Journal of Geodesy (2020) 94:52
https://doi.org/10.1007/s00190-020-01379-3

ORIGINAL ARTICLE

Towards an International Height Reference System: insights from the Colorado geoid experiment using AUSGeoid computation methods

S. J. Claessens[1] · **M. S. Filmer**[1]

Received: 20 December 2019 / Accepted: 26 April 2020 / Published online: 11 May 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

We apply the AUSGeoid data processing and computation methodologies to data provided for the International Height Reference System (IHRS) Colorado experiment as part of the International Association of Geodesy Joint Working Groups 0.1.2 and 2.2.2. This experiment is undertaken to test a range of different geoid computation methods from international research groups with a view to standardising these methods to form a set of conventions that can be established as an IHRS. The IHRS can realise an International Height Reference Frame to be used to study physical changes on and within the Earth. The Colorado experiment study site is much more mountainous (maximum height 4401 m) than the mostly flat Australian continent (maximum height 2228 m), and the available data over Colorado are different from Australian data (e.g. much more extensive airborne gravity coverage). Hence, we have tested and applied several modifications to the AUSGeoid approach, which had been tailored to the Australian situation. This includes different methods for the computation of terrain corrections, the gridding of terrestrial gravity data, the treatment of long-wavelength errors in the gravity anomaly grid and the combination of terrestrial and airborne data. A new method that has not previously been tested is the application of a spherical harmonic high-pass filter to residual anomalies. The results indicate that the AUSGeoid methods can successfully be used to compute a high accuracy geoid in challenging mountainous conditions. Modifications to the AUSGeoid approach lead to root-mean-square differences between geoid models up to ~0.028 m and agreement with GNSS-levelling data to ~0.044 m, but the benefits of these modifications cannot be rigorously assessed due to the limitation of the GNSS-levelling accuracy over the computation area.

Keywords Geoid · Quasigeoid · Stokes · GNSS-levelling · Orthometric heights

1 Introduction

A key component in defining an International Height Reference System (IHRS) is establishing a set of standards and conventions that can be applied consistently on a global basis (Ihde et al. 2017). An area where there are different conventions applied is regional geoid computation by different research groups around the world. For example, these groups may use different methods to reduce gravity anomalies, and to different surfaces, with some computing geoid models and others quasigeoids using Molodensky's principles. In an effort to determine the differences between these computation methods and software, the International Association of Geodesy (IAG) resolved to establish a collaborative effort between IAG Joint Working Groups 0.1.2 and 2.2.2 which has become the IHRS Colorado experiment. The objective of the experiment is for numerous groups to use their computation methods and software to compute limited area geoid models in challenging terrain and compare and analyse the differences. From this analysis, the best performing methodologies can be adopted as the standards for the formulation of the IHRS, leading to an International Height Reference Frame (IHRF) as the realisation of a global unified physical height reference based on gravity potential.

The experiment required all groups to compute height anomalies, geoid heights, and potential values over a 5° × 8° area (35°–40°N, 102°–110°W) covering Colorado and adjacent US states. Our contribution to this experiment is to apply,

✉ S. J. Claessens
S.Claessens@curtin.edu.au

[1] School of Earth and Planetary Sciences and The Institute for Geoscience Research, Curtin University, GPO Box U1987, Perth, WA 6845, Australia

as much as possible, the approach that has recently been used in the computation of AUSGeoid2020 (Featherstone et al. 2018) so that the results from this approach applied to the experiment can be compared to the results from other international groups.

AUSGeoid2020 consists of two components: a gravimetric component called the Australian gravimetric quasigeoid 2017 (AGQG2017; Featherstone et al. 2018) and a geometric component (Brown et al. 2018) that fits this gravimetric quasigeoid to the Australian Height Datum (AHD71; Roelse et al. 1971). The geometric component is not of interest for the IHRS Colorado experiment, so here we only use the approach used in the computation of AGQG2017, but refer to it as 'the AUSGeoid approach'. This approach is based on remove-compute-restore (RCR) quasigeoid computation with deterministically modified Stokes integration, where the modification parameters are determined from parameter sweeps and comparison to GNSS-levelling data. Earlier versions of this approach were used for the computation of AUSGeoid09 (Featherstone et al. 2011) and NZGeoid09 (Claessens et al. 2011).

