

Testing Stokes-Helmert geoid model computation on a synthetic gravity field: experiences and shortcomings

P. Vaniček

Department of Geodesy and Geomatics Engineering, University of New Brunswick,
P.O. Box 4400, Fredericton, New Brunswick, E3B 5A3, Canada

R. Kingdon

Department of Geodesy and Geomatics Engineering, University of New Brunswick,
P.O. Box 4400, Fredericton, New Brunswick, E3B 5A3, Canada
(currently at: Fugro Airborne Surveys, 2191 Thurston Drive, Ottawa, Ontario, K1G 6C9, Canada)

M. Kuhn

Western Australian Centre for Geodesy & The Institute for Geoscience Research,
Curtin University of Technology, GPO Box U1987, Perth WA 6845, Australia

A. Ellmann

Faculty of Civil Engineering, Tallinn University of Technology,
Ehitajate tee 5, 19086 Tallinn, Estonia

W. E. Featherstone

Western Australian Centre for Geodesy & The Institute for Geoscience Research,
Curtin University of Technology, GPO Box U1987, Perth WA 6845, Australia

M. C. Santos

Department of Geodesy and Geomatics Engineering, University of New Brunswick,
P.O. Box 4400, Fredericton, New Brunswick, E3B 5A3, Canada

Z. Martinec

Geophysics Section, Dublin Institute for Advanced Studies,
5 Merrion Square, Dublin 2, Ireland

C. Hirt

Western Australian Centre for Geodesy & The Institute for Geoscience Research,
Curtin University of Technology, GPO Box U1987, Perth WA 6845, Australia

D. Avalos-Naranjo

Department of Geodesy and Geomatics Engineering, University of New Brunswick,
P.O. Box 4400, Fredericton, New Brunswick, E3B 5A3, Canada
(currently at: Instituto Nacional de Estadística y Geografía, Av. Héroe de Nacozari Sur 2301, C.P.
20276, Aguascalientes, Mexico)

Abstract

We report on testing the UNB (University of New Brunswick) software suite for accurate regional geoid model determination by use of Stokes-Helmert's method against an Australian synthetic field (ASF) as "ground truth". This testing has taken several years and has led to discoveries of several significant errors (larger than 5mm in the resulting geoid models) both in the UNB software as well as the ASF. It was our hope that, after correcting the errors in UNB software, we would be able to come up with some definite numbers as far as the achievable accuracy for a geoid model computed by the UNB software. Unfortunately, it turned out that the ASF contained errors, some of as yet unknown origin, that will have to be removed before that ultimate goal can be reached. Regardless, the testing has taught us some valuable lessons, which we describe in this paper. As matters stand now, it seems that given errorless gravity data on 1' by 1' grid, a digital elevation model of a reasonable accuracy and no topographical density variations, the Stokes-Helmert approach as realised in the UNB software suite is capable of delivering an accuracy of the geoid model of no constant bias, standard deviation of about 25 mm and a maximum range of about 200 mm. We note that the UNB software suite does not use any corrective measures, such as biases and tilts or surface fitting, so the resulting errors reflect only the errors in modelling the geoid.

Keywords: synthetic gravity models, geoid, topographic effects, Stokes-Helmert method

1. Introduction

Foreword: Readers not already familiar with the Stokes-Helmert technique should read the summary in Ellmann and Vaniček (2007) as a primer; readers not already familiar with the Australian synthetic gravity field should read Baran et al. (2006) as a primer. However, we provide summaries of both in Sections 2 and 3, respectively, of this paper.

Almost exclusively, the precision and accuracy of regional geoid or quasigeoid modelling results have been validated empirically by using co-located points whose heights were determined both by differential levelling and by GPS (or by any other GNSS). However, the heights of these control points have their own errors, which stem from the levelling (e.g., height system used, distortions in the vertical datum, etc.) and GPS measurements (e.g., short occupations, inconsistent reference frames, poor geometry, etc). GPS-levelling control points are generally over the dry land only and usually their distribution is rather inhomogeneous.

One application of a synthetic Earth gravity model (SEGM) is to provide another means of validating geoid determination theories, techniques and computer software. Early SEGMs were vicarious, assuming spherical harmonic models were error-free and able to generate self-consistent gravity and geoid (e.g., Tziavos 1996, Novák et al. 2001; Featherstone 2002). The limitation of these is that they are unable to allow for testing of the treatment of the topography as they lack the sufficiently high frequencies that are caused predominantly by topography. As such, more sophisticated SEGMs that contain topography and its gravitational effects have been produced (e.g., Haagmans 2000, Kuhn

and Featherstone 2005, Baran et al. 2006). Also, several higher-degree dissertations have been produced on the topic (e.g., Pail 2000, Claessens 2003, Ågren 2004). However, literature searches show that none of these SEGMs have gained widespread acceptance; we are unsure as to why. Some other efforts to ‘benchmark’ different geoid computation techniques have been made, notably at the Auvergne test site in France (Valty et al. 2012), but these use observational data so are subject to the limitations outlined above.

Over the past three decades, UNB has been involved in regional gravimetric geoid determination using the Stokes (1849) – Helmert (1884) method (e.g., Vaniček and Sjöberg 1991, Vaniček and Martinec 1994, Martinec and Vaniček 1994a, Martinec and Vaniček 1994b, Martinec 1996, Vaniček et al. 1995, 1996, 1999). A summary of the theoretical principles and corresponding mathematical expressions is in Ellmann and Vaniček (2007) and abbreviated in Section 2 of this article. In parallel with the theoretical refinement of the Stokes-Helmert approach, UNB has developed the SHGeo (Stokes-Helmert precise Geoid) software package, which remains a work in progress, but has been used to deliver several Canadian geoid models (e.g., Vaniček and Kleusberg 1987, Vaniček et al. 1991, 1996b). It has also been used elsewhere (e.g. Kadir et al. 1999; Janak et al. 2006, Avalos-Naranjo et al. 2011, Blitzkow et al. 2012) and has now been licensed to a private geophysical survey company.

In this contribution, we present our experiences (both good and bad) when testing UNB’s current Stokes-Helmert techniques using AusSEGM (Baran et al. 2006), herein abbreviated yet further to ASF (Australian synthetic field). This testing was done for three reasons: (i) construction of the ASF was a joint Curtin-UNB effort, (ii) the ASF includes a topographic component that aims to avoid the limitations imposed by SEGMs based solely on spherical harmonic models, and (iii) the area covered by ASF is similar in size to Canada, where the UNB Stokes-Helmert technique has been used most extensively.

The testing methodology is rather straightforward: The synthetic gravity anomalies and topography are input to the SHGeo suite as would be done with observational data. A geoid model is computed from these synthetic input data and then compared to the synthetic geoid from the ASF. The differences are inspected and attempts made to isolate the causes of any differences that are larger than expected.

This has proven to be a most useful exercise because it has been a give and take process, which led to vigilance and introspection over both the UNB Stokes-Helmert technique and the construction of the ASF, as well as various refinements to both group’s computer codes. Basically, every stage in the computational chain was examined for correctness and accuracy. The simplest example of this process was ensuring that computer code developed in different hemispheres and for different grid spacing was transferable. Another example was the seemingly simple evaluation of the reference field (i.e., geoid contributions from a spherical harmonic model), which led to both groups’ software being made consistent, including the zero degree and zonal terms (Vaniček and Kleusberg 1987, Sect 5; cf., Smith 1998). Other stages caused us to revert to using EGM2008 (Pavlis et al. 2012) alone (cf. Tziavos 1996, Novak et al. 2001) to reconcile differences where the influence of the topography was not essential.

2. Outline of the Stokes-Helmert technique

2.1. Disturbing potential and gravity anomaly in the Helmert space

Over continental regions, the geoid is mostly located inside the topographic masses, whereas gravity measurements (and levelling to compute gravity anomalies) are taken on the topographic surface. To satisfy the boundary condition of Stokes's problem, the surface anomalies need to be continued downward to the geoid level. Downward continuation requires the knowledge of the mathematical prescription of the physical behavior of the quantity to be continued downwards, which is most often harmonicity. For harmonic behavior, the well-known Laplace equation (e.g., Heiskanen and Moritz 1967, Eq. 1-15) must hold everywhere external to the geoid.

The presence of topography violates the harmonicity condition. Therefore, in order to achieve harmonicity of the disturbing potential, UNB uses Helmert's second condensation method (Helmert 1884), where the masses located above the geoid are condensed to form a surface layer on the geoid. Due to this condensation, however, the gravity field in Helmert's space, here expressed by the gravitational potential $W^h(r, \Omega)$, becomes slightly different from the actual gravity potential, $W(r, \Omega)$, in real space. The corresponding Helmert disturbing potential, $T^h(r, \Omega)$, is obtained from:

$$T^h(r, \Omega) = W^h(r, \Omega) - U(r, \Omega) = T(r, \Omega) - [V^t(r, \Omega) - V^{ct}(r, \Omega)] - [V^a(r, \Omega) - V^{ca}(r, \Omega)] \quad (1)$$

i.e., by subtracting the potentials of topographic/atmospheric masses (i.e., $V^t(r, \Omega)$ or $V^a(r, \Omega)$) from the actual disturbing potential, $T(r, \Omega) = W(r, \Omega) - U(r, \Omega)$, and adding at the same time the corresponding potentials of the condensation layer (i.e., $V^{ct}(r, \Omega)$ and $V^{ca}(r, \Omega)$). Normal gravity potential, as usual, is denoted by $U(r, \Omega)$.

The following notation and quantification will be used throughout this paper. The geocentric position (r, Ω) of any point is represented by the geocentric radius $r(\Omega)$ and a pair of geocentric spherical coordinates $\Omega = (\phi, \lambda)$, where ϕ and λ are the spherical latitude and longitude, respectively. The region of interest of this study is the space external to the geoid and the geoid itself. Thus Eq. (1) and all subsequently developed expressions are valid for $r(\Omega) \geq r_g(\Omega)$, where $r_g(\Omega)$ is the geocentric radius of the geoid surface. Solving for the geoid and other related corrections, the mathematical operations need often be taken over the total solid angle $\Omega_0 = [\phi \in \langle -\pi/2, \pi/2 \rangle, \lambda \in \langle 0, 2\pi \rangle]$. The cases employing a different spatial domain will be pointed out specifically when used.

It has been shown (Vaníček and Martinec 1994) that Helmert's disturbing potential is harmonic at every point outside the geoid to an accuracy consistent with the assumption of a constant topographical mass density. Since the disturbing potential cannot be measured directly, the boundary value problem (BVP) of the third kind (e.g., Heiskanen and Moritz 1967, Chap. 1-5) has to be formulated and solved. In geoid determination, some type of gravity anomalies, referred to the geoid level, serve as the boundary values for this problem. The question here is: what are the appropriate gravity anomalies required for solving the geodetic BVP in Helmert's space? The fundamental gravimetric equation (e.g., Heiskanen and Moritz, 1967, Eq. 2-148) for the Helmert anomaly, $\Delta g^h(r, \Omega)$, takes the form (Vaníček et al. 1999)

$$\begin{aligned}\Delta g^h(r_t, \Omega) &= -\frac{\partial T^h(r, \Omega)}{\partial n} \Big|_{r=r_t(\Omega)} + \frac{\partial \gamma(r, \Omega)}{\partial n} \Big|_{r=r_t(\Omega)} \frac{T^h(r_t, \Omega)}{\gamma[r_0(\phi) + H(\Omega)]} = \\ &= -\frac{\partial T^h(r, \Omega)}{\partial r} \Big|_{r=r_t(\Omega)} + \varepsilon_{\delta_g}(r_t, \Omega) - \frac{2}{r_t(\Omega)} T^h(r_t, \Omega) - \varepsilon_n(r_t, \Omega)\end{aligned}\quad (2)$$

where $\varepsilon_{\delta_g}(r_t, \Omega)$ is the ellipsoidal correction to the gravity disturbance (due to replacing the derivative with respect to the ellipsoidal normal n by a more convenient radial derivative along the radius r , see the first term on the right-hand side of Eq. (2)); $\partial\gamma/\partial n$ is the linear approximation of vertical gravity gradient of the normal gravity field. The ellipsoidal correction $\varepsilon_n(r_t, \Omega)$ for the spherical approximation allows replacement of the “ellipsoidal” term by a more simple “spherical” term $-\frac{2}{r_t(\Omega)}T^h(r_t, \Omega)$. These two ellipsoidal corrections are needed to account for the deviation of the elliptical shape of the Earth from the spherical approximation, which is usually adopted in the formulation of the BVP (e.g., Stokes 1849). Therefore, it should be emphasized that these ellipsoidal effects do not stem from the application of Helmert’s condensation scheme. A detailed derivation of the two ellipsoidal corrections and their exact physical meaning along with numerical examples are given in Vaniček et al. (1999).

