid.

In order to describe this process, we assume that $N = 2^n$ for some positive integer n , and denote the coefficients in (4) by $\{s_{i,j,k}^n\}$. From the coefficients $\{s_{i,j,k}^n\}$, by applying one step of multiresolution decomposition, we compute the coefficients $\{s_{i,j,k}^{n-1}\}$ on the next coarser scale ($n - 1$), that is, the scale corresponding to twice the step-size of the scale n . From N^2 coefficients per shell on the scale n we obtain $(N/2)^2$ coefficients per shell on the scale $(n - 1)$. The procedure is carried out in the Fourier domain, and consists of the following steps:

- Apply a two-dimensional FFT to the coefficient matrix $\{s_{i,j,k}^n\}$ to obtain the matrix $\{\hat{s}_{i,j,k}^n\}$.
- Apply a one-dimensional decomposition in each index (first on rows, then on columns) of the matrix $\{\hat{s}_{i,j,k}^n\}$. The one-dimensional transform is defined by

$$\hat{s}_m^{n-1} = \frac{1}{2} \left[\tilde{m}_0 \left(\frac{2\pi m}{N} \right) \hat{s}_m^n + \tilde{m}_0 \left(\frac{2\pi m}{N} + \pi \right) \hat{s}_{m+N/2}^n \right]$$

for $m = 0, \dots, N/2 - 1$. The function \tilde{m}_0 is defined by

$$\tilde{m}_0(\xi) = \left(\cos \frac{\xi}{2} \right)^M \frac{a(\xi)}{a(2\xi)},$$

where M is the order of the B-spline $\beta(x)$, and where $a(\xi)$ is defined by

$$a(\xi) = \sum_{1-M/2}^{M/2-1} \beta(k) e^{ik\xi}.$$

- Apply the inverse two-dimensional FFT to the matrix $\{\hat{S}_{i,j,k}^{n-1}\}$ to obtain the coefficient matrix $\{S_{i,j,k}^{n-1}\}$ on the coarser scale.

We note that multiresolution decomposition can be used in conjunction with any of the models described here, not just the doubly-periodic model.

The decomposition algorithm provides a simple and robust method for filtering the high frequency content of the model – high and low frequency can be extracted and represented separately. For example, Figure 1 in Appendix A illustrates how multiresolution decomposition is used to represent the low frequency portion of the field.

Such decomposition can be useful for the purpose of localizing the high frequency contribution of the gravity field. The need for such decomposition is apparent, see for example [2] or [3]. In [3] a deterministic modification of Stokes kernel is constructed for this purpose, with the goal of applying the results to computation of a gravimetric geoid. We feel that the models presented here, especially when coupled with the decomposition algorithm, might be useful in addressing such problems.

2.2. MULTI-WAVELET CUBE MODEL

In this model, the surface of the sphere is mapped to the surface of a cube. Thus, the concentric shells form a sequence of nested cubes. A point on the surface of a sphere is mapped to a point on the reference cube (which has faces perpendicular to the coordinate axes and at a distance of one unit from the origin) using the following simple algorithm:

- Input coordinates (r, ϕ, θ) on the spherical surface of radius r .
- Compute $(x = r \sin \theta \cos \phi, y = r \sin \theta \sin \phi, z = r \cos \theta)$.
- Find $d = \max\{|x|, |y|, |z|\}$.
- Coordinates on the reference cube are $(\xi, \eta, \zeta) = (x/d, y/d, z/d)$.

Geometrically, we can think of a ray that emanates from the origin and intersects the sphere and the reference cube each in a single point. These two points are then mapped one to the other. (Mappings between cubic and spherical surfaces have been considered by other authors as well, see e.g. [4].)

We next place an equispaced rectangular grid on each face of the cube, which can then be mapped backwards onto the sphere. We note that distortion of the grid on the spherical surface caused by the curvature of the sphere is limited, and does not cause any problem near the poles.