However, due to differences in terrain, data, and product requirements, some changes to the AUSGeoid approach were required. In this paper, we describe how these differences affect the optimal geoid computation strategy. This includes comparison of alternative methods for the computation of terrain corrections, the gridding of terrestrial gravity data, the treatment of long-wavelength errors in the gravity anomaly grid, and the combination of terrestrial and airborne data.

A first and obvious difference between the Colorado computation area and Australia is the terrain. The Colorado terrain is high and mountainous [mean elevation of 2013 m and maximum elevation (Mount Elbert) of 4401 m]. Australia, on the other hand, is the lowest and flattest of continents (e.g. Sandiford and Quigley 2009) with vast plains and a maximum elevation of only 2228 m (Mount Kosciuszko), despite its much larger area. This affects the computation strategy employed for terrain corrections and gravity gridding (Sects. 2.1 and 2.2).

A second and very important difference is in the available data. As very little airborne gravity data are available over Australia, the AUSGeoid approach does not include sophisticated methods for the incorporation of airborne gravity data. However, much of the Colorado area is covered by airborne gravity observations from the GRAV-D project (Smith 2007). We have therefore computed two types of geoid models: ones that exclude all airborne data and ones that include the airborne data. For the inclusion of the airborne gravity data, the NZGeoid2017 approach (McCubbine et al. 2018) was largely followed (Sects. 2.3 and 2.4). In this approach, terrestrial and airborne data are combined using 3D least-squares collocation (LSC) with a planar logarithmic covariance function prior to Stokes integration.

Another difference in terrain and data availability is that in the AUSGeoid computation approach, marine gravity anomalies from satellite altimetry and the challenges of gravity modelling in the coastal zone (e.g. Wu et al. 2019) play an important role. Since the Colorado area does not involve any marine zones, this challenge does not apply here.

A final issue that necessitated an adaptation to the AUSGeoid approach was the product requirements. Firstly, since the AHD71 uses a normal-orthometric height system (e.g. Filmer et al. 2010), the gravimetric component of AUSGeoid2020 is a quasigeoid model. However, heights with respect to the North American Vertical Datum of 1988 (NAVD88; Zilkoski et al. 1992), which are used in the Colorado area, are given in Helmert orthometric heights (e.g. Jekeli 2000). To be consistent with Helmert orthometric heights, the basic requirements for the IHRS Colorado experiment stipulated the use of a simple, approximate geoid-to-quasigeoid separation term (Sect. 2.6). Note, however, that the use of geopotential numbers or normal heights, not orthometric heights, is recommended for the IHRS (Ihde et al. 2017). As per the experiment requirements, we have computed a quasigeoid model, geoid model and potential value model.

The different height system also required adaptation of the GNSS-levelling comparison method. We have tested two different methods for the interpolation of geoid undulations to benchmarks (Sect. 2.7).

The data and methods employed in this study are detailed in Sect. 2. Results excluding airborne gravity data are presented in Sect. 3. This includes the effects of different terrain correction computation methods, different data gridding methods, different methods for treatment of long-wavelength errors in the gravity anomaly grid, Stokes integration parameter optimisation and different interpolation methods in the GNSS-levelling comparisons. In Sect. 4, the results including airborne data are presented, as well as a comparison between the solutions with and without airborne data. A discussion of the results and conclusions are presented in Sects. 5 and 6.