All the quantities in Eq. (2) refer to the surface of the Earth, $r_t(\Omega) \cong r_g(\Omega) + H(\Omega)$, where $H(\Omega)$ is the height of the topographic surface. Note that in Eq. (2) the normal gravity $\gamma[r_0(\phi) + H(\Omega)]$ is referred to the telluroid $r_0(\phi) + H(\Omega)$ where $r_0(\phi)$ is the geocentric radius (a function of latitude only) of the surface of the reference ellipsoid. Note that orthometric heights are used everywhere in the Stokes-Helmert technique.

Importantly, the Helmert gravity anomaly can also be expressed via commonly used free-air gravity anomalies $\Delta g(r, \Omega)$ (Molodensky 1945). After a few manipulations one arrives at (cf. Vaniček et al. 1999)

$$\begin{aligned}\Delta g^h(r_t, \Omega) &= \Delta g(r_t, \Omega) + \frac{\partial [V^t(r_t, \Omega) - V^{ct}(r_t, \Omega)]}{\partial r} + \frac{2}{r_t(\Omega)} [V^t(r_t, \Omega) - V^{ct}(r_t, \Omega)] + \\ &+ \frac{\partial [V^a(r_t, \Omega) - V^{ca}(r_t, \Omega)]}{\partial r} + \frac{2}{r_t(\Omega)} [V^a(r_t, \Omega) - V^{ca}(r_t, \Omega)] + \varepsilon_{\delta_g}(r_t, \Omega) - \varepsilon_n(r_t, \Omega)\end{aligned}\quad (3)$$

where the second and third terms on the right-hand side are the direct and secondary indirect topographic effects (denoted hereafter as DTE and SITE, respectively) on the gravitational attraction. The second line of Eq. (3) contains their atmospheric counterparts and the two ellipsoidal corrections. As the ASF does not contain any atmospheric masses, this particular feature of the Stokes-Helmert technique is not used in this testing. Equation (3) is an important formula, as it contains all necessary components to convert free-air gravity anomalies (these are usually derived directly from gravity and height observations) into a quantity more suitable for geoid determination.

More specifically, the product of Helmert anomaly and geocentric radius, $\Delta g^h \cdot r$, is harmonic (Vaniček et al. 1996), and therefore Helmert’s anomalies can be continued downward from the Earth’s surface to the geoid level. This is not the case for free-air gravity anomalies.

2.2. Topographic effects on gravity and downward continuation of gravity anomalies

One focus of the ASF was on the inclusion of topographic effects, which is the most demanding task in precise geoid determination. In the Stokes-Helmert approach, the topographical effects on gravity are formulated (and evaluated) in their spherical form, instead of the widely used planar approximations. The spherical models are closer to reality; in addition to that, planar models cannot be used if the harmonization of the Earth's gravity field is required; see Vaniček et al. (2001).

The DTE and SITE in Eq. (3) are evaluated on the Earth's surface. Newton's integral in spatial form can be used for evaluating the gravitational potential of the topographical masses; the gravitational potential of the condensation layer can be treated similarly. The SITE is then computed as the difference between the two expressions; see the third term on the right-hand side of Eq. (3). The DTE is obtained by taking the radial derivatives of the corresponding potentials of the topography and the condensed layer. The derivations of and final expressions for the DTE and SITE can be found in (Martinec and Vaniček 1994b, Martinec 1993, 1998, Vaniček and Martinec 1994, Vaniček et al. 1999 and Ellmann and Vaniček 2007, Eqs. 18-23).

Once the Helmert gravity anomaly is evaluated on the Earth's surface, it has to be continued downward (or upward, when the Earth's surface is located beneath the geoid) to the geoid. In the UNB realisation of the Stokes-Helmert approach, the downward continuation is evaluated using Poisson's integral equation (e.g., Heiskanen and Moritz 1967, p. 317). Downward continuation is the inverse problem to the original Poisson integral. The integral equation is re-written as a system of linear equations in matrix-vector form (Martinec 1996), which is then used for solving the inverse problem, i.e., computing the unknown elements of the vector $\Delta \mathbf{g}^h(r_g, \Omega)$. For more details, the interested reader is referred to Sun and Vaniček (1998) and Vaniček et al. (1996).

2.3. Solution to Stokes's boundary value problem

Using the original Stokes (1849) formula, gravity anomalies over the entire Earth are required for geoid determination. In practice, however, the area of available anomalies, i.e., the area of numerical integration, is limited to some neighborhood (Ω_{ψ_0}) around the computation point. The truncation bias that occurs when the far zone ($\Omega_0 - \Omega_{\psi_0}$) is neglected in the integration, can be reduced by modifying Stokes's formula, e.g., following Molodensky et al. (1960). The circular neighbourhood (Ω_{ψ_0}) of radius Ψ_0 is, in the sequel, called the Near Zone (NZ) and the Far Zone is denoted by FZ. The UNB approach also employs the generalized Stokes scheme, which spectrally decomposes the geoidal height into the long-wavelength reference spheroid of degree L (computed from satellite-only derived spherical harmonics) and the high-frequency geoid, which is computed from the terrestrial data (Vaniček and Sjöberg 1991):

$$N(\Omega) = \frac{R}{4\pi\gamma_0(\phi)} \iint_{\Omega_{\psi_0}} S^L(\psi_0, \psi(\Omega, \Omega')) \left(\Delta g^h(r_g, \Omega) - \sum_{n=2}^L \Delta g_n^h(r_g, \Omega) \right) d\Omega' + \frac{R}{2\gamma_0(\phi)} \sum_{n=2}^L \frac{2}{n-1} \Delta g_n^h(r_g, \Omega) + \frac{V^t(r_g, \Omega) - V^{ct}(r_g, \Omega)}{\gamma_0(\phi)}, \quad (4)$$

where R is the mean radius of the Earth, the modified Stokes function $S^L(\psi_0, \psi(\Omega, \Omega'))$ can be computed according to Vaníček and Kleusberg (1987), $\psi(\Omega, \Omega')$ is the geocentric angle between the computation and integration points, and $d\Omega'$ is the infinitesimal area of the integration element. The modification limit L can be selected by diverse criteria, which were already discussed in (ibid).

Note that the global geopotential model (GGM)-derived anomaly, $\Delta g_n^h(r_g, \Omega)$, in Eq. (4) is related to the geoid level. Since GGMs rely upon external-type series of spherical harmonics, then inside the topographical masses, the GGM-derived gravity quantities may be biased. To account for this bias, the topographic masses should be mathematically removed or condensed. The effect of Helmert's second condensation method on the long-wavelength spectrum of the geopotential is accounted for by Vaníček et al. (1995, Eq. 20), which utilizes the spherical harmonic coefficients of the squared topographic heights, H_{nm}^2 . Thus, Helmertised geopotential coefficients can be used for computing the harmonics $\Delta g_n^h(r_g, \Omega)$ in Eq. (4).

The Stokesian integration with Helmert's residual anomalies (i.e., the first term on the right-hand side of Eq. (4)) results in the Helmert residual co-geoid. Since the low-degree reference field is removed from the anomalies before Stokes's integration, the long-wavelength contribution to the geoidal height (e.g., Heiskanen and Moritz 1967, p. 97), i.e., the reference spheroid, must be added to the residual geoid.

As already noted, the condensation of the topographic masses yields the Helmert potential that is slightly different from the actual potential. Consequently, the Helmert co-geoid does not exactly coincide with the geoid in real space. The effect that causes this change is called primary indirect topographical effect for Helmert's condensation (PITE^{Hc}). Thus, the last term on the right-hand side of Eq. (4), PITE^{Hc}, transforms the Helmert co-geoid into the real space geoid. An important advantage of Helmert's second condensation method is that PITE^{Hc} is nowhere larger than a few decimeters, perhaps a metre, worldwide.

It should be noted that the density of the condensation layer was chosen here in agreement with the principle of mass conservation in spherical approximation. In this case, the residual topographic potential $\delta V'$ does not include the zero-degree spherical harmonic, but it does include the first degree spherical harmonics (Martinec 1998, Chap. 2.1). Even though their effect is quite small (not exceeding a few cm in the study area), it was evaluated by the expressions derived by Martinec (1998, Eq. 2.17) as we shall see later.

3. Construction of the Australian Synthetic Field

This section provides a brief overview on the methodology used to construct the ASF. For more details, the interested reader is referred to Baran et al. (2006). ASF is a regional high-resolution (1 by 1 arc-min) SEGM based on a combination of a source and effect model. The assumed errorless EGM96 GGM (Lemoine et al. 1998) has been taken as the effect model, while the source model is based on a high-resolution (3 by 3 arc-sec) synthetic digital elevation model (SDEM, Fig. 1), which is a fractal surface based on the GLOBE v1 DEM (Hastings and Dunbar 1998). While the SDEM extends beyond the Australian mainland (e.g., 112-155°E, 8-45°S) ASF is only intended to be used over the Australian mainland.

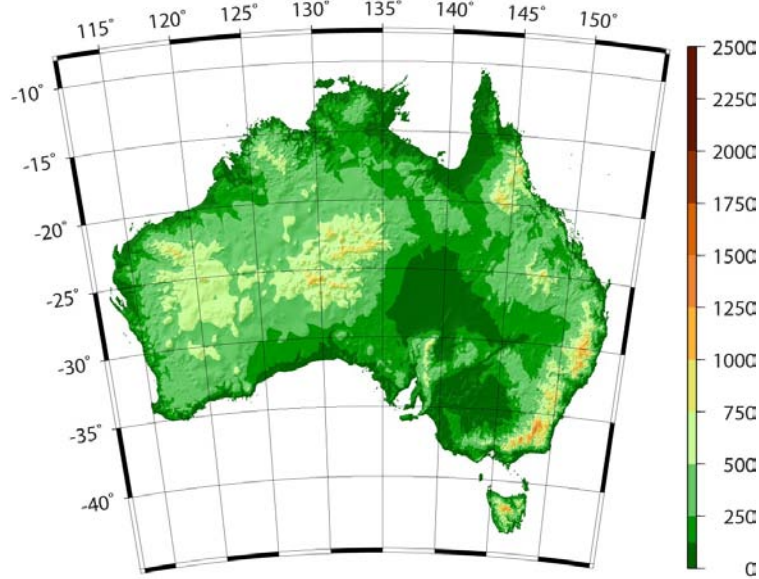


Fig. 1: Synthetic Digital Elevation Model (3-arc-sec by 3-arc-sec) over Australia (Min: 0.0 m, Max: 2,405.4 m, Mean: 128.2 m, Std. Dev.: 194.0 m). Lambert projection. Units in metres.

The long-wavelength effect part (degrees ≤ 360) is taken from EGM96 evaluated on the surface of the SDEM for free-air gravity anomalies and at zero elevation for the geoid height. As both parameters are given within the theory of Molodensky, free-air gravity anomalies are of Molodensky-type and geoid heights are quasi-geoid heights (e.g., Heiskanen and Moritz 1967). No corrections were applied to convert the quasi-geoid heights into geoid heights. Also no ellipsoidal corrections were applied, but the user can apply them following the approaches of, e.g., Jekeli (1983), Vaniček et al. (1999), Ardalán and Grafarend (2001) and Sjöberg (2003).

Using forward gravity modelling techniques, the short-wavelength source part (degrees > 360) is derived from the SDEM over Australia augmented globally by the 5 by 5 arc-min JGP95E (Joint Gravity Project) DEM (Lemoine et al. 1998, Chap. 2). For the purpose of deriving the effect on potential and gravity, the SDEM heights have been interpreted as heights above the (quasi-)geoid approximated by the surface of the GRS80 reference ellipsoid. Furthermore, within the Newtonian integration, the ellipsoidal mass elements (defined by the geographic data grid) have been replaced by mass-equal prisms of the same height (e.g., Kuhn 2000, 2003) using a constant mass density of 2670 kg/m^3 .

As the so-derived gravitational effects contain signals at all spectral frequencies (only limited by the 3 by 3 arc-min resolution of the SDEM), the gravitational effects have been spectrally separated into long- and short-wavelength parts. The spatial representation of the long-wavelength part (≤ 360) was derived through a spherical harmonic analysis and subsequent synthesis evaluated on a 1 by 1 arc-min grid and subsequently removed from the full forward gravity modeled signal to leave the short-wavelength part (>360). Finally, ASF free-air gravity anomalies $\Delta g_{AusSEGM}$ and quasigeoid heights $N_{AusSEGM}$ (cf., Figs. 2 and 3) are derived through a superposition of long- and short-wavelength components given by (Baran et al., 2006, Eqs. 9 and 11)

$$\Delta g_{AusSEGM} = \Delta g_{EGM96}^{360} + \delta g_{SDEM}^{>360} \quad (5)$$

$$N_{AusSEGM} = N_{EGM96}^{360} + \delta N_{SDEM}^{>360} \quad (6)$$

where Δg_{EGM96}^{360} and N_{EGM96}^{360} are the long-wavelength parts for the Molodensky free-air anomaly and quasigeoid height obtained by evaluating EGM96 (to degree 360), respectively. The corresponding short-wavelength parts, obtained through forward gravity modelling of the SDEM-implied masses, are $\delta g_{SDEM}^{>360}$ for gravity and $\delta N_{SDEM}^{>360}$ for quasigeoid heights. The latter is obtained by applying Bruns's formula (e.g., Heiskanen and Moritz 1967 p 85) to the corresponding effect on gravitational potential.