The rectangular grid partitions each face of the cube into a number of square subdivisions, and we build a wavelet representation of the geopotential on each subdivision. Currently, the basis functions we use are the multiwavelets (see [5] and [6]), chosen because they form an orthonormal basis on a square subdivision without overlapping into adjacent subdivisions.

The scaling functions in a multiwavelet basis are orthogonal polynomials, for example, Legendre polynomials. Formulas for interpolating the geopotential at the Gaussian nodes within the subinterval are easily obtained using the Gaussian integration rules.

To interpolate the field between shells, we can use the polynomial interpolation scheme as described in Section 2.1.2. Alternatively, we can expand in terms of multiwavelets in height as well.

This model seems best suited for the estimation problem, because the expansions on individual subdivisions of space are independent of the expansions in other regions. This independence provides greater flexibility for using different resolutions in different regions.

2.3. SPLINE CUBE MODEL

Concerning performance of the two models described thus far, two observations have been made. Namely, (1) the multiwavelet cube model has a more efficient memory access than the doubly-periodic spline model, resulting in better speed for evaluation, and (2) the doubly-periodic spline model requires substantially fewer coefficients to achieve the same accuracy as the multiwavelet cube model. The model described in this section has been successful in combining the best features of both.

The importance of efficient memory organization is the following. To evaluate our models it is necessary to interpolate between the tabulated shells, and to do this we must evaluate the two-dimensional expansions at a given point on several consecutive shells. However, the computation involves only a few coefficients on each shell, and speed is gained in memory access if these can be stored closer together.

This can be done by subdividing the surface into a number of ‘panels’ which, taken together, cover the spherical surface. Panels for the same angles but different heights are stored contiguously in memory, which provides faster memory access.

2.3.1. *Partition of the spherical surface*

In this model we subdivide each spherical shell into six panels which may be regarded as the six faces of a cube. This is done in such a way that the grid spacing for the polar regions is the same as that for the equatorial regions, and no excessive distortion of the grid on the spherical surface occurs.

To obtain six square panels, we subdivide the surface of a sphere as indicated in Table I. Panels are numbered 1–6, and the angle ranges of spherical surface co-

TABLE I
Subdivision of the spherical surface to obtain six panels

Panel	Angle ranges	x -coordinate	y -coordinate
1	$-\pi \leq \phi < -\pi/2, \quad -\pi/4 \leq \theta \leq \pi/4$	$\alpha \phi + 3$	$\alpha \theta$
2	$-\pi/2 \leq \phi < 0, \quad -\pi/4 \leq \theta \leq \pi/4$	$\alpha \phi + 1$	$\alpha \theta$
3	$0 \leq \phi < \pi/2, \quad -\pi/4 \leq \theta \leq \pi/4$	$\alpha \phi - 1$	$\alpha \theta$
4	$\pi/2 \leq \phi < \pi, \quad -\pi/4 \leq \theta \leq \pi/4$	$\alpha \phi - 3$	$\alpha \theta$
5	$ \gamma \leq 1, \omega \geq 1, \theta > 0$	$\alpha \tan^{-1}(\omega) \pm 2$	$-\alpha \sin^{-1}(\gamma)$
6	$ \gamma \leq 1, \omega \geq 1, \theta < 0$	$\alpha \tan^{-1}(\omega) \pm 2$	$-\alpha \sin^{-1}(\gamma)$

ordinates (ϕ, θ) for each panel appear in the second column. Canonical coordinates on the face of each panel are (x, y) , where $-1 \leq x, y \leq 1$, and these are obtained as indicated in columns 3 and 4 of the table with $\alpha = 4/\pi$. Panels designated 5 and 6 contain the north and south poles, respectively, and we define $\omega = \tan \theta / \cos \phi$ and $\gamma = \cos \theta \sin \phi$ in rows 5 and 6 of the table. For the x -coordinate in rows 5 and 6, we use the minus sign if $\omega > 0$ and the plus sign if $\omega < 0$.

We note that the B-spline expansion for each panel overlaps its immediate neighbors. Thus, to use this model for the estimation problem, we would need to add a certain number of equations to ensure that the representation near the boundaries of each panel is consistent with that of its neighbors.