2 Data and methods

2.1 Terrain corrections

A $3'' \times 3''$ DEM over Colorado was provided by United States National Geodetic Survey (NGS), based on SRTM v4.1 (Jarvis et al. 2008), covering an area that exceeds the computation area by 2° in all directions (33°–42°N, 100°–112°W). It was used to compute planar terrain corrections over the test area using two different algorithms: (1) the algorithm used in AUSGeoid2020 (McCubbine et al. 2017) and (2) a more recently derived algorithm (Goyal et al. 2019) developed for the Indian geoid project. Both methods use the fast Fourier

transform that relies on applying a binomial expansion to the integrand. As the terrain over most of Australia is relatively flat, the algorithm by McCubbine et al. (2017) considers only the first-order term of the expansion. However, in the more mountainous terrain of Colorado, the first-order approximation is not sufficiently accurate. Therefore, the method of Goyal et al. (2019) was selected for this study. The terrain correction is herein calculated using a binomial expansion

$$\mathrm{TC} = G\rho \int_{x_1}^{x_2}\int_{y_1}^{y_2}\left[\frac{1}{l}\sum_{k=1}^{\infty}\left(\frac{k-\frac{1}{2}}{k}\right)\left(-\frac{\Delta z}{l}\right)^2\right]\mathrm{d}x\mathrm{d}y \qquad (1)$$

where G is the universal gravitational constant, ρ is the terrain density, l is the horizontal distance, and Δz is the vertical distance between the computation point and the integration point. The binomial expansion was used up to sixth order ($k = 6$). Brute-force computation not relying on the binomial expansion was used at locations with gradients exceeding 45°, where the expansion may diverge.

2.2 Terrestrial gravity data

A total of 59,303 terrestrial gravity observations were made available by NGS (Fig. 1). This equates to an average of 0.15 observations/km^2, which, for comparison, is less than the 0.23 terrestrial gravity observations/km^2 used in the computation of AUSGeoid2020 (Featherstone et al. 2018).

Atmospheric corrections according to Moritz (2000) were applied to these observations. Then, the gravity observations were transformed into Molodensky-type free-air anomalies, where Somigliana–Pizzetti normal gravity was computed rigorously using GRS80 reference ellipsoid parameters

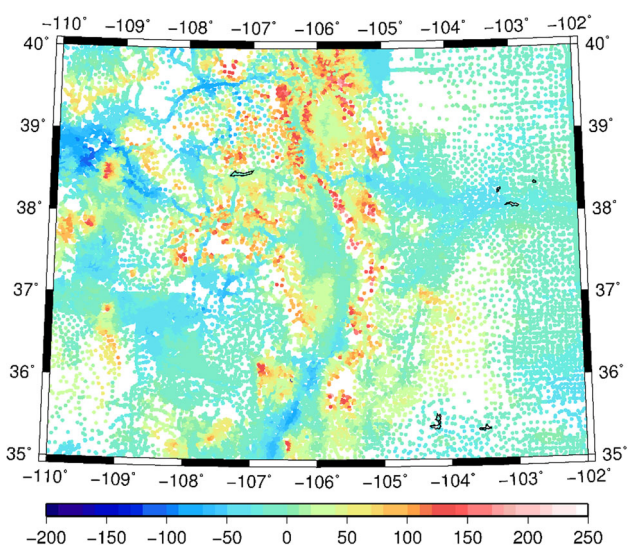

Fig. 1 Free-air anomalies showing the coverage of terrestrial gravity observations (units in mGal)

(Moritz 2000). In this process, the ellipsoidal height of the telluroid was approximated by the provided orthometric heights of the gravity observations. Simple planar Bouguer gravity anomalies were subsequently computed by applying the Bouguer plate correction with a constant topographic mass density of 2670 kg m^{-3}. (This mass density value was stipulated in the IHRS Colorado experiment requirements.)

For the gridding of the gravity data, two different methods were used, so as to test the suitability of the method specifically adapted for the AUSGeoid computation against a conventional method in the challenging Colorado terrain. The first method is a straightforward gridding of refined Bouguer anomalies and is herein referred to as the simple gridding method. In this method, terrain corrections were bi-cubically interpolated to the gravity observation locations and applied to the simple planar Bouguer anomalies to obtain refined Bouguer anomalies (e.g. Hackney and Featherstone 2003). These were then gridded onto a $1' \times 1'$ grid using the tensioned spline (Smith and Wessel 1990) routine in the Generic Mapping Tools (GMT; Wessel et al. 2013), with a tension factor of 0.25. Subsequently, the gridded refined Bouguer anomalies were transformed into Faye anomalies by addition of the simple Bouguer correction.