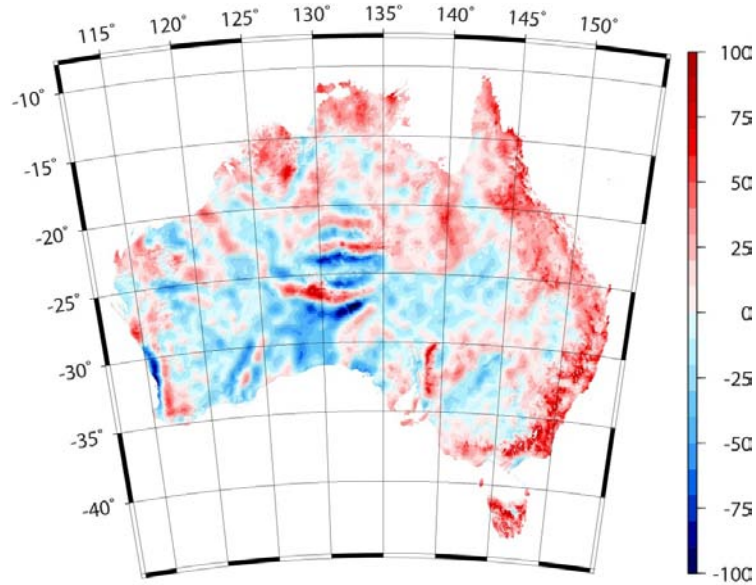


Fig. 2: ASF free-air gravity anomalies (Min: -113.2 mGal, Max: 190.9 mGal, Mean: 4.1 mGal, Std. Dev.: 23.7 mGal). Lambert projection. Units in mGal.

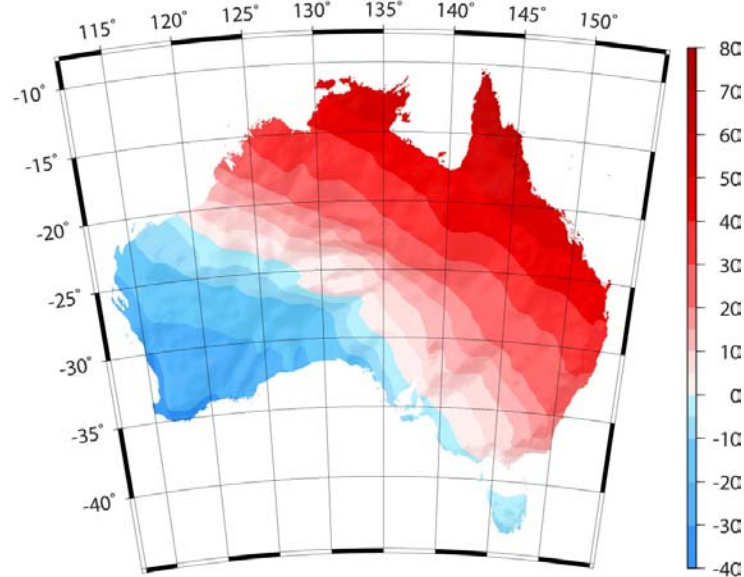


Fig. 3: ASF geoid height (Min: -34.3 m, Max: 73.0 m, Mean: 14.6 m, Std. Dev.: 24.1 m). Lambert projection. Units in metres.

It should be mentioned that the spectral separation is done via a surface spherical harmonic expansion instead of solid spherical harmonics, thus implicitly assuming a harmonic continuation between values given at the surface of the SDEM and the quasigeoid. However, this is considered not critical when doing the spectral separation as the functional values are not of interest between these two surfaces. Related to this it is stated in (Baran et al. 2006): “*As long as there exists a one-to-one correspondence between the coordinates of points at the simulated Earth’s surface and the spherical polar coordinates (which is the case for ASF), the spectral separation can be achieved using surface spherical harmonics (e.g. Jekeli 1988).*”

4. Philosophical differences

It is now instructive to point out some key philosophical differences between the UNB Stokes-Helmert technique and the construction of the ASF, which will explain some - but not all - of the discrepancies presented later. The ASF took the long- and medium-wavelength global image of the Earth’s gravity field from EGM96 (Lemoine et al. 1998) and high-pass filtered gravitational effects induced by high-resolution local and lower-resolution global topography to produce: (i) synthetic gravity acceleration values on the Earth’s surface, and (ii) quasi-geoid heights referred to the reference ellipsoid consistent with EGM96 (Baran et al. 2006). On the other hand, the UNB Stokes-Helmert technique is supposed to produce geoid heights referred to any chosen reference ellipsoid when a satellite-only reference field, local topography and topographic density are given. For the experiments herein, no 3D topographic density model and no atmospheric attraction were used in either approach.

Next, different approaches have been taken in Australia and Canada for historical reasons of their respective height systems and the need to come up with different ‘products’ (cf. Vermeer 1998). Canada is officially adopting a geoid-based vertical system (Véronneau and Huang 2011, Vaniček et al. 2012) and all Canadian data used in this study refer to a rigorous Helmert orthometric height system (cf., Kingdon et al. 2005). Australia uses an orthometric system based on normal rather than actual gravity (called the “normal-orthometric” height system) which seems to give heights that are closer to normal rather than orthometric heights (e.g., Featherstone and Kuhn 2006). As such, the geoid is more appropriate in Canada, whereas the quasigeoid is more appropriate in Australia. However, there is a spatially dependent inconsistency between the quasigeoid and the normal-orthometric heights system (Filmer et al. 2010). Also, there is an inconsistency between the classical geoid and the Helmert approximation of the orthometric height, driven principally by the lack of 3D topographical density information (e.g., Tenzer et al. 2005, Santos et al. 2006). This had some unintended consequences, as outlined below and elaborated upon in the sequel.

The ASF was computed at the time when EGM96 was the ‘state of the art’ GGM (Lemoine et al. 1998). We acknowledge that newer data sets are now available to re-compute an SEGM, e.g., EGM2008 (Pavlis et al. 2012) or SRTM (Farr et al. 2007), but it is a computationally demanding task and, moreover, the ASF had already been available to us. In addition to this, we believe that the dates associated with the data sets used are irrelevant for the purpose we used the data sets for.

The choice to use the degree-360 expansion of EGM96 was largely driven by the Australian approach to quasigeoid determination before that time (e.g., Featherstone et al. 2001). However, this

resulted in a problem that - in retrospect - is undesirable. Baran et al. (2006), assuming that the EGM96 spherical harmonic coefficients were error free, generated quasigeoid heights to degree 360, then added high-pass filtered geoid heights induced by the synthetic topography. To convert EGM96 quasigeoid heights to geoid heights, one can make use of the simple approximation by Rapp (1997), also see (Heiskanen and Moritz 1967), and - as such - it does not provide a fully rigorous test of topographical effects on the geoid.

This is at odds with the UNB approach of using a low-degree satellite-only reference field (RF). There is hence an argument for constructing a synthetic field that is more aligned with the geoid (or quasigeoid) computation philosophies to be tested. However, the brute-force numerical integration to forward model high-resolution topographic effects was quite considerable although this is no longer such a restriction with the wider availability of supercomputing facilities (e.g., Kuhn et al. 2009, Fellner et al. 2011).

It should be pointed out that this study is not concerned with testing the computation of the coefficients that go in to the RF contribution, instead assuming the EGM96 coefficients to be error-free. We also only consider an error-free ASF, rather than simulating errors to see how they propagate. It was considered much more important to first test the UNB technique in a scenario unclouded by errors. Once a satisfactory result is achieved, then error simulation could be included, but determining reasonable and reliable error estimates (e.g., for the low-frequency terrestrial data) is notoriously difficult.

5. Improvements to the Australian Synthetic Field

In our two-way process, we have discovered some problems with the ASF, necessitating some corrections and also raising some questions (these remain unanswered to date, but will be discussed later). A small error in the application of the spherical harmonic analysis and synthesis software (determination of coefficients from spatial data and production of spatial data) has been discovered related to a mismatch between cell/pixel and grid/node registered values on a geographical grid. The magnitudes of these aliasing errors are well below 1 mGal and 10 mm for the synthetic gravity and geoid height, respectively.

A new spherical harmonic analysis of the high-pass filtered (>360) gravitational acceleration and potential also revealed leakage into the <360 field (Fig 4), probably caused by the inexactitude of the registration of values on a geographical grid. In addition, the modelling of some discontinuous gravitational effects of topography when constructing the ASF caused Gibbs fringing in the high-pass filtered synthetic gravity data of magnitudes less than 1 mGal. These errors necessitated the generation of correction grids (Fig 5) to ASF that have been applied in the tests presented in Section 7 and later.

No isostatic compensation modelling was applied during the construction of the ASF, so the forward modelling of the >360 topographical effects on gravity and geoid generates a larger gravitational signal than a compensated topography. However, this is not such a bad thing as it means the residual gravity anomalies, after being Stokes-integrated, generate more pronounced geoid undulations, thus placing more of a challenge on the numerical aspect of the computations. Finally, as

the ASF is in a completely spherical approximation with no atmosphere, ellipsoidal corrections, or atmospheric effects, so the evaluation of these terms was ‘turned off’ in the SHGeo software.

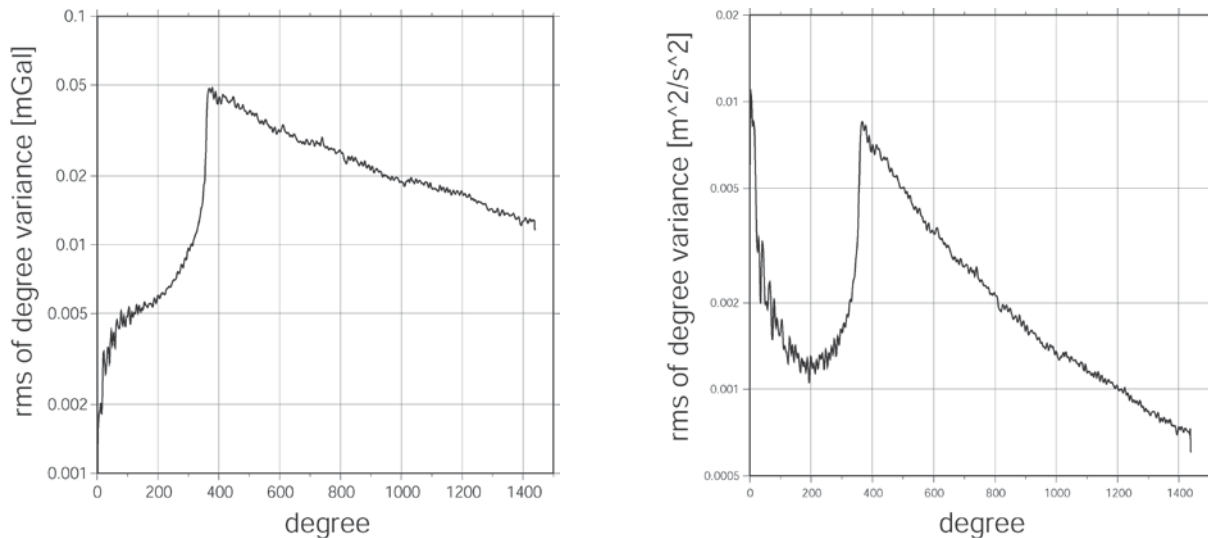


Fig. 4: Degree variance of the coefficients generated from a spherical harmonic analysis of the ASF’s high-pass filtered gravitational acceleration (left panel) and potential (right panel). Note the spurious energy build-up in the low degrees.

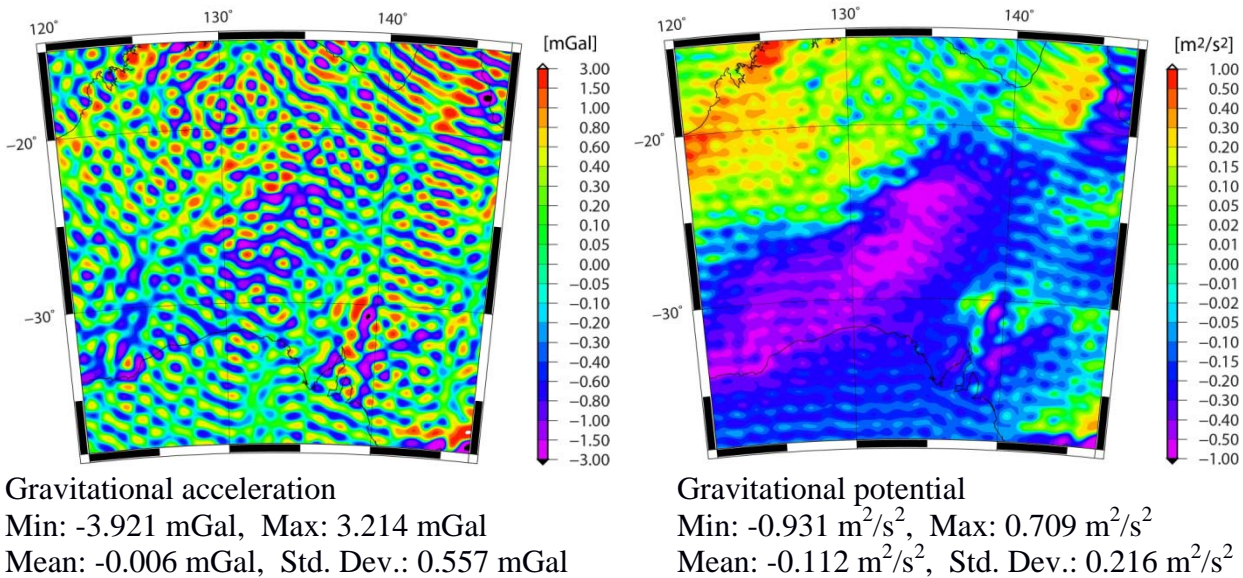


Fig 5: Correction grids applied to ASF to remove the errors mentioned in the text except the aliasing and the Gibbs effect; gravitational acceleration (left panel) and potential (right panel).