2.3.2. Computing the B-spline expansion coefficients

Coefficients for this model are computed in the same manner as described in Section 2.3.2. After computing B-spline expansion coefficients to cover a doubly-periodic extension of the sphere, coefficients to cover each of the first four panels are extracted and stored. The sphere is then rotated, another doubly-periodic expansion computed, then coefficients are extracted to cover the two remaining panels. The rotation is necessary to avoid excessive distortion of the grid near the poles (mentioned in §2.1).

We note that it is inconvenient to store the gravity force field for the polar regions in spherical coordinates. In this coordinate system, the representation of a vector at the poles is not unique, since the longitudinal angle ϕ is arbitrary at the poles.

2.4. PERFORMANCE RESULTS

Here we present performance results obtained with the local model described in Section 2.3. Table II shows memory requirements for the spline cube model. In the first column, ‘order of model’ refers to the degree and order of the underlying spherical harmonic model, and in the second column we list the size in megabytes for a spline cube model which approximates it. As expected, local models require

TABLE II
Size of local models

Order of model	Size in Mbytes
18	10.29
41	27.86
70	70.28

TABLE III
Comparison of evaluation times

Order of model	Time for sph. harm	Time for local model	Speed-up factor	Largest l_2 error
18	3.2745	0.8988	3.6430	0.12e-12
41	16.427	0.8410	19.532	0.16e-11
70	47.547	0.8471	56.130	0.13e-10

substantially more memory than spherical harmonic models. However, modern computers have sufficient RAM to handle the increased size easily.

Tests for timing and accuracy reported in Table III were done by comparing performance of the local model directly to the WGS84 spherical harmonic model [7]. The test consists of computing gravity vectors for 10,000 randomly generated points, then comparing execution times and accuracy. Accuracy is measured by computing the l_2 -norm of the difference between the gravity vectors produced by the two different models, and reporting the largest error. Execution times are in seconds, and the speed-up factor is obtained by dividing the execution time for the spherical harmonic model by the execution time for the local model.

Observe that, while evaluation time for the spherical harmonic model increases quadratically with the order of the model, evaluation time for the local model is constant.

3. Estimation of the Geopotential

The estimation problem may be formulated as an L_2 -minimization problem as follows: given a set of measurements $\{u_p\}$, corresponding to a set of positions $\{\mathbf{r}_p\}$ (where $\mathbf{r}_p = (r_p, \phi_p, \theta_p)$ in spherical coordinates), we want to construct a function u that satisfies Laplace's equation in addition to interpolating the data,

$$u_p = u(\mathbf{r}_p), \quad \text{for each } p. \quad (7)$$

We refer to this problem as the harmonic interpolation problem. The condition (7) is usually replaced by a minimization criterion, for example minimize the error in

L^2 -norm. In practice, the problem of finding u is often replaced by the problem of finding the corrections for u .

The actual estimation problem is more complicated since the input data often consists of measurements made from satellites, which contain errors due to various sources and, thus, at best can only be expected to be close to values of the actual geopotential. This situation brings in additional complications which we would like to ignore for now and, therefore, we confine our discussion to the problem of harmonic interpolation in order to illustrate several important points.

A traditional approach for solving the problem is to represent u with spherical harmonics and estimate the coefficients. We could thus make use of expression (1), which contains spherical harmonics up through degree and order N . This representation obviously solves the Laplace equation, and the problem is one of computing the coefficients to solve the interpolation problem (7). Difficulties occur if the measurements contain spatial frequencies higher than can be accounted for by the expression (1), in particular local changes of high frequency. Condition of the resulting system is a serious problem as well.