The second method is that of Featherstone and Kirby (2000), which is referred to as the reconstruction method. It was, for example, used in the computation of AUSGeoid09 (Featherstone et al. 2011) and AUSGeoid2020 (Featherstone et al. 2018). In this method, the simple planar Bouguer anomalies were bi-cubically interpolated to a regular $3'' \times 3''$ grid using the GMT tensioned spline routine (Smith and Wessel 1990; Wessel et al. 2013), again with a tension factor of 0.25. Note that terrain corrections were not applied to the point anomalies before gridding. While application of the terrain corrections, resulting in refined Bouguer anomalies, would have provided a smoother surface for gridding, the method of Featherstone and Kirby (2000) allows for the addition of the mean terrain correction afterwards. This reduces aliasing which may arise when more gravity observations are taken in lowland areas (Featherstone and Kirby 2000). This method has been shown to be beneficial over Australia, but has until now not been tested over rougher terrain. Gridded Molodensky-type free-air gravity anomalies on the topography were generated on a $3'' \times 3''$ grid, to which planar terrain corrections (Sect. 2.1) were added to provide Faye anomalies on a $3'' \times 3''$ grid, which were then block-averaged to a $1' \times 1'$ grid.

The $1' \times 1'$ grids resulting from the simple gridding method and the reconstruction method were each used to compute two "terrestrial-only" geoid models for Colorado, without inclusion of available airborne gravity data, which are compared and evaluated (Sect. 3.2).

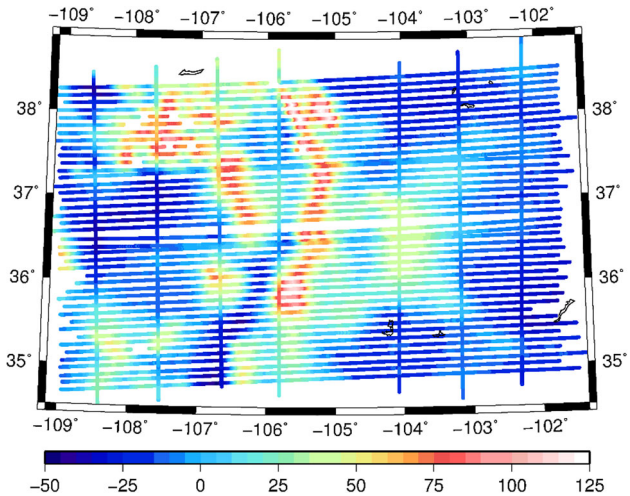

Fig. 2 Free-air anomalies showing the coverage of airborne gravity observations (units in mGal)

2.3 Airborne gravity data

A total of 283,716 debiased GRAV-D (Smith 2007) airborne gravity observations, resampled to a 1 Hz observation frequency, were provided by NGS. The data consist of 56 flight lines with elevations ranging between 5000 and 8000 m. Figure 2 shows the spatial coverage of the data. We have applied the approach used for the computation of NZGeoid2017 (McCubbine et al. 2018) to process the airborne gravity data, which also includes a different procedure to process the terrestrial gravity data.

First, we converted the airborne gravity observations into Molodensky-type free-air gravity anomalies. Each flight line was then low-pass filtered with a 1D Gaussian filter to remove high-frequency noise. McCubbine et al. (2018) use a filter length of 120 s to achieve an along-track spatial resolution of ~ 8 km. However, the airborne gravity over Colorado was flown at higher speed (average speed of ~ 200 knots compared to ~ 130 knots over New Zealand), and therefore, a shorter filter length of 80 s was used here instead. This resulted in a similar spatial resolution of ~ 8 km. We cautiously estimate the accuracy of the airborne data at ± 3 mGal, based on the RMS error from crossover analysis of the data of 2.32 mGal (GRAV-D Team 2017) and analysis of other GRAV-D airborne gravity data over mountainous terrain by Huang et al. (2017).