6. Improvements to Stokes-Helmert’s technique

The input data for all the computations described herein are bound by 112°E-155°E, 8°S-45°S (Baran et al., 2006), but the results are only compared within six degrees (the maximum Ψ_0 considered) of each boundary (i.e., 118°E-149°E, 14°S-39°S) so as to avoid the propagation of spurious differences due to edge effects. When reporting the differences between the SHGeo results and the ASF, we do it by giving four numbers computed for the comparison area: the mean (bias), standard deviation, the minimum and maximum errors. Sometimes, we refer to the “range”, which is, of course, the difference between the maximum and the minimum in the area of interest.

After doing a first-run computation on the ASF published by Baran et al. (2006), and without the corrections described above, we encountered differences between the SHGeo geoid model computed from the synthetic gravity and topography and the ASF geoid that were larger than expected (Fig 6). The largest point differences were encountered in the region of the Australian Alps, located in the south-east corner of the study area; also the bias of 34mm is worrisome. This result obtained with $\Psi_0 = 6^\circ$ and $L = 90$ (a combination later discarded in favour of the four combinations: $\Psi_0 = 2^\circ$ and 6° , $L = 20$ and 60 used hence; and as such not even recomputed with the corrections to ASF applied), led to an examination of each stage of the computation chain, and each package in the UNB SHGeo software suite was re-examined and checked for bugs. Several small errors in SHGeo were discovered and corrected, as well as the development of new approaches, all of which will be described in the sequel.

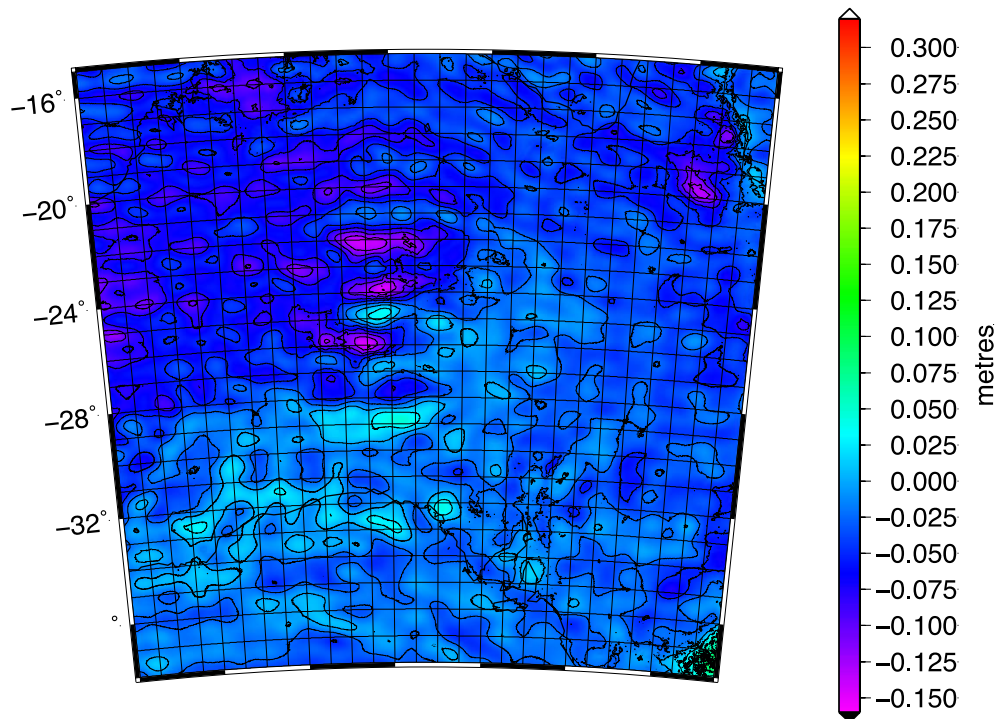


Fig 6: Differences (in metres) between the geoid model computed by SHGeo software and the ASF “ground truth” (Min = -162 mm, Max = 157 mm, Mean = -34 mm, Std. Dev. = 25 mm.)

The primary candidate for testing after this preliminary comparison was treatment of the notoriously ill-conditioned downward continuation (DWC) of gravity anomalies on the topography to the geoid (e.g., Sun and Vaníček 1998). Even though the DWC process is well posed (for a detailed discussion of this problem, see Section 8), the computations are invariably laborious and the results suffer from spurious high-frequency noise. Thus, a new technique was designed (Kingdon and Vaníček 2011) that was supposed to cure this problem, and it seems that – to a large degree – it did. Importantly, this work may not have been conducted had the ASF not led to the re-examination of the Stokes-Helmert technique. More about this new technique will be given in Section 8, after all of the tests on the SHGeo technique have been discussed.

7. Further tests on SHGeo and discussions

7.1 Contribution of the reference field

As alluded to in the Introduction, even the seemingly simple contribution of the RF had to be attended to in our process of systematic checking. The ASF had been computed with respect to the defining constants (i.e., gravity-mass constant GM , even zonal harmonic J_2 , and equatorial radius a) of EGM96 and not GRS80. Thus the zero-degree term and scale change via zonal harmonics of the reference ellipsoid (Vaníček and Kleusberg 1987, Sect 5), cf., (Smith 1998) were not implemented. When the geoid model was evaluated with respect to GRS80, there was a bias (zero-degree term) amounting to -0.93 m. As such, the option of “referring the solution to GRS80” was disabled in the SHGeo software so that the geoid model solutions were compatible with the reference ellipsoid implicit to EGM96.

7.2 Hörmander correction

When the topographic masses are condensed onto the geoid, the Helmertised gravity anomalies Δg^h , used subsequently to compute the Helmertised anomalous gravitational potential by Stokes’s integration, contain first-degree spherical harmonics. This causes the failure of the origin of coordinate system to coincide with the centre of mass of the Helmertised Earth. Since the Stokes solution, used for both the geoid and quasigeoid determination, is only valid in a coordinate system whose origin coincides with the centre of mass of the Earth, Hörmander (1976, 2nd eq., p.30) introduced three extra unknowns to balance the number of unknowns and conditions.

The three unknowns, expressing the shift in the centre of mass, called hereafter the Hörmander corrections (called also “Hörmander’s trick” by Holota and Nesvadba, 2008) were determined by Martinec (1993, 1998) based on the preservation of mass scheme. His determination of (-6; -15; 2) mm was for the TUG87 (Technical University Graz 1987) DEM, but since this result depends mainly on well-established low-frequency terms, and since the magnitude of the vector is small to start with, this should be quite close to the result from any more modern DEM (e.g., JGP95E augmented with a higher resolution Australian SDEM used for the ASF). The vertical displacement of the geoid at a particular point is given to a reasonable accuracy by the projection of the displacement vector in the direction of the ellipsoidal normal. Over the ASF test area, the Hörmander correction varies from -4 mm to 10 mm, exhibiting a tilt in the east-west direction.

7.3 Modified kernel and truncation coefficients

The spectral patterns in the first-run test (Fig 6) led us to question whether the spheroidal Molodensky modification and truncation coefficients used in the Stokesian integration over the cap of radius ψ_0 (Vaniček and Kleusberg 1987; Vaniček and Sjöberg 1991) were generated properly, and to what spherical harmonic degree the FZ contribution needed to be evaluated. Although in all of the UNB computations, Stokes's integral is effectively evaluated globally (i.e., not just within a spherical cap; i.e., the "FZ contribution," "truncation bias," or "truncation error" is accounted for – see Appendix A), it was still desirable to minimize the FZ component as much as possible. This is done to minimize the error passed into the solution from the RF.

In addition, the UNB approach exclusively uses a RF whose coefficients are determined by satellite methods alone, for the following two reasons:

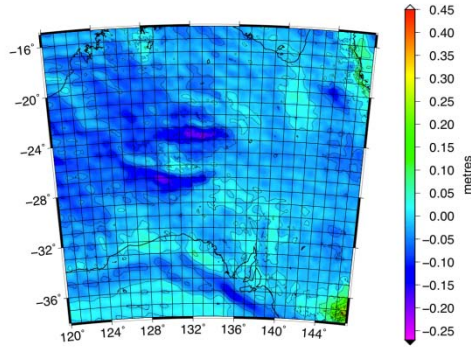
1. It is desirable to have no correlation between the information contained in the global spherical harmonic model and terrestrial data (used in the evaluation of the NZ contribution, i.e., over the spherical cap in Stokes's integral);
2. A global spherical harmonic model, no matter how good it may be, approximates the local information only in an asymptotic manner, i.e., getting closer and closer to the truth only with increasing degree of expansion. This is, of course, only true if and only if the data, e.g., terrestrial gravity or satellite altimetry that went into the coefficient determination, were error-free *over the entire globe*. This, in reality, is not the case.

The spheroidal Molodensky modification coefficients were used for kernel modification in the NZ and the truncation coefficients used in the evaluation of the corresponding FZ contribution. They were validated by comparison, coefficient by coefficient, with a set generated from the *trunc.f* program developed at Curtin University (Featherstone et al. 2004). This revealed that, when calculated for the same modification degree and NZ radius, the coefficients from both programs were identical for all values of radius ψ_0 and degree L tested.

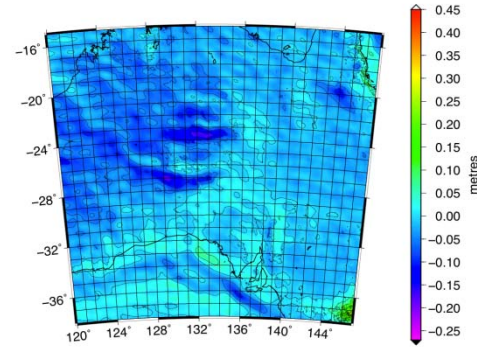
7.4 Effect of varying ψ_0 and L in the NZ and FZ contributions

The need to test for invariance of the solution with respect to the choice of ψ_0 and L came in reaction to the determination of the geoid carried out for four different combinations of ψ_0 and L . The four results, for different combinations of $\psi_0 = 2^\circ$ or 6° and $L = 20$ or 60 , were appreciably different from one another (Fig 7). Theoretically, these results should be invariant with respect to the particular choice of ψ_0 and L (Appendix A), but numerically they are not. The reasons for these differences are not known; see the later discussion.

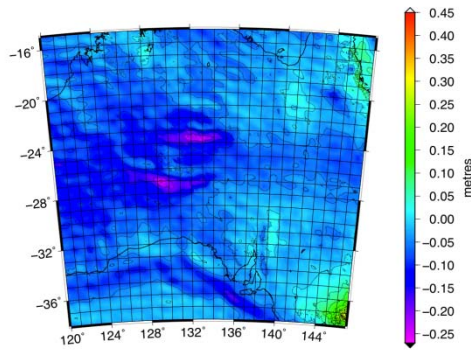
The interesting features of these differences are the persistent maximum over the Australian Alps and the washboard-like feature that persists over central Australia. The former is probably related to topography, whilst the latter is somewhat correlated with the high-resolution part of the ASF. We also mention here that the combination $\psi_0 = 2^\circ$ and $L = 20$ probably should not be used in practice because the cap is too small in relation to the spectral information provided by the RF.