We propose another approach, namely, to build a multiresolution representation of the harmonic interpolant u . Let us express the solution of Laplace's equation as

$$u(r, \phi, \theta) = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} K(r, \theta, \theta', \phi - \phi') u(a, \phi', \theta') d\theta' d\phi'. \quad (8)$$

for $r > a$, where

$$K(r, \theta, \theta', \phi - \phi') = \frac{1}{4\pi} \frac{a(r^2 - a^2) \cos \theta'}{(r^2 - 2ar \cos \gamma + a^2)^{3/2}},$$

is the kernel that appears in (8), and

$$\cos \gamma = \sin \theta \sin \theta' + \cos \theta \cos \theta' \cos(\phi - \phi').$$

We now seek a wavelet expansion for the function u in (8) of the form

$$u(a, \phi', \theta') = \sum_{k,l} s_{k,l}^n \varphi_k^n(\theta') \varphi_l^n(\phi') + \sum_{j=1}^n \sum_{k,l} \{d_{k,l}^{1,j} \psi_k^j(\theta') \psi_l^j(\phi') + d_{k,l}^{2,j} \varphi_k^j(\theta') \psi_l^j(\phi') + d_{k,l}^{3,j} \psi_k^j(\theta') \varphi_l^j(\phi')\} \quad (9)$$

which is an expansion of u over $j = 1, \dots, n$ scales, where n is an adjustable parameter. Substituting (9) into (8) and applying the condition (7), we obtain

$$u(\mathbf{r}_p) = \sum_{k,l} s_{k,l}^n t_{k,l}^{n,p} + \sum_{j=1}^n \sum_{k,l} \{d_{k,l}^{1,j} \alpha_{k,l}^{j,p} + d_{k,l}^{2,j} \beta_{k,l}^{j,p} + d_{k,l}^{3,j} \gamma_{k,l}^{j,p}\} \quad (10)$$

where, for example,

$$\alpha_{k,l}^{j,p} = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} K(r_p, \theta_p, \theta', \phi_p - \phi') \psi_k^j(\theta') \psi_l^j(\phi') d\theta' d\phi'.$$

The definitions of $t_{k,l}^{n,p}$, $\beta_{k,l}^{j,p}$ and $\gamma_{k,l}^{j,p}$ are similar, and are obtained by substituting (9) into (8) and comparing the result to (10).

Collecting all such equations, for $p = 1, \dots, N_p$, where N_p is the number of observations, we have a linear system of the form

$$Tx = u, \quad (11)$$

where x is a vector representing all coefficients $d_{k,l}^{1,j}$, $d_{k,l}^{2,j}$, $d_{k,l}^{3,j}$, and $s_{k,l}^n$, and T is a matrix containing the elements $\alpha_{k,l}^{j,p}$, $\beta_{k,l}^{j,p}$, $\gamma_{k,l}^{j,p}$, and $t_{k,l}^{n,p}$. The right-hand side $u = \{u_1, \dots, u_{N_p}\}$ is a vector of observations.

The matrix T of the linear system (11) is sparse and typically very ill-conditioned. We are now faced with solving an ill-conditioned least squares problem. The standard approach to such problems is to use the rank-revealing QR algorithm which yields a solution of minimum l^2 -norm. However, existence of a large (numerical) null space often renders minimum l^2 -norm solution unsatisfactory.

We have developed an algorithm for solving ill-conditioned least squares problems [8] which produces what we call the ‘minimum detail’ solution. This algorithm is a QR-type algorithm where the unknowns are organized by their spectral content. In a wavelet system of coordinates it simply means that we work one scale at a time.

The algorithm makes two passes. First, it restructures the matrix by selecting the unknowns which can be resolved given a particular set of observations. This pass proceeds from the coarsest scale to the finest. Second, it computes the values of the selected unknowns, starting from the finest scale and proceeding to the coarsest.

The algorithm depends critically on the fact that the basis functions are localized both in space and wave number domains. Such an algorithm would be impossible if, for example, the basis functions were spherical harmonics.

We provide this discussion partly to emphasize the fact that it is not sufficient to consider only representation of the gravity field. One must also consider all algorithms that will be used in conjunction with the representation, as well as for the estimation, since these aspects represent the two sides of the same problem.