2.4 Combination of terrestrial and airborne gravity data

Airborne gravity data were downward continued to the topography using 3D least-squares collocation (LSC) with planar logarithmic covariance function (Forsberg 1987). Residual gravity anomalies Δg_{res} were used as input, and these were computed using

$$\Delta g_{\mathrm{res}} = \Delta g - \Delta g_{\mathrm{GGM}} - \delta g_{\mathrm{TC}} \tag{2}$$

where Δg are the Molodensky-type free-air gravity anomalies, Δg_{GGM} are the long-wavelength gravity anomalies from the tide-free satellite-only Global Gravity Model (GGM) GO_CONS_GCF_2_DIR_R6 (Förste et al. 2019) until degree and order (d/o) 300, and δg_{TC} is the gravitational effect of the topographic masses determined from a long-wavelength DEM. These residual gravity anomalies were computed for both the (non-gridded) terrestrial and airborne data sets.

Two different methods for the computation of the topographic correction δg_{TC} were tested, with the aim of assessing their adaptability to this mountainous test area. The difference between these methods is the technique employed to generate the long-wavelength DEM: (1) using a 2D Gaussian filter and (2) using a spherical harmonic box filter. In method 1, a 2D Gaussian filter with a standard deviation of 18′ was used, following Forsberg et al. (2014) and McCubbine et al. (2018). In method 2, a spherical harmonic box filter was used as an alternative, to test the influence of the long-wavelength topography on the geoid. The 3″ × 3″ DEM was first block-averaged to 2′ × 2′ resolution and subsequently extended to a global 2′ × 2′ grid, where all cells outside the study area were set equal to the mean DEM height over the study area. This global grid was expanded into a surface spherical harmonic series up to d/o 2699 using the *SHTools* software (Wieczorek and Meschede 2018), applying the algorithm of Driscoll and Healy (1994). The long-wavelength DEM was then created by spherical harmonic synthesis of this series up to d/o 300.

Rectangular prism integration (Nagy et al. 2000) was used to compute the gravitational effect of topographic masses from the long-wavelength DEM. The gravitational effect on the airborne data was filtered with the same 1D Gaussian filter as the observations, so that short-wavelength terrain effects do not produce noise in the airborne data (e.g. Forsberg 2002).

3D LSC was applied to simultaneously downward continue the airborne gravity data to the Earth surface and combine it with the terrestrial gravity data, directly creating a 1′ × 1′ grid of residual gravity anomalies on the surface of the topography. For the LSC process, the planar logarithmic covariance function of Forsberg (1987) was used

$$C\left(\Delta g_{\mathrm{res},1}, \Delta g_{\mathrm{res},2}\right) = -f \sum_{k=0}^{3} \alpha_k \log\left(D_k + \sqrt{r^2 + (D_k + h_1 + h_2)^2}\right) \tag{3}$$

where

$$f = C_0 \log\left(\frac{(D+T)^3(D+3T)}{D(D+2T)^3}\right) \tag{4}$$

Here, C_0, D and T are three constants defining the covariance function, which were determined empirically as $C_0 = 407.86\,\text{mGal}^2$, $D = 16.6\,\text{km}$ and $T = 46.2\,\text{km}$, with $\alpha_k = [1, -3, 3, -1]$, $D_k = D + kT$, r is the planar distance between any two points, $\Delta g_{\text{res},1}$ and $\Delta g_{\text{res},2}$ are residual gravity anomalies at those two points, and h_1 and h_2 are the heights of the two points.

The Δg_{GGM} and δg_{TC}. components are then added back to the gridded residual gravity anomalies, and the anomalies are subsequently transformed into Faye anomalies. The $1' \times 1'$ grid of Faye anomalies forms the input for the remove-compute-restore (quasi)geoid computation in the AUSGeoid approach.

2.5 Stokes integration

The remove-compute-restore (RCR) technique for quasi-geoid computation is used in the AUSGeoid approach. The GGM used here is Cnm_refB_v050317a_s2-2190zt_4, a preliminary zero-tide version of EGM2020 (Barnes et al. 2015) to spherical harmonic d/o $n_{\text{max}} = 2190$. The model was converted to a tide-free model by applying a correction to the C_{20}-term of the model. The correction was determined by subtracting the C_{20}-term of the zero-tide version of EGM2008 from the C_{20}-term of the tide-free version of EGM2008 (Pavlis et al. 2012). In this way, we have used the same tide conversion as was used in EGM2008, which is based on the equations of Rapp (1989) and Ekman (1989).