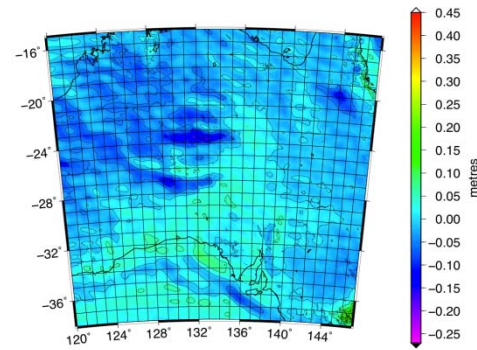
Combination $\psi_0=2 / L=20$
 Min = -224 mm, Max = 435 mm,
 Mean = -29 mm and Std. Dev. = 37 mm



Combination $\psi_0=2 / L=60$
 Min = -205 mm, Max = 382 mm,
 Mean = -22 mm and Std. Dev. = 32 mm



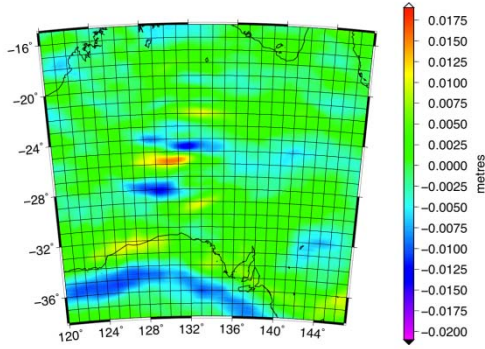
Combination $\psi_0=6 / L=20$
 Min = -260 mm, Max = 448 mm,
 Mean = -53 mm and Std. Dev. = 45 mm



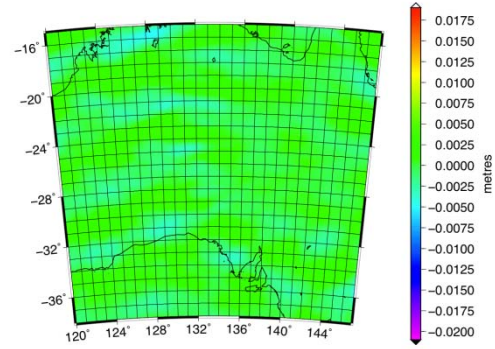
Combination $\psi_0=6 / L=60$
 Min = -186 mm, Max = 361 mm,
 Mean = -8 mm and Std. Dev. = 34 mm

Fig 7: Differences between the geoid model computed by SHGEO and the ASF geoid for four different selections of parameters ψ_0 and L .

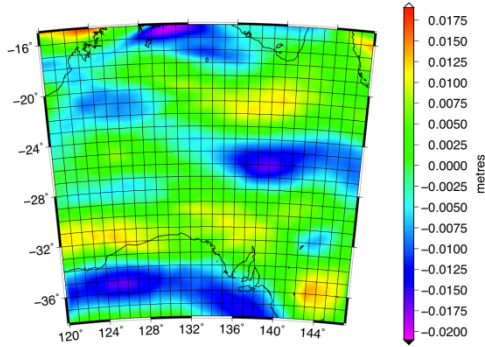
The next step was to determine whether SHGeo program suite might be defective in this regard. From Appendix A, the expectation is that the programs are working properly if the NZ and FZ contributions add to the same residual geoid for any combination of ψ_0 and L . This subsection demonstrates a case where we use a SEGM without topography (cf., Tziavos 1996, Novák et al. 2001; Featherstone 2002), namely gravity anomalies generated from EGM2008 (Pavlis et al. 2012), expanded to degree 2190 in the spatial Stokes integration over NZ, and the EGM2008 spectral form in the FZ. Some discretisation error is expected to appear in these results, but this should be perhaps a few tens of mm, as it is seen in Fig.8.



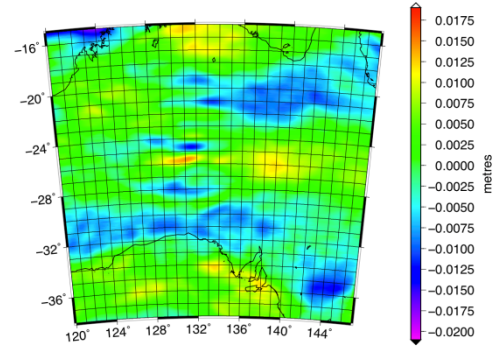
Combination $\psi_0=2 / L=20 - \psi_0=2 / L=60$
 Min = -14.6 mm, Max = 16.4 mm,
 Mean = -0.2 mm and Std. Dev. = 3.7 mm



Combination $\psi_0=2 / L=60 - \psi_0=6 / L=60$
 Min = -5.6 mm, Max = 4.1 mm,
 Mean = 0 mm and Std. Dev. = 1.2 mm



Combination $\psi_0=6 / L=20 - \psi_0=6 / L=60$
 Min = -20.2 mm, Max = 18.2 mm,
 Mean = -0.4 mm and Std. Dev. = 6.7 mm



Combination $\psi_0=2 / L=20 - \psi_0=6 / L=20$
 Min = -17.5 mm, Max = 14.9 mm,
 Mean = 0.2 mm and Std. Dev. = 5.2 mm

Fig 8: Differences between the geoids computed by SHGEO from EGM2008 gravity data, using different values of ψ_0 and L .

Figure 8 is equivalent to Fig 7, except that EGM2008 is used as the SEGM, thus removing the need to compute any topographical effects and DWC of Helmert's anomalies. Figure 8 shows that there is some variance with ψ_0 and L , but it is of lower frequency than that shown in Fig 7. This is what one would expect when using a smoother SEGM. The long-wavelength nature of the discrepancies suggests that their source is not high frequency. This fact points away from the evaluation of topographical effects or the DWC as sources of the differences, since these are mainly high-frequency effects. However, they may be influenced by the aliasing of energy in the low-degree constituents of the high-pass filtered synthetic gravitational potential (cf. Figs. 4 and 5).

7.5 Helmertisation

We also tested whether the Helmertisation of the spectral expression of the RF had been done properly, even though the low-frequency differences in Fig 7 did not raise any suspicion in this regard. The split of the field into low-frequency RF and high-frequency residual field is done in Helmert's space. Therefore, both the spatial results and the spectral results have to be from Helmert's space: the spatially determined gravity anomalies have to be of the Helmert variety (involving the DTE of Helmert's kind, and the appropriate corrections (Martinec and Vaníček 1994b)); likewise for the spectral expression (Vaníček et al. 1995). For clarification, spectral expressions are used in the SHGeo suite in two slightly different contexts: as RF and in the evaluation of the FZ contribution.

We verified that the DTE, the most significant component of the transformation to Helmert's space, was evaluated correctly and applied to EGM96 for the RF calculations. In doing so, we discovered that it was not applied to the spherical harmonic model used for the FZ contribution. The most convenient way to incorporate the Helmertization into SHGeo was to convert the EGM96 coefficients to Helmert's space using the spectral model of squared topographic heights from JGP95E. This was then used as input both to the RF and FZ computations. The change in the FZ contribution as a result of this addition is characterised by the following values: min = -6 mm, max= 4 mm, mean = 0 mm, standard deviation (std. dev.) = 2 mm, which is almost negligible. The RF results were left unchanged because SHGeo had computed these correctly.

8. The downward continuation stability problem

The DWC was re-examined, which led to the design and implementation of a fast and adequately accurate regularisation procedure. DWC is probably the trickiest operation in the SHGeo suite. Even though the problem of DWC is well posed in the Hadamard sense, i.e., the solution is finite, unique and it is a continuous function of the values on the right hand side of the equations (Hadamard 1923), as shown by, e.g., Wong (2002), the system of linear equations created to solve the Fredholm integral equation grows increasingly less stable as the topographical heights increase and the grid spacing decreases. Thus, an accurate solution, which always exists, necessitates special algorithms for matrix inversion and consumes much computer time, yet the results are burdened with high-frequency noise for which there is no physical justification.

Many geodesists use various regularisation techniques (e.g., Tikhonov, 1977) to smooth this high-frequency noise (e.g., Martinec 1998; Novak et al. 2001). The accuracy of the DWC results is controlled by the condition number of the matrix of \mathbf{B} coefficients (Martinec 1996) of the system of linear equations resulting from the transformation of the Fredholm equation to discrete form. Hence, if one knows the accuracy of the quantities to be downward continued, one can set a limit to the accuracy that can be obtained after the DWC process. This is achieved using for instance the simple Jacobi iterative technique for matrix inversion through setting up appropriate convergence criteria, either using the L_2 norm or using the Tchebyshev norm, L_∞ , or maximum absolute value norm L_1 (Kingdon and Vaníček 2011). If the condition number is estimated using its empirically derived relationship with height, this technique is fast. The DWC gravity using this approach is up to 900 mGal different from the results using exact matrix inversion.

As an illustration, this new DWC technique (we wish to call it “tolerance-regularized downward continuation”) was applied to the ASF in two test areas, of 1° and 3° squared, respectively, with data at 1 by 1 arc-min spacing (i.e., the same as the ASF). Area A is at the northern extremity of the ASF data, and spans 9°S to 10°S latitude and 147.5°E to 148.5°E longitude. While this area is outside the ASF’s validity (Sect 3) it provides a challenging topography, thus provides some extreme situation for the DWC. Area B is in southeastern Australia and spans 35°S to 38°S latitude and 147°E to 150°E longitude. This area extends over the most rugged topography in Australia.

Figure 9 shows the DWC effect (equivalent to gravity anomalies on the geoid minus gravity anomalies on the topographic surface) for test area A. The convergence tolerance for the Jacobi method was set to 0.005 mGal, which is roughly equal to one half of the accuracy to which the input values are given. In the relaxed version, this tolerance is scaled by the condition number to deliver the best result that the numerical apparatus can deliver. From Fig 9, the DWC effect is unacceptably noisier when using the precise inversion method, which is identical to an exact solution using LU decomposition, while the new relaxed Jacobi method produces significantly less noise.

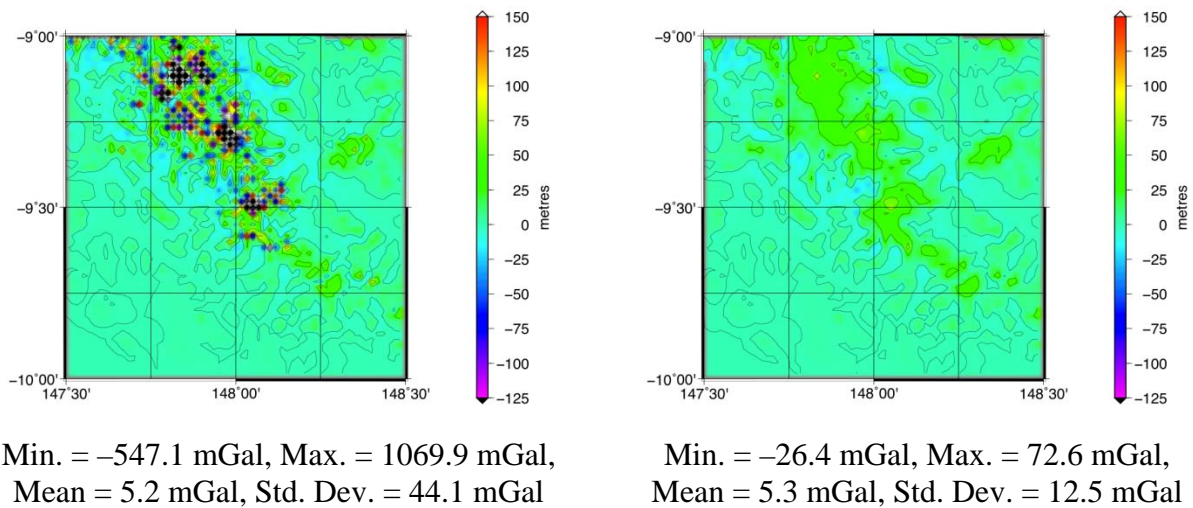


Fig 9: DWC effect for test area A, using rigorous inversion (left panel) and the relaxed Jacobi method (right panel), both with a 0.005 mGal tolerance (scaled up to 6.7 mGal in the relaxed Jacobi method, based on the matrix condition number).

The difference between the geoid models that use these two methods, now in area B, is shown in Fig 10. The Jacobi result used to create Fig 10 was applied with a tolerance of 1 mGal, which is a more common choice than the 0.005 mGal used to illustrate the performance of the relaxed Jacobi method above. Higher tolerances would produce larger noise. The noise in the DWC can have a significant detrimental effect on the geoid model, but effect is quite localised and diminishes as one moves away from the spikes because of the smoothing operation performed by Stokes’s integral.

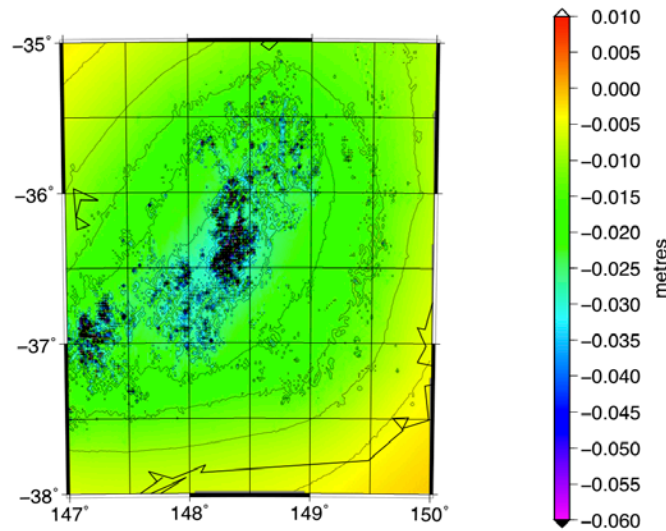


Fig 10: Difference in geoid model results computed from rigorous and relaxed Jacobi methods, for the 2/20 case, in test area B (Min = -498 mm, Max = 183 mm, Mean = -16 mm, Std. Dev. = 9 mm)

Table 1 summarises the computation times for areas A and B. The very similar time consumption for the relaxed Jacobi methods in both areas is because it performs only as many iterations as the conditioning of the **B** matrix would allow, which is normally fewer than 10, and requiring a fraction of the computation time compared to the exact inversion method. By contrast, for areas with large elevations, the exact inversion method uses numerous iterations trying to meet an unattainable tolerance that the numerical apparatus cannot support, hence introducing noise. There are probably other better ways of regularising the inverse solution, but that topic is beyond the scope of this paper. Nevertheless, the tests on the ASF led the UNB team to investigate this improved inversion.