4. Conclusions and Future Work

At this point we fully understand the interplay between the choices available for representing functions (e.g. the geopotential) and for constructing estimation algorithms. We are currently involved in constructing optimal representations for the geopotential, as well as implementing our approach to estimation using actual data, after which we will compare the resulting multiresolution models with current spherical harmonic models. The results will be reported elsewhere.

Acknowledgements

We gratefully acknowledge the invaluable contributions made to this work by Dr. Joseph Liu, who was instrumental in putting this program in motion, and also by Mark Storz, both of USAFSC. G. Beylkin; This research was partially supported by DARPA/AFOSR Grant DOD F49620-97-1-0017, DARPA/AFOSR Grant DOD F49620-98-1-0491 and DARPA/NASA Grant S43 5-28646. R. Cramer; This research was partially supported by DARPA/AFOSR Grant DOD F49620-97-1-0017 and DARPA/AFOSR Grant DOD F49620-98-1-0491.

Appendix A. Illustration of the Multiresolution Decomposition

In Figure 1 we illustrate the multiresolution decomposition of the geopotential (see §2.1.4) within the B-spline model. The underlying spherical harmonic model is EGM96 (360×360 order and degree model). The plot in the upper left corner is a high resolution display of a selected region of Earth's geopotential. This particular region was chosen because of the wide bandwidth of its features. Going from left to right and top to bottom, the sequence shows the successive coarsening using the multiresolution decomposition.

Appendix B. An Analogue of Gibbs' Phenomenon Due to Abrupt Truncation of the Spherical Harmonic Expansion

Let us present a comparison that illuminates difficulties inherent in current spherical harmonic models. We consider the difference between two models, WGS84-41 and WGS84-70 [7]. As the coefficients of WGS84-41 are identical with the coefficients of WGS84-70 up through order 41, we can view WGS84-41 as a truncated spherical harmonic expansion of the full WGS84-70 model. In the same way, WGS84-70 can be viewed as a truncation of even higher order models.

To illustrate our comparison, we plot the function V for both models WGS84-70 and WGS84-41 on the surface of the Earth around its equator, taking $r = R$, $\theta = \pi/2$, and $0 \leq \phi \leq 2\pi$ in (1). This particular choice provides a good representative of the typical behavior.

Figure 2 shows the WGS84-70 model together with the WGS84-41 model and the difference between them. The relative difference is plotted in Figure 3.

The question that one might ask is what part of this difference is due to the analogue of Gibbs' phenomenon. (Gibbs' phenomenon is an oscillatory error which occurs if a Fourier series with slowly decaying coefficients is truncated abruptly. Here we refer to the abrupt truncation of the spherical harmonic expansion.) The difference between the two graphs is most likely due to the abrupt truncation. Spherical harmonics, being global, oscillatory functions, depend on cancellation