Ellipsoidal gravity anomalies from the GGM were rigorously synthesised on a $1' \times 1'$ grid at the surface of the topography, as per Featherstone et al. (2018). However, a difference with the procedure described in Featherstone et al. (2018) is that the spherical harmonic synthesis was performed using the *harmonic_synth* software (provided by the EGM2008 development team) rather than *isGrafLab* (Bucha and Janák 2014). While the rigorous computation method in *harmonic_synth* is computationally less efficient, the procedure to synthesise ellipsoidal gravity anomalies is more straightforward, and the computational burden was not prohibitive as the Colorado computation area is considerably smaller than the continent of Australia.

Subtracting the GGM anomalies from the gridded Faye anomalies gives residual gravity anomalies. To avoid confusion with the residual gravity anomalies from Eq. (2), we herein call these the residual Faye anomalies. These residual Faye anomalies are typically not completely free of signal content below d/o n_{max}. This is primarily due to errors in both the GGM and in the gridded Faye gravity anomalies. If there are long-wavelength errors in the Faye anomalies, this will lead to long-wavelength errors in the computed quasi-geoid after the restore step. To overcome this issue, instead of relying solely on Stokes kernel modification to reduce long-wavelength errors (e.g. Vaníček and Featherstone 1998), we have here also tested a new procedure in which a spherical harmonic high-pass filter is applied to the residual Faye anomalies before Stokes integration. This is a procedure not previously used in the computation of AUSGeoid models, but is implemented for the Colorado experiment because of the apparent long-wavelength errors in the residual Faye anomalies.

The spherical harmonic high-pass filter was implemented as follows. The $1' \times 1'$ grid of residual Faye anomalies was first block-averaged to $2' \times 2'$ resolution and subsequently extended to a global $2' \times 2'$ grid, where all cells outside the study area were set equal to the mean residual Faye anomaly over the computation area. The block-averaging to $2' \times 2'$ resolution is performed, because this allows exact spherical harmonic analysis to d/o 2699 using the algorithm of Driscoll and Healy (1994) without use of extended range arithmetic. The long-wavelength residual Faye anomaly synthesised on the $1' \times 1'$ grid up to a specific degree n was then subtracted from the residual Faye anomalies.

1D-FFT Stokes integration was tested with two deterministically modified integration kernels: the kernel of Wong and Gore (1969), herein called the WG-kernel, and the kernel of Featherstone et al. (1998), herein called the FEO-kernel. The WG-kernel $S_{\text{WG}}(\psi)$ is defined as

$$S_{\text{WG}}(\psi) = S(\psi) - \sum_{n=2}^{M-1} \frac{2n+1}{n-1} P_n(\cos\psi) \tag{5}$$

where ψ is the angular distance between the computation point and the observation point, $S(\psi)$ is the unmodified Stokes kernel, M is the modification degree, and $P_n(\cos\psi)$ is the Legendre polynomial of degree n. The FEO-kernel $S_{\text{FEO}}(\psi)$ is defined as

$$S_{\text{FEO}}(\psi) = S_{\text{VK}}(\psi) - S_{\text{VK}}(\psi_0) \quad \text{for } 0 \le \psi \le \psi_0 \tag{6}$$

where ψ_0 is the integration cap size, $S_{\text{VK}}(\psi)$ is the Vaníček and Kleusberg (1987) kernel

$$S_{\text{VK}}(\psi) = S_{\text{WG}}(\psi) - \sum_{k=2}^{M-1} \frac{2k+1}{2} t_k(\psi_0) P_n(\cos\psi) \tag{7}$$

and $t_k(\psi_0)$ is the modification coefficient defined in Vaníček and Kleusberg (1987).

Parameter sweeps were performed to identