



Rigorous Evaluation of Gravity Field Functionals from Satellite-Only Gravitational Models Within Topography

Michael B. Sheng, Petr Vaníček, Robert Kingdon, and Ismael Foroughi

Abstract

Currently, extensive work is being done in the field of geodesy on producing better gravitational models using purely space-based techniques. With the large datasets spanning a long timeframe, thanks to the GOCE and GRACE missions, it is now possible to compute high quality global gravitational models and publish them in a convenient form: spherical harmonics. For regional geoid modeling, this is advantageous as these models provide a useful reference which can be improved with terrestrial observations. In order for these global models to be usable below the topographical surface, certain considerations are required; topographical masses cause the function that describes the gravity potential to be non-harmonic in the space between the topographical surface and the geoid. This violates the mathematical assumptions behind solid spherical harmonics.

This paper aims to look at the error caused by evaluating solid spherical harmonics when topography is present. It thus provides a more rigorous methodology than the commonly used approach of computing the quasigeoid and then applying an approximate correction term for the geoid-quasigeoid separation. It is therefore well-suited for the Stokes-Helmert approach to high-precision regional geoid computation. Comparisons between the more rigorous methodology and the generally used algorithm are made in order to study the error that is committed. With a range of 23.6 cm and a standard deviation of 0.8 cm, this is a non-trivial error if the ultimate goal is to compute a regional geoid with an accuracy of better than 1 cm.

Keywords

Geoid • Gravitational model • Harmonic synthesis • Stokes-Helmert

1 Introduction

This paper describes a methodology for rigorously evaluating functionals from global gravitational models over land while accounting for topographical density and is designed specifically for use in the well documented SHGeo high-precision regional geoid computational scheme and

associated software package (Vaníček and Martinec 1994). This paper also investigates the error that is committed by neglecting the requirement of harmonicity of the gravity potential within topography when one wishes to work with the global model below the topographic surface, or in our case, on the geoid. A similar methodology is presented in Wang et al. (2012) although the numerical results are not rigorously evaluated as they opt for an alternative approach more tailored for use with ultra-high degree reference geopotential models.

Initially, several satellite-only global models were considered in the computations: DIR-R5, GGM05S, and GOCO05S. These models were considered due to the

M.B. Sheng (✉) • P. Vaníček • R. Kingdon • I. Foroughi
Department of Geodesy and Geomatics Engineering, University
of New Brunswick, Fredericton, Canada
e-mail: michaelsheng1@gmail.com

differing data sources used in their formulation but only the results from DIR-R5 will be discussed in this paper as the results confirmed that the error committed by assuming that a non-harmonic function will behave as if it were harmonic depends on topography and is independent of the potential model. Due to the limited useful spectral content of the satellite-only models, the studied model is limited to degree and order 160 which is sufficient for providing a satellite-only reference field for computing regional high-precision geoids (Abdalla et al. 2012; Foroughi et al. 2016).

The model spherical harmonic coefficients are generally referred to the sphere with radius equal to the major-semi axis of the geocentric ellipsoid of revolution (a), such as the GRS80 reference ellipsoid. For various applications, we need to know the potential field described by the series on the geoid (i.e. inside of topography) and not on this sphere. This complicates matters as the field cannot be continued downward to the geoid because it does not behave in a harmonic manner within the topographical masses located between the Brillouin sphere and the geoid (Heiskanen and Moritz 1967). In practice, various functionals of the gravity field are often evaluated within the topography directly from the harmonic coefficients (Barthelmes 2009) which violates the fundamental assumptions of harmonic series (that the field they represent is harmonic in the space where the field is considered).

2 Methodology

The first fact to be considered is that topography protrudes in places above the sphere of radius a thus making the gravity potential in the space above the sphere inherently non-harmonic. In order to work around this problem, the spectral coefficients should be scaled to the Brillouin sphere of somewhat larger radius R_B . Then, by making use of Helmert's second condensation method, it is possible to more or less rigorously (the topographical density is known only to a certain degree) account for the effects of topographical masses and make the gravity potential everywhere above the geoid more or less harmonic (Helmert 1884; Vaníček and Martinec 1994). It should be noted that there are an infinite number of ways to compensate for the topographic masses (Helmert's condensations, Pratt-Hayford, Airy-Heiskanen, etc. (Heiskanen and Moritz 1967)), and that Helmert's second condensation was chosen here due to the direct applicability to the SHGeo computational scheme. The difference between the topographical and the condensed topographical potentials – the so-called direct topographical effect (DTE, δV) – is expressed in spectral form derived by Vaníček et al. (1995). Its subtraction transforms the gravity potential from the real space to the Helmert space in which the gravity