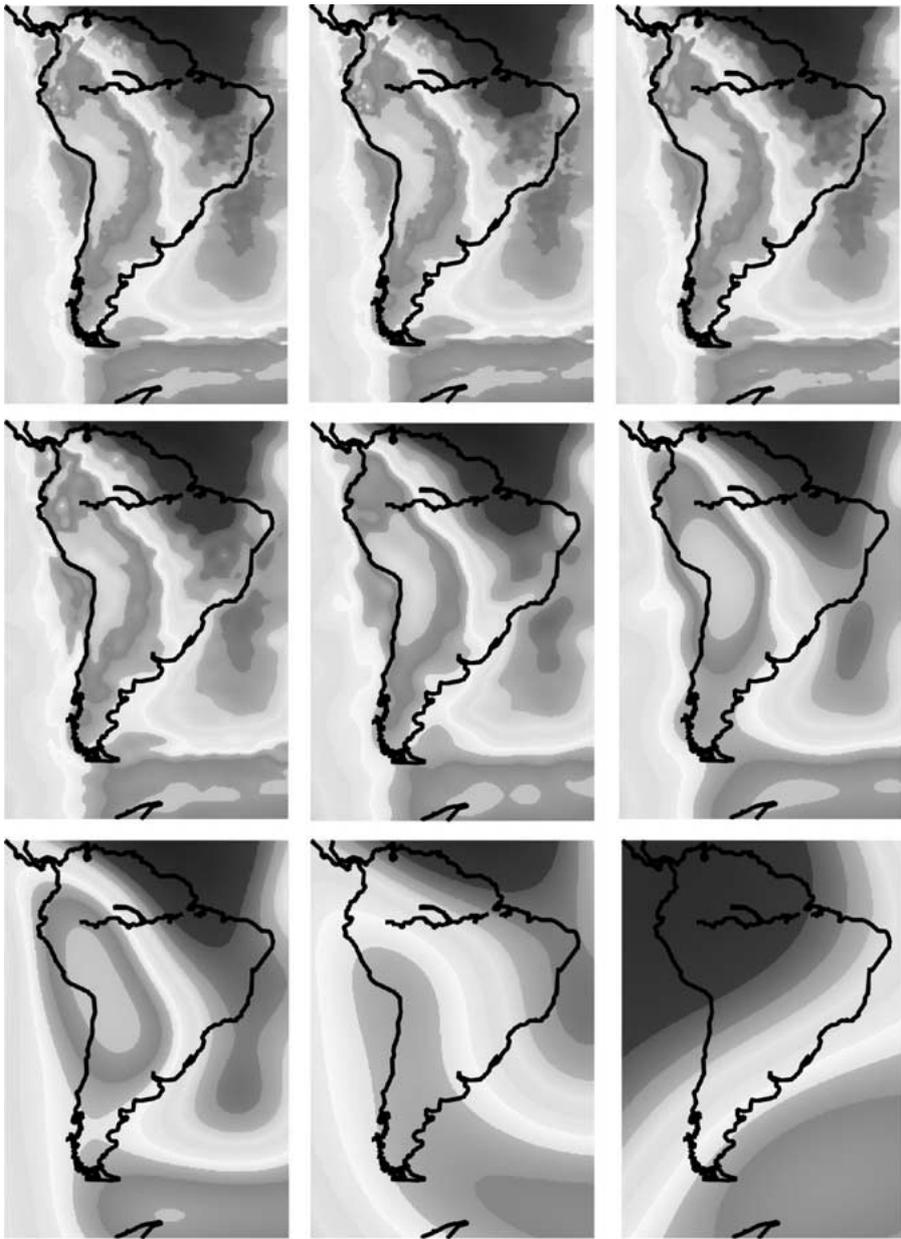


Figure 1. Multiresolution decomposition of a selected region using the B-spline representation of the spherical harmonic geopotential model EGM96.

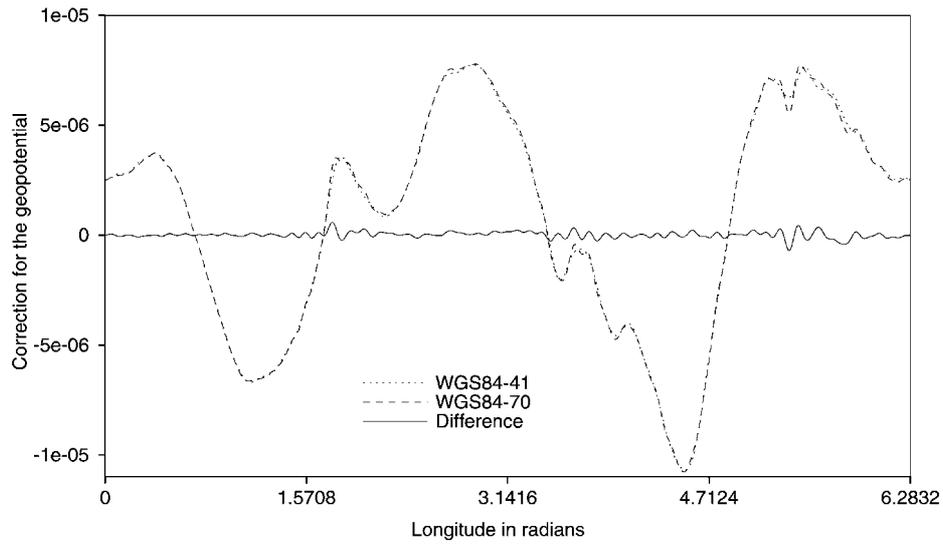


Figure 2. Geopotentials on the equator computed using the WGS84-70 model and the WGS84-41 model displayed together with the difference between them. In this and the following figures we plot only the ‘correction’ terms of V , that is, only those for $n \geq 3$.