Test Area	Maximum Height (m)	Previous DWC Time (min)	New Jacobi DWC Time (min)
A	3277	>60.00	2.03
B	2289	15.63	2.02

Table 1 CPU time consumption for the rigorous Jacobi vs. relaxed Jacobi methods, using 0.5 mGal tolerance for both.

9. Further tests on the ASF and discussion

After applying the new DWC technique, the combination of $\psi_0 = 6 / L = 60$ fits the ASF “ground truth” quite well, with the descriptive statistics of the differences shown in Fig 11. From the theoretical point of view (Appendix A), the essential requirement for the invariance of the residual geoid model with respect to the choice of ψ_0 and L is that the analysed field be identical when expressed in the spatial way (as given by gravity data) and the spectral way (as by spherical harmonics for the RF and FZ contributions).

The differences between the geoid model values obtained from spectral and the spatial expressions of the ASF are quite small and somewhat correlated with topography. On the other hand, the differences between gravity expressed in spectral and spatial ways are more significant. Note that there is no spectral representation of the ASF beyond degree and order 360, but for the purpose of the above comparison, the ASF’s high-frequency part has been spectrally analysed so to check there is no energy in the corresponding low-degree (≤ 360) coefficients.

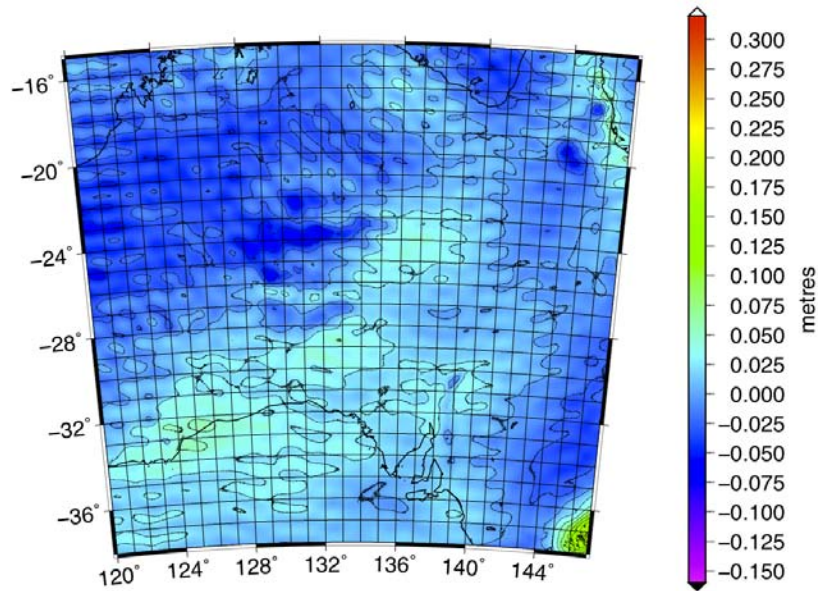


Fig 11: Difference between our best fitting solution (Ψ_0 and L combination 6/60) and the ASF.
(Min = -77 mm, Max = 204 mm, Mean = 1 mm Std. Dev.= 24 mm)

The simplest way to test whether the difference between these two descriptions of the ASF cause a significant error was to calculate the difference between them using the spatial and spectral forms in the portions of the UNB software where the spectral form was used: the RF and the FZ. These differences for degrees up to 360 (beyond which no spectral form of the ASF is prescribed), are shown in Figs. 12 and 13. The largest influence these discrepancies had on the FZ contribution is 13 mm, for the $\psi_0 = 2 / L = 20$ case (Fig 12); the largest influence on an $L = 60$ RF is 61 mm (Fig 13).

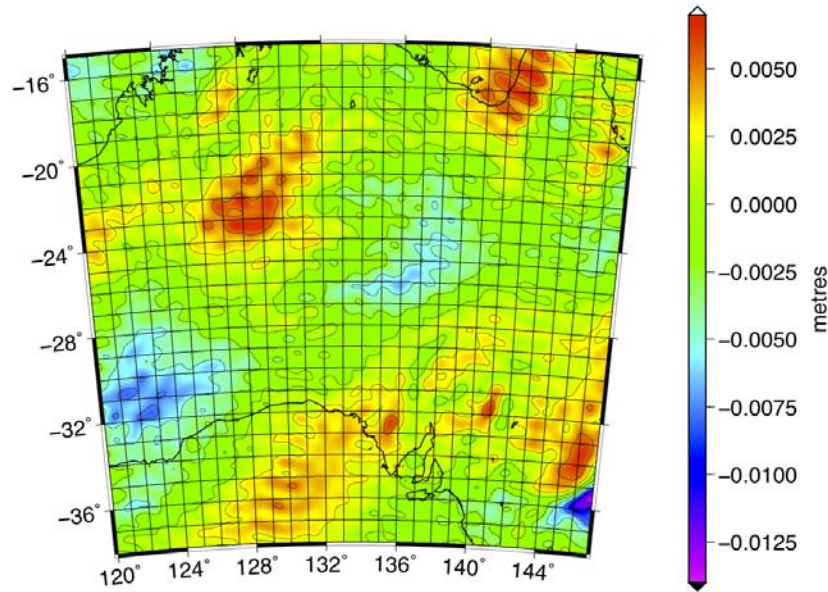


Fig 12: Effect on 2/20 FZ integration of spatial/spectral discrepancy (Min = -13 mm, Max. = 7 mm, Mean = 0 mm, Std. Dev. = 3 mm.)

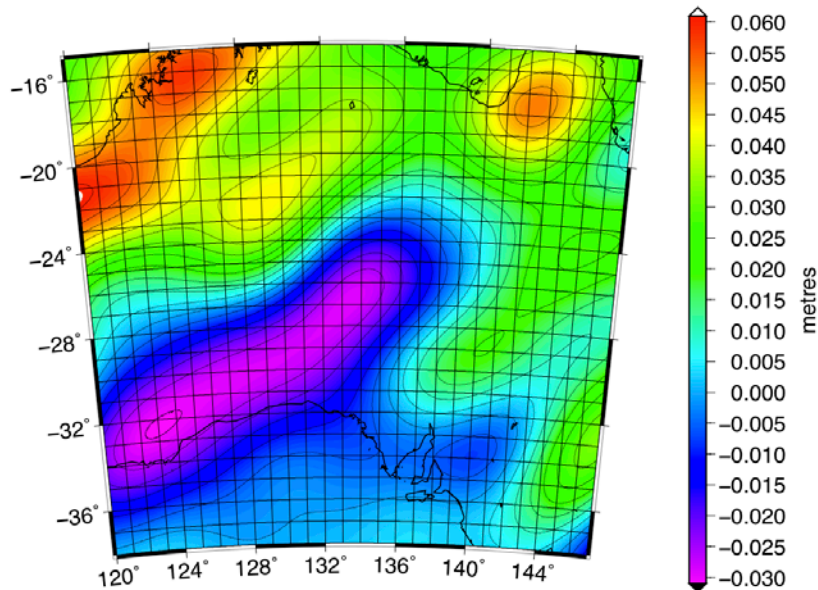


Fig 13: Effect on degree 60 reference field of spatial vs. spectral discrepancy (Min. = -31 mm, Max. = 61 mm, Mean = 12 mm, Std. Dev. = 23 mm.)

Since the source of these discrepancies is not clearly known, a correction cannot be devised and applied for them. However, even if a correction were applied, the difference between different solutions would not be affected significantly. Thus, the spatial/spectral difference can explain some - but not all - of the differences between different results of Stokes's integration. Of concern, the mean of the

discrepancies between different combinations of ψ_0 and L reach several tens of millimetres. An explanation is that it is caused by the difference between the spectral and spatial expressions of ASF as mentioned above.

We hoped that the application of the above differences and the other corrections to the ASF (Section 5) would eradicate the discrepancies among solutions with different combinations of ψ_0 and L . Unfortunately, one additional problem with the ASF that clouds the comparisons is the ASF's use of the EGM96 quasigeoid. The resulting hybrid of quasigeoid and the geoid computed by means of the Newton integration over topographical masses means that there is no real way of telling what problems are caused by this.

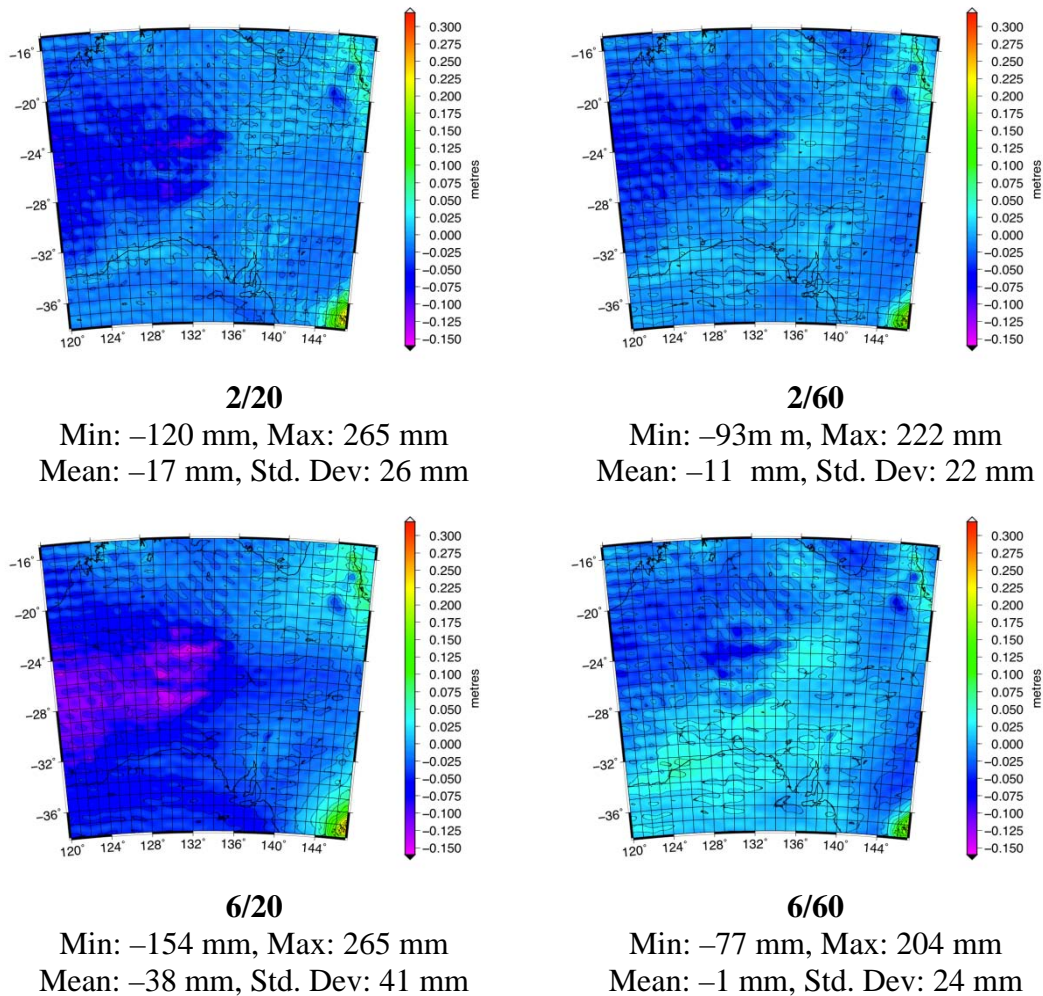


Fig 14: SHGeo geoid model minus ASF geoid, based on quasigeoid coefficients.

Using the EGM96 quasigeoid coefficients in the calculations for the four sets of ψ_0 and L , with inclusion of all the improvements from the earlier stages and the corrections to the ASF, produced the results shown in Fig.14. Yet again, the results did not show invariance with respect to the choice of ψ_0

and L . It is thus impossible to decide which EGM96 spectral expression (quasigeoid or geoid) should be used in a re-evaluation. Results from the earlier testing stages used the quasigeoid coefficients so for illustration we also used the geoid coefficients (cf., Vaniček et al. 1995) as shown in Fig.15. The difference between the two approaches, shown in Fig. 16 is - in our opinion - inconclusive.