potential behaves (more or less) harmonically everywhere above the geoid. We have:

$$\forall n \geq 0, m = 0, 1, \dots, n : W_{nm}^H = W_{nm} - \delta V_{nm} \quad (1)$$

$$\forall n \geq 0, m = 0, 1, \dots, n : \delta V_{nm} \cong 2\pi G \rho_0 \frac{n}{2n+1} (H^2)_{nm} \quad (2)$$

where G is the gravitational constant, ρ_0 is the global mean topographic density of $2,670 \text{ kg m}^{-3}$, and $(H^2)_{nm}$ are the coefficients of topographical height squared. The astute reader will notice the differences between Eq. (2) presented above and Eq. (20) from Vaníček et al. (1995); the difference comes from the choice to conserve the total mass of the Earth as opposed to conserving the centroid when transforming to Helmert's space (Wichiencharoen 1982; Martinec 1993). Once we have obtained Helmert's gravity potential W^H , the normal potential (U) generated by the GRS80 ellipsoid of revolution is subtracted from it to arrive at Helmert's disturbing potential:

$$\forall n \geq 0, m = 0, 1, \dots, n : T_{nm}^H = W_{nm}^H - U_{n0}, \quad (3)$$

where U_{n0} has only even zonal harmonics different from 0.

Because the space through which the potential is to be "downward continued" has been made as free of mass as we know how, the gravitational potential behaves more or less harmonically and it is now possible to express Helmert's disturbing potential on the Helmert co-geoid by evaluating the harmonic coefficients of T^H for $r = r_g$:

$$T^H(r, \varphi, \lambda) = \frac{GM}{r} \sum_{n=0}^{n_{max}} \left(\frac{R}{r}\right)^n \sum_{m=0}^n \times (C_{nm}^H \cos m\lambda + S_{nm}^H \sin m\lambda) P_{nm}(\sin \varphi) \quad (4)$$

where (φ, λ) are the geocentric latitude and longitude, C^H and S^H are the harmonic coefficients of Helmert's disturbing potential (T_{nm}^H) in the geodetic norm (Varshalovich et al. 1988), P_{nm} are the associated Legendre polynomials, R is the radius to which the coefficients are referred (in our case the radius of the Brillouin sphere), and r is the radius of the surface where the disturbing potential T^H is synthesized (in our case, the radius of the geoid, r_g , approximated by r_e with a resulting error of less than 1 cm) (Barthelmes 2009; Vaníček et al. 1995). Let us note that the same expression in spectral form can be used in any space where the gravitational potential behaves harmonically above the surface r defined in Eq. (4) and Helmert's space was chosen for this study simply due to the ease of implementation with the SHGeo software package.

We note that a ‘‘Helmertisation’’ of a global field had already been attempted by Vaníček et al. (1995). In their derivations, Vaníček et al. assumed, incorrectly as it turns out, that the global field would be known at the reference ellipsoid level and be valid at the geoid. As this cannot be the case, their results are not quite correct.

Once ‘‘downward continued’’, Helmert’s disturbing potential is transformed to give us the Helmert co-geoid (the geoid in Helmert’s space) using Bruns’s formula. If needed, the Helmert co-geoid is transformed to the real space by accounting, once again, for the residual topographical potential of the difference between topography and the condensation layer, also known as the primary indirect topographical effect on the geoid (PITE) or $\frac{\delta V}{\gamma_0}$. Thus, geoidal undulation (N) in the real space is obtained as:

$$N(\varphi, \lambda) = \frac{T^H(r_g, \varphi, \lambda)}{\gamma_0(\varphi)} + \frac{\delta V(r_g, \varphi, \lambda)}{\gamma_0(\varphi)} \quad (5)$$

where:

$$\begin{aligned} \delta V(r_g, \varphi, \lambda) = & \frac{GM}{r_g} \sum_{n=0}^{n_{max}} \left(\frac{R}{r_g}\right)^n \sum_{m=0}^n \\ & \times \left(C^{\delta V}_{nm} \cos m\lambda + S^{\delta V}_{nm} \sin m\lambda\right) P_{nm}(\sin \varphi) \end{aligned} \quad (6)$$