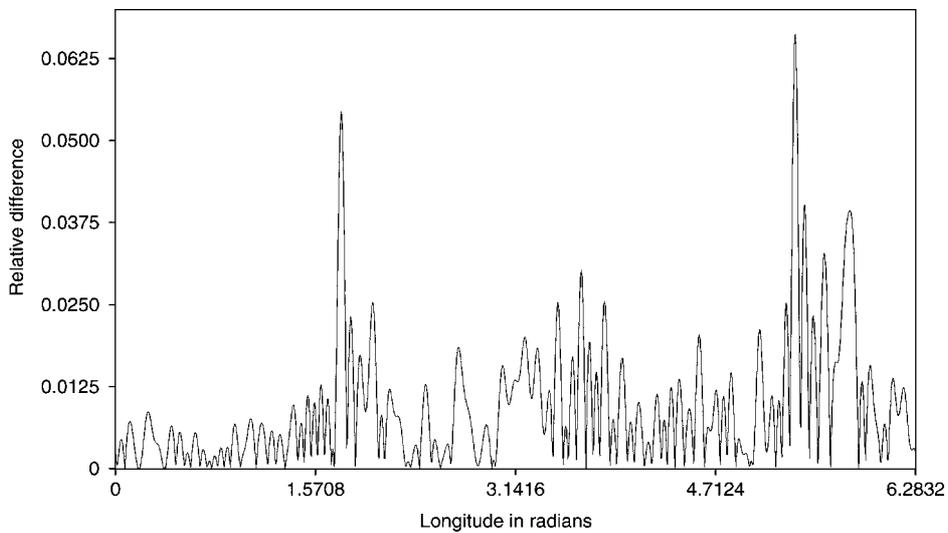


Figure 3. Relative difference between WGS84-70 and WGS84-41. We attribute the peaks, which are above 5% locally, to the Gibbs type phenomenon near the locations of the rapid change of the geopotential.

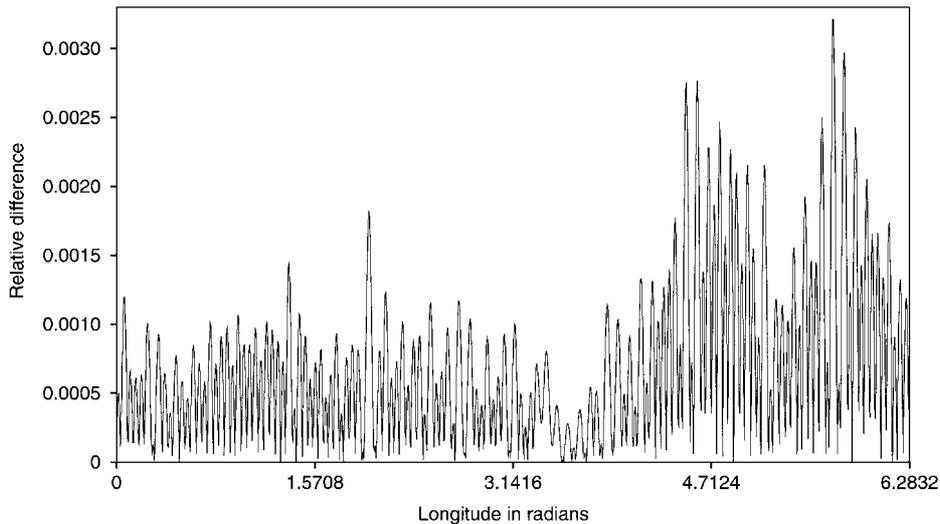


Figure 4. Relative difference between WGS84-70 and LWGS84-70R. Note that the difference is at most about 0.3%.

(destructive interference) to achieve the approximation. If high frequencies are removed by truncation then Gibbs' phenomenon occurs (in full analogy with Fourier series).

Now let us consider the B-spline approximation of the surface potential $V(r = R$ in (1)), obtained by interpolating V on an equispaced, two-dimensional grid. The size of the grid is roughly $4N \times 4N$, where N is the order and degree of the model, which is sufficient for a highly accurate representation of the potential function V (or its derivatives). For example, we use a 160×160 grid to represent the WGS84-41 model with accuracy $\approx 10^{-11}$ and represent the WGS84-70 model with the same accuracy on a 280×280 grid. Let us refer to the former B-spline representation as LWGS84-41 (local WGS84-41) and to the latter as LWGS84-70 (local WGS84-70). Model LWGS84-70 is, for practical purposes, indistinguishable from model WGS84-70, and LWGS84-41 is likewise indistinguishable from WGS84-41.

Performing one step of multiresolution decomposition (see §2.1.4) on the representation LWGS84-70, we obtain the reduced model, which we refer to as LWGS84-70R. This reduced model is supported on a 140×140 grid and thus is somewhat smaller in size than the LWGS84-41 model. The relative difference between LWGS84-70R and WGS84-70 is shown in Figure 4.

Note that the relative difference in Figure 4 is roughly 20 times smaller than that in Figure 3. This indicates a low information content of roughly half of the spherical harmonic coefficients in WGS84-70 model, corresponding to the high spatial frequencies. The B-spline model LWGS84-70R preserves the essential features of WGS84-70 much more faithfully than does WGS84-41, and yet it is practically devoid of high frequency content, and is essentially equivalent to WGS84-41 in its resolution. The difference is that LWGS84-70R was not obtained by an abrupt

truncation, but rather by a smoothing process which, unlike truncation, avoids the introduction of strong oscillatory artifacts (Gibbs' phenomenon).

In summary, we have seen that there exist significant differences between WGS84-41 and WGS84-70 and that these differences are indicative of errors introduced by abrupt truncation. It then seems likely that WGS84-70 would contain the same kinds of errors if it were obtained as a truncation of a higher-order model. The higher order and degree models, such as EGM96, do not avoid this problem, since the abrupt truncation is still present. Using a smoothing process based on multiresolution analysis (e.g. B-splines) for generation of lower resolution models, rather than a truncation process, yields significantly better results.

References

1. Freedden, W. and Windheuser, U.: 1997, 'Combined spherical harmonic and wavelet expansion – a future concept in earth's gravitational determination', *Appl. Comput. Harm. Anal.* **4**, 1–37.
2. Vatrtr, V.: 1999, 'Truncation error due to geopotential model EGM96', *Studia Geophysica et Geodaetica* **43**, 223–227.
3. Featherstone, W., Evans, J. and Olliver, J.: 1998, 'A Meissl-modified Vaníček and Kleusberg kernel to reduce the truncation error in gravimetric geoid computations', *J. Geod.* **72**, 154–160.
4. Ronchi, C., Iacono, R. and Paolucci, P.: 1995, 'The 'cubed sphere': a new method for the solution of partial differential equations in spherical geometry', *J. Comput. Phy.* **124**, 93–114.
5. Alpert, B.: 1993, 'A class of bases in L^2 for the sparse representation of integral operators', *SIAM J. Math. Anal.* **24**(1), 246–262.
6. Alpert, B., Beylkin, G., Coifman, R. and Rokhlin, V.: 1993, 'Wavelet-like bases for the fast solution of second-kind integral equations', *SIAM J. Sci. Comput.* **14**(1), 159–184.
7. Dept. of Defense World Geodetic System: 1984, 'Defense Mapping Agency Technical Report', 1987. DMA TR 8350.2.
8. Beylkin, G., Gines, D. and Coult, N.: 2001, 'Scattered data interpolation using multi-resolution QR factorization', *APPM Report 475*, Univ. of Colo. at Boulder, July 2001.