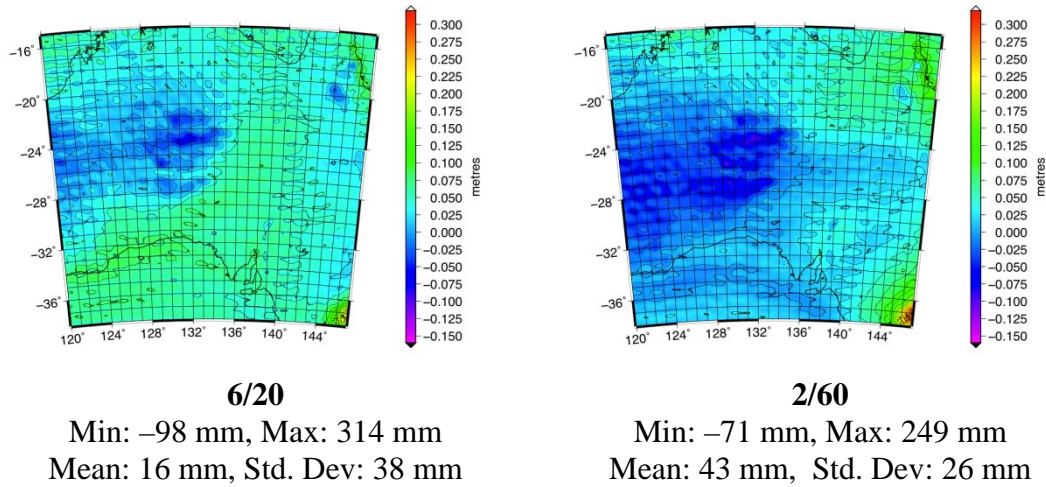


Fig 15: SHGeo geoid model minus ASF geoid, based on geoid coefficients

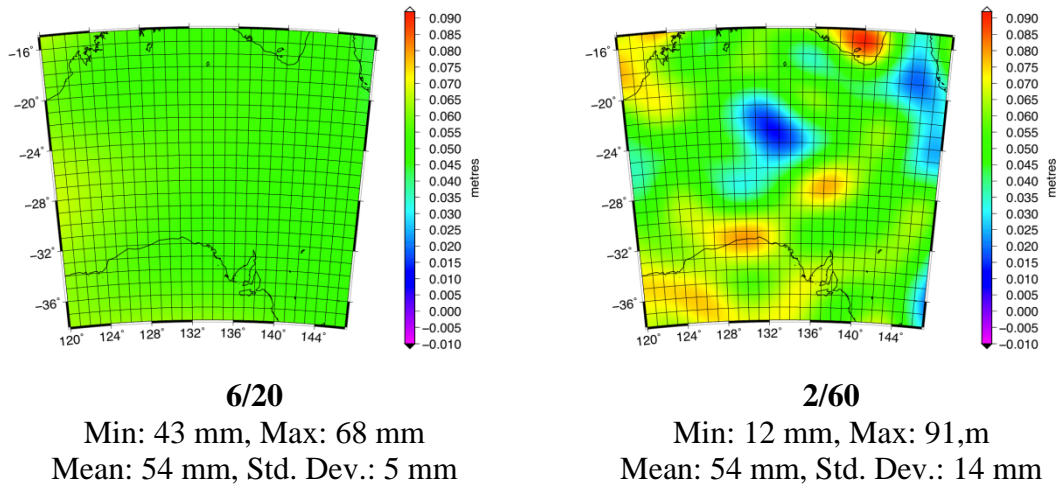


Fig 16: SHGeo geoid model based on geoid coefficients, minus SHGeo geoid model based on quasigeoid coefficients.

In parallel to the UNB-based testing of invariance for ψ_0 and L in the SHGeo software, the Curtin team used their FFT1Dmod software to investigate the same invariance. The Curtin testing also included EGM2008 to degree and order 2190. This then led to the use of whole-of-cell mean kernel values by numerical and analytical methods (Hirt 2011). This was expanded further to consider many

more geodetic kernels via Gauss-Legendre quadrature (Hirt et al. 2011). Yet again, this is a case of work that might not have been conducted had we not been experimenting with the ASF.

10. The ψ_0 and L invariance enigma

The derivation in Appendix A shows that the sum NZ + FZ should remain the same whatever values one uses for ψ_0 and L . However, the results in Sections 7.4 and 9 do show some variances. Featherstone (2003) speculated that for small values of ψ_0 and large values of L that the determination of the spheroidal Molodensky modification coefficients was numerically unstable. This did not turn out to be the case; his speculation was based only on the determinant of the matrix, whereas the condition number used here (see below) is a more powerful indicator of numerical stability.

We therefore computed the condition numbers in single precision computer algebra for the four sets of ψ_0 and L combinations used in the test presented herein (Table 2). The condition number for each case appears acceptable, so the numerical stability is not a plausible culprit for the lack of invariance shown in Sections 7.4 and 9. If the condition numbers are used, it can be shown that the matrix of normal equations for Molodensky coefficients is always well conditioned. More recently, Li and Wang (2011) determined that Alaskan geoid models computed from the spheroidal Molodensky modification coefficients were variable depending on the choices of ψ_0 and L . They, however, used the real rather than a synthetic field for their tests and such results should have been expected.

Degree L of modification	Cap radius ψ_0 (degrees)	Condition number
20	2	1.111
20	6	1.864
60	2	1.831
60	6	3.224

Table 2 Condition numbers for the combinations tested in this paper

Another, perhaps more plausible, explanation is the size of the integration domain ψ_0 in relation to the choice of L . Specifically, if a small ψ_0 is used, the NZ integration cannot sense lower frequencies that may have not been removed by the degree- L RF. This might have well been an explanation if a real gravity field were used. When a synthetic (errorless) field is involved, this explanation would not be applicable and the answer given in Appendix A must be used. At this time, we can only speculate about the sources of these variances but the errors and conceptual differences in the ASF should not be disregarded. Further research into these is clearly needed.

11. Conclusions and recommendations

Testing the Stokes-Helmert technique and the SHGeo software on the ASF has been quite an instructive experience for both the UNB and Curtin teams. The two-way process has uncovered some seemingly simple and thus easily overlooked aspects of regional geoid modelling, such as correct grid registrations and computation of the RF for a particular reference ellipsoid. Of most benefit, however, it has led to a new level of introspection over the entire geoid modelling chain, as well as to the development of an improved numerical procedure for the UNB DWC and refined numerical integration in the Curtin codes. The two-way process has also been beneficial in dispelling some previously held beliefs, as exemplified by the [incorrect] assumption that the computation of the spheroidal Molodensky coefficients was numerically unstable in some cases. Geoid computation, being a thoroughly numerical process, is replete with approximations in all the numerical processes. Therefore, in spite of the thorough testing of the software, the possibility of having additional sources of error cannot be ruled out completely.

This long-term series of investigations, involving investigators that are separated among several time zones, has probably created more questions than answers, specifically the so-called “ ψ_0 and L enigma”. It has also brought into question the exact design of a synthetic field that really should be tailored to test a particular geoid modelling method. The ASF used here suffers problems from Gibbs fringing, aliasing, omission of ellipsoidal and atmospheric terms, and the very undesirable mixture of geoid and quasigeoid concepts. Importantly, however, this exercise has shown that the construction of synthetic fields, though seemingly simple in concept, is more of a challenge. This might explain why the uptake has been rather low. It has provided valuable lessons for the construction of future synthetic fields. It is a nice example of the adage that “one size does not fit all”.

Acknowledgements

PV, MK and WEF were supported under the Australian Research Council’s (ARC) funding schemes (project numbers A00001127 and LX0989220). PV was also supported by the National Sciences and Engineering Research Council of Canada (NSERC) and, in the past, by GEOIDE, Canadian Centre of Excellence. Most figures were produced using Generic Mapping Tools (GMT; Wessel and Smith 1998) in the Lambert conical projection. We are wholeheartedly grateful to the three anonymous reviewers for their constructive critiques.

Appendix A

The derivation of the expressions used for the testing is simple. The residual geoid is given by

$$N^L(\Omega) = N^L(\Omega; \psi_0) + \delta N^L(\Omega; \psi_0) \quad (\text{A1})$$

where the NZ contribution $N^L(\Omega; \psi_0)$ is evaluated by the known Stokes's surface integral

$$N^L(\Omega; \psi_0) = \frac{R}{4\pi G} \iint_{C_{\psi_0}} \Delta g^L(\Omega') S'^L(\psi) d\Omega' \quad (\text{A2})$$

over the spherical cap of radius ψ_0 and the (modified) spheroidal Stokes's function S'^L is given by

$$S'^L(\psi; \psi_0) = S^L(\psi) - M'_L(\psi; \psi_0) = S^L(\psi) - \sum_{k=2}^L \frac{2k+1}{2} t'_k(\psi_0) P_k(\cos\psi) \quad (\text{A3})$$

where S^L is the spheroidal Stokes's function and $t'_k(\psi_0)$ for $k = 2, \dots, L$, are coefficients to be selected in some, unspecified (perhaps pre-computed by some algorithm), truncation coefficients.

It can be shown that the FZ contribution $\delta N^L(\Omega; \psi_0)$ can be determined exactly from

$$\delta N^L(\Omega; \psi_0) = \frac{R}{2G} \sum_{n=L+1}^{\infty} q_n(t'_n) \Delta g_n^L(\Omega) \quad (\text{A4})$$

where the modified Molodensky truncation q_n have to be evaluated from

$$\forall n > L: q_n(t'_n) = Q_n(\psi_0) - \frac{1}{2} \sum_{k=2}^L (2k+1) \left(\frac{2}{k-1} + t'_n(\psi_0) \right) e_{kn}(\psi_0). \quad (\text{A5})$$

Here $Q_n(\psi_0)$ and $e_{kn}(\psi_0)$ (a.k.a., $R_{kn}(\psi_0)$) are both functions of the radius of the integration cap. They can be computed using the algorithms of Paul (1973).

Equation (A4) calls for the summation to go to infinity. The coefficients q_n shrink rather rapidly with growing n and the summation can be truncated relatively early. To assure that the error in $\delta N^L(\Omega; \psi_0)$ that originates from the truncation remains smaller than 5 mm (in RMS), the summation has to be carried only to n_{max} . We have formulated the highest necessary degrees n_{max} for the spectral form of the field to be used in the evaluation of FZ for different Ψ_0 as: $n_{max} = 120$ for $\Psi_0 = 6^\circ$ (as already determined by Martinec 1993) and $n_{max} = 360$ for $\Psi_0 = 2^\circ$. The n_{max} for $\Psi_0 = 2^\circ$ was determined by us from the ranges and RMS differences between two solutions, 2/20 and 2/60, for 3 different degrees of the EGM: 120, 240 and 360. The results were: 333 mm (36 mm), 144 mm (16 mm) and 97 mm (11 mm), where the first number is the range and the number in brackets is the RMS. Our guess is that the real n_{max} is somewhere between 240 and 360.

References

- Ågren J (2004) Regional geoid determination methods for the era of satellite gravimetry: numerical investigations using synthetic Earth gravity models, doctoral dissertation, Department of Infrastructure, Royal Institute of Technology (KTH), Stockholm, Sweden, 246pp.
- Ardalan AA, Grafarend EW (2001) Ellipsoidal geoidal undulations (ellipsoidal Bruns formula): case studies. *Journal of Geodesy*, 75(9-10):544-552, doi: 10.1007/s001900100212
- Avalos-Naranjo D, Sosa-Gaytan M and Munoz-Abundes R (2011) Aspectos metodológicos de las cartas y modelos gravimétricos 2010, Instituto Nacional de Estadística y Geografía, Aguascalientes, Mexico, October 2011.
- Baran I, Kuhn M, Claessens SJ, Featherstone WF, Holmes SA, Vaníček P (2006) A synthetic Earth gravity model designed specifically for testing regional gravimetric geoid determination algorithms, *Journal of Geodesy* 80(1):1-16, doi: 10.1007/s00190-005-0002-z.
- Blitzkow D., de Matos ACOC, Fairhead JD, Pacino MC, Lobianco MCB and Campos IO (2012). "The progress of the geoid model computation for South America under GRACE and EGM2008 models." *Geodesy for Planet Earth, International Association of Geodesy Symposia*, Vol. 136, pp. 893-899, doi:10.1007/978-3-642-20338-1_112.
- Claessens SJ (2003) A synthetic Earth model, Delft University Press, Delft, The Netherlands, 61pp
- Ellmann A, Vaníček P (2007) UNB application of Stokes-Helmert's approach to geoid computation, *Journal of Geodynamics* 43(2): 200-213, doi: 10.1016/j.jog.2006.09.019.
- Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, Kobrick M, Paller M, Rodriguez E, Roth L, Seal D, Shaffer S, Shimada J, Umland J, Werner M, Oskin M, Burbank D, and Alsdorf D (2007), The Shuttle Radar Topography Mission, *Reviews of Geophysics* 45, RG2004, doi: 10.1029/2005RG000183.
- Featherstone WE (2002) Tests of two forms of Stokes's integral using a synthetic gravity field based on spherical harmonics. In: Grafarend EW, Krumm FW, Schwarze VS (eds), *Geodesy - The Challenge for the Third Millennium*, Springer, Berlin Heidelberg, 163-171.
- Featherstone WE (2003) Software for computing five existing types of deterministically modified integration kernel for gravimetric geoid determination, *Computers & Geosciences* 29(2): 183-193, doi: 10.1016/S0098-3004(02)00074-2.
- Featherstone, W.E. and M. Kuhn (2006) Height systems and vertical datums: a review in the Australian context, *Journal of Spatial Science* 51(1): 21-42, doi: 10.1080/14498596.2006.9635062.
- Featherstone WE, Kirby JF, Kearsley AHW, Gilliland JR, Johnston GM, Steed J, Forsberg R, Sideris MG (2001) The AUSGeoid98 geoid model of Australia: data treatment, computations and comparisons with GPS-levelling data, *Journal of Geodesy* 75(5-6): 313-330, doi: 10.1007/s001900100177.
- Featherstone, W.E., S.A. Holmes, J.F. Kirby and M. Kuhn (2004) Comparison of remove-compute-restore and University of New Brunswick techniques to geoid determination over Australia, and inclusion of Wiener-type filters in reference field contribution, *Journal of Surveying Engineering* 130(1): 40-47, doi: 10.1061/~ASCE!0733-9453~2004!130:1~40!.
- Fellner JJ, M Kuhn, WE Featherstone (2011) Development of a Synthetic Earth Gravity Model by 3D mass optimisation based on forward modelling', *Earth Planets Space* 63: 1-8.
- Filmer MS, Featherstone WE, Kuhn M (2010) The effect of EGM2008-based normal, normal-orthometric and Helmert orthometric height systems on the Australian levelling network, *Journal of Geodesy* 84(8): 501-513, doi: 10.1007/s00190-010-0388-0.
- Haagmans R (2000) A synthetic Earth model for use in geodesy, *Journal of Geodesy* 74(7-8): 503-511, doi: 10.1007/s001900000112.