3 Numerical Example

In this section several comparisons are made between the methodology described in the previous section and the not so rigorous methodology that directly estimates geoidal undulation without converting topography into a space in which the gravity potential is harmonic. This is done approximately by applying Eq. (116) from Barthelmes (2009) presented as:

$$\begin{aligned} N(\varphi, \lambda) & \cong \frac{GM}{r_e \gamma(r_e, \varphi)} \sum_{n=0}^{n_{max}} \left(\frac{R}{r_e}\right)^n \sum_{m=0}^n \\ & \times \left(C^T_{nm} \cos m\lambda + S^T_{nm} \sin m\lambda\right) P_{nm}(\sin \varphi) - \frac{2\pi G\rho_0}{\gamma(r_e, \varphi)} \\ & \times \left[R \sum_{n=0}^{n_{max}} \sum_{m=0}^n \left(C^H_{nm} \cos m\lambda + S^H_{nm} \sin m\lambda\right) P_{nm}(\sin \varphi) \right]^2 \end{aligned} \quad (7)$$

where C^H_{nm} and S^H_{nm} are the spectral coefficients of topographical heights scaled by R and C^T_{nm} and S^T_{nm} are the spectral coefficients of disturbing potential after applying Eq. (3) in the real space. The first term of this equation is supposed to give the quasigeoid and the second term adds an approximation of the geoid-quasigeoid separation. Apart from the fact that the quasigeoid cannot be properly expressed either by a harmonic series, nor any other series,

there are several mathematical and physical problems with this approach. Firstly, the quasigeoid is a folded surface; therefore, the quasigeoidal height cannot be described as a function of horizontal position. Secondly it seems peculiar to use an approximate formula for the geoid-quasigeoid separation in the computation of an accurate geoid model.

For the comparisons, DIR-R5 was used up to degree and order 160 using the two differing methodologies. This comparison will give us the error committed by overlooking the laws of physics and mathematics. The following assumptions were made for the purpose of these numerical tests:

1. Topographical surface of the Earth (being a discontinuous function) can be represented by an infinite summation of continuous functions in the mean sense.
2. Real density variations within topography can be approximated by mean topographic density; the effect of topodensity inhomogeneity should also be considered if this methodology is to be applied properly but was not available when computations were done for this paper.
3. The radius of the geoid (r_g) can be approximated by the geocentric radius of the reference ellipsoid (r_e) with a sufficient accuracy (Vaníček et al. 1995).

Height squared coefficients from DTM2006.0 (Pavlis et al. 2007) were used in Eq. (2). The differences quoted here are those between the geoidal undulations computed by applying Eqs. (1)–(5) and the geoidal undulations computed by applying Eq. (7) in the real space. Figure 1 shows the differences between the non-rigorous process and our more rigorous process; the statistical summary of the differences between the two approaches are presented in Table 1.

For this numerical example, geoidal undulation was chosen as the functional of the gravity potential that must naturally be computed within topography. It should be noted that similar errors will arise when attempting to evaluate any other functional of the gravity potential within topography.

4 Discussion and Conclusions

One problem people may perceive with the aforementioned methodology is that nominally, the interior gravity potential series (meaning interior to the Brillouin sphere) is divergent because the radial functions R/r are everywhere larger than 1 and the sequence of radial functions $(R/r)^n$ grows beyond all limits with growing n . Here we make the following reasonable assumption: that the series does in fact converge for all r between the Brillouin sphere and the geoid because the Helmert disturbing potential coefficient sequence converges to 0 faster than the sequence $(R/r)^n$ diverges. It is intuitively clear that this condition is always satisfied as gravity in the space between the geoid and the Brillouin sphere has been always observed to be well within finite limits. For

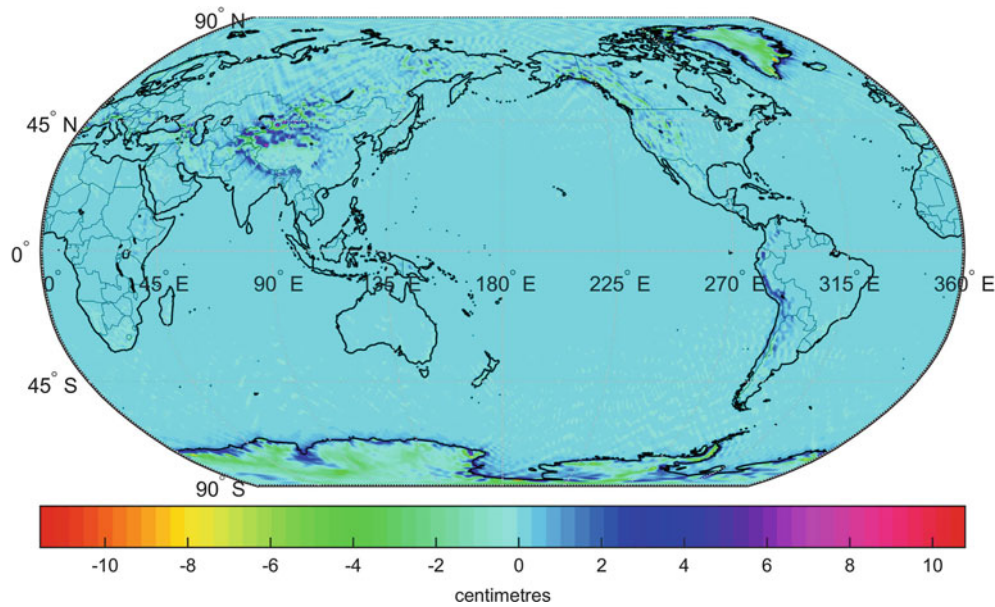


Fig. 1 Differences between approaches – DIR-R5

Table 1 Statistics of the differences between non-rigorous and rigorous methodologies

	Maximum	Minimum	Mean	Standard deviation
DIR-R5	11.3 cm	−12.3 cm	0.00 cm	0.8 cm

the purpose of this project, it is also clear that the series cannot numerically diverge as we truncate the series at a finite degree and order and therefore even the sequence $(R/r)^n$ cannot diverge to infinity.

As can be seen from the values shown in Table 1, the differences in the two methodologies are significant. From Fig. 1 we can see that the differences are directly correlated to topographical heights as should be expected from Eq. (2). The large differences in Antarctica and Greenland are likely due to the assumption that the mean topographical density represents real density. If a topo-density inhomogeneity model were used, as it should have been, the differences between the rigorous and more commonly-used methodologies would have been smaller in these regions but possibly larger in others.

Numerically, it would appear that the error committed by neglecting the requirement of harmonicity of the gravitational potential is not negligible if one aims to compute a regional geoid model with accuracy of 1 cm; errors present in the global model will have a direct effect on the resulting regional model. The range of the differences is 23.6 cm with a mean that tends to zero and a standard deviation of 0.8 cm. It seems only logical that one should prefer to use a methodology that strictly adheres to physical and mathematical principles rather than one which does not.

The next step in testing would be to incorporate the effect of lateral density variations to better capture the effects of real topography. The continuation of this work should also include the assessment of the geoidal undulations over a sample of GPS/levelling benchmarks globally; this would allow us to gain a better sense of the geometrical fit and it would also make the assessment more trustworthy by virtue of using independent data. Finally, the effect on regional geoid modeling of neglecting the requirements of harmonicity of global models should be investigated. Global models are implemented in various stages of the regional formulations; therefore, the propagation of errors might be more complex than initially assumed.

References

- Abdalla A, Fashir H, Ali A, Fairhead D (2012) Validation of recent GOCE/GRACE geopotential models over Khartoum state – Sudan. *J Geodetic Sci* 2(2):88–97
- Barthelmes F (2009) Definition of functionals of the geopotential and their calculation from spherical harmonic models. Scientific Technical Report STR09/02, Potsdam
- Foroughi I, Vaníček P, Novák P, Kingdon R, Sheng M, Santos M (2016) Optimal combination of satellite and terrestrial gravity data for regional geoid determination using Stokes-Helmert method. Paper presented at the IAG Commission 2 meeting, 22 September 2016, Thessaloniki
- Heiskanen WH, Moritz H (1967) *Physical geodesy*. W.H. Freeman and Co., San Francisco
- Helmert FR (1884) *Die mathematischen und physikalischen Theorien der höheren Geodäsie*, vol 2. B.G. Teubner, Leipzig. (Reprinted in 1962 by Minerva GMBH, Frankfurt/Main)
- Martinec Z (1993) Effect of lateral density variations of topographical masses in view of improving geoid model accuracy over Canada.

- DSS contract 23244-2-4356/01-SS, Geodetic Survey of Canada, Ottawa
- Pavlis NK, Factor JK, Holmes SA (2007) Terrain-related gravimetric quantities computed for the next EGM. In: Proceedings of the 1st international symposium of the international gravity field service, vol 18, Harita Dergisi, pp 318–323
- Vaníček P, Martinec Z (1994) The Stokes-Helmert scheme for the evaluation of a precise geoid. *Manuscr Geodaet* 19:119–128
- Vaníček P, Najafi M, Martinec Z, Harrie L, Sjöberg LE (1995) Higher-degree reference field in the generalized Stokes-Helmert scheme for geoid computation. *J Geod* 70:176–182
- Varshalovich DA, Moskalev AN, Khersonskii VK (1988) Quantum theory of angular momentum. World Scientific Publishing Co., Singapore
- Wang YM et al (2012) The US Gravimetric Geoid of 2009 (USGG2009): model development and evaluation. *J Geod* 86:165–180
- Wichiencharoen C (1982) The indirect effects on the computation of geoid undulations. Science Report 336, Department of Geodesy, Ohio State University, Columbus