- Hadamard J (1923) Lectures on the Cauchy Problem in Linear Partial Differential Equations, Yale University Press, New Haven.
- Hastings DA, Dunbar PK (1998) Development & Assessment of the Global Land One-km Base Elevation Digital Elevation Model (GLOBE), ISPRS Archives, 32(4): 218-221
<http://www.ngdc.noaa.gov/mgg/topo/globe.html>
- Heiskanen WH, Moritz H (1967) Physical geodesy. W.H. Freeman and Co., San Francisco.
- Helmert FR (1884) Die mathematischen und physicalischen Theorien der höheren Geodäsie, Vol. 2, B.G. Teubner, Leipzig (reprinted in 1962 by Minerva GMBH, Frankfurt/Main).
- Hirt C (2011) Mean kernels to improve gravimetric geoid determination based on modified Stokes's integration, Computers & Geosciences 37(11):1836-1842, doi: 10.1016/j.cageo.2011.01.005.
- Hirt C, Featherstone WE, Claessens SJ (2011) On the accurate numerical evaluation of geodetic convolution integrals, Journal of Geodesy 85(8): 519-538, doi: 10.1007/s00190-011-0451-5.
- Holota P Nesvadba O (2008) Model Refinements and Numerical Solutions of Weakly Formulated Boundary-Value Problems in Physical Geodesy. In: Xu P, Liu J, Dermanis A (Eds.): VI Hotine-Marussi Symposium of Theoretical and Computational Geodesy. Wuhan, China, 29 May - 2 June, 2006. International Association of Geodesy Symposia, Volume 132, Springer, Berlin, Heidelberg, 320-326.
- Hörmander L (1976) The boundary problems of physical geodesy, Archive for Rational Mechanics and Analysis 62(1): 1-52, doi: 10.1007/BF00251855.
- Janak J, Pasteka R, Zahorec P, Loviska Z (2006) Terrain correction in extremely disturbed terrain. Contributions to Geophysics and Geodesy, 36:, 41-52.
- Jekeli C (1988) The exact transformation between ellipsoidal and spherical harmonic expansions. manuscripta geodaetica, 13:106-113
- Kadir MA, Fashir HH, Omar K (1999) A regional gravimetric co-geoid over South East Asia, Geomatics Research Australasia 71: 37-56.
- Kingdon R, Vaníček P (2011) Poisson downward continuation solution by the Jacobi method, Journal of Geodetic Science 1(1): 74-81, doi: 10.2478/v10156-010-0009-0.
- Kingdon R, Vaníček P, Santos M, Ellmann A, Tenzer R (2005) Toward an improved height system for Canada, Geomatica 59(3): 241-249.
- Kuhn M (2000) Geoidbestimmung unter Verwendung verschiedener Dichtehypothesen. Deutsche Geodätische Kommission, Reihe C, no. 520, München
- Kuhn M (2003) Geoid determination with density hypotheses from isostatic models and geological information, Journal of Geodesy 77(1-2): 50-65, doi: 10.1007/s00190-002-0297-y
- Kuhn M, WE Featherstone (2005) Construction of a synthetic Earth gravity model by forward gravity modelling, in: Sansò, F. (ed) A Window on the Future of Geodesy, Springer, Berlin, Heidelberg, 350-355.
- Kuhn M, WE Featherstone, JF Kirby (2009) Complete spherical Bouguer gravity anomalies over Australia. Australian Journal of Earth Sciences, 56(2): 213-223, doi: 10.1080/08120090802547041.
- Li X, Wang Y-M (2011) Comparisons of geoid models over Alaska computed with different Stokes's kernel modifications, Journal of Geodetic Science 1(2): 136-142, doi: 10.2478/v10156-010-0016-1.
- 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 EG, Rapp RH, Olson TR (1998) The development of the joint NASA GSFC and NIMA Geopotential Model EGM96, NASA/TP-1998-206861, NASA Goddard Space Flight Center, Greenbelt Maryland, USA.
- Martinec, Z., (1993) Effect of lateral density variations of topographical masses in view of improving geoid model accuracy over Canada. Energy, Mines and Resources, (now Geomatics Canada), Geodetic Survey Division, Ottawa, Contract Report, No. 93-002, 112 pp.

- Martinec Z (1996) Stability investigations of a discrete downward continuation problem for geoid determination in the Canadian Rocky Mountains, *Journal of Geodesy* 70(11): 805-828, doi: 10.1007/BF00867158.
- Martinec Z (1998) Boundary-Value Problems for Gravimetric Determination of a Precise Geoid. *Lecture Notes in Earth Sciences*, Springer, Berlin, Heidelberg, 223 pp.
- Martinec Z, Vaníček P (1994a) The indirect effect of Stokes-Helmert's technique for a spherical approximation of the geoid. *manuscripta geodaetica* 19(2): 213-219.
- Martinec Z, Vaníček P (1994b) Direct topographical effect of Helmert's condensation for a spherical approximation of the geoid. *manuscripta geodaetica* 19(3): 257-268.
- Molodensky MS (1945) *Fundamental Problems of Geodetic Geavimetry* (in Russian) Trudy TsNIIGAIK 42, Geodezizdat, Moscow.
- Molodensky MS, Yeremeev VF and Yurkina MI (1960) *Methods for Study of the External Gravitational Field and Figure of the Earth*. TRUDY Ts NIIGAIK, 131, Geodezizdat, Moscow. English transl.: Israel Program for Scientific Translation, 248 pp, Jerusalem 1962.
- Moritz H (1980) Geodetic Reference System 1980, *Bulletin Géodésique* 54(3): 187-192, doi: 10.1007/s001900050278.
- Novák P, Vaníček P, Véronneau M, Holmes S, Featherstone WE (2001) On the accuracy of modified Stokes's integration in high-frequency gravimetric geoid determination, *Journal of Geodesy* 74(11): 644-654, doi: 10.1007/s001900000126.
- Pail R (2000) *Synthetic global gravity model for planetary bodies and applications in satellite gravity gradiometry*, PhD thesis, Technical University of Graz, Graz, Austria, 133pp.
- Paul MK (1973) A method of evaluating the truncation error coefficients for geoidal height, *Bulletin Géodésique* 47(4): 413-425, doi: 10.1007/BF02521951.
- Pavlis NK, Holmes SA, Kenyon SC, Factor JF (2012), The development and evaluation of Earth Gravitational Model (EGM2008), *Journal of Geophysical Research* 117, B04406, doi: 10.1029/2011JB008916.
- Rapp RH (1997) Use of potential coefficient models for geoid undulation determinations using a spherical harmonic representation of the height anomaly/geoid undulation difference, *Journal of Geodesy* 71(5): 282-289, doi: 10.1007/s001900050096.
- Santos MC, Vaníček P, Featherstone W, Kingdon R, Ellmann A, Martin BA, Kuhn M, Tenzer R (2006) The relation between the rigorous and Helmert's definitions of orthometric heights, *Journal of Geodesy* 80(12): 691-704, doi: 10.1007/s00190-006-0086-0.
- Sjöberg LE (2003) Ellipsoidal corrections to order e^2 of geopotential coefficients and Stokes' formula. *Journal of Geodesy*, 77(3): 139-147, doi: 10.1007/s00190-003-0321-x.
- Smith DA (1998) There is no such thing as "The" EGM96 geoid: Subtle points on the use of a global geopotential model, *International Geoid Service Bulletin* 8: 17-28.
- Stokes GG (1849) On the variation of gravity at the surface of the Earth, *Transactions of the Cambridge Philosophical Society* 8: 672-695.
- Sun W, Vaníček P (1998) On some problems of the downward continuation of the 5' x 5' mean Helmert gravity disturbance, *Journal of Geodesy* 72(7-8): 11-420, doi: 10.1007/s001900050180.
- Tenzer R, Vaníček P, Santos MC, Featherstone WE, Kuhn M (2005) The rigorous determination of orthometric heights, *Journal of Geodesy* 79(1-3): 82-92, doi: 10.1007/s00190-005-0445-2.
- Tikhonov A N, Arsenin VY (1977) *Solutions of Ill-posed Problems*, John Wiley & Sons, New York.
- Tziavos IN (1996) Comparisons of spectral techniques for geoid computations over large regions, *Journal of Geodesy* 70(6): 357-373, doi: 10.1007/s001900050027.

- Valty P, Duquenne H, Panet I (2012) Auvergne dataset: testing several geoid computation methods, in Kenyon S, Pacino MC, Marti U (eds.), *Geodesy for Planet Earth*, Springer, Berlin, Heidelberg, 465-472.
- Vaniček P, Kleusberg A (1987) The Canadian geoid - Stokesian approach, *manuscripta geodaetica* 12(2): 86-98.
- Vaniček P, Sjöberg LE (1991) Reformulation of Stokes's theory for higher than second-degree reference field and modification of integration kernels, *Journal of Geophysical Research* 96(B4): 6529-6339, doi: 10.1029/90JB02782.
- Vaniček P, Martinec Z (1994) Stokes-Helmert scheme for the evaluation of a precise geoid, *manuscripta geodaetica* 19(2): 119-128.
- Vaniček P, Ong P, Zhang C (1991) New gravimetric geoid for Canada: the "UNB90" solution. In Rapp RH, Sanso F (Eds.): *Determination of the Geoid: Present and Future*. Springer, Berlin, Heidelberg, 214-219.
- Vaniček P, Najafi M, Martinec Z, Harrie L, Sjöberg LE (1995) Higher-degree reference field in the generalised Stokes-Helmert scheme for geoid computation, *Journal of Geodesy* 70(3): 176-182, doi: 10.1007/BF00943693.
- Vaniček P, Sun W, Ong P, Martinec Z, Najafi M, Vajda P, Horst B (1996) Downward continuation of Helmert's gravity, *Journal of Geodesy* 71(1): 21-34, doi: 10.1007/s001900050072.
- Vaniček P, Huang J, Novák P, Pagiatakis SD, Véronneau M, Martinec Z, Featherstone WE (1999) Determination of the boundary values for the Stokes-Helmert problem, *Journal of Geodesy* 73(4): 160-192, doi: 10.1007/s001900050235.
- Vaniček P, Novák P and Martinec Z (2001) Geoid, topography, and the Bouguer plate or shell. *Journal of Geodesy* 75 (4): 210-215, doi: 10.1007/s001900100165.
- Vaniček P, Kingdon R, Santos M (2012) Geoid versus quasi-geoid: a case of physics versus geometry, *Contributions to Geophysics and Geodesy* 42(1): 101-119.
- Véronneau M, Huang J (2011) The reference surface and maintenance of the up-coming height system in Canada, *American Geophysical Union, Fall Meeting, San Francisco, USA*.
- Vermeer M (1998) The geoid as a product. *Proceedings, Second Continental Workshop on the Geoid in Europe, March 10-14, 1998, Report 98:4, Finnish Geodetic Institute, 63-69, Masala*.
- Wessel P, Smith WHF (1998) New, improved version of Generic Mapping Tools released, *EOS: Transactions of the American Geophysical Union* 79(47): 579.
- Wong J (2002) On Picard Criterion and the Well-Posed Nature of Harmonic Downward Continuation, *Technical Report 213, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, Canada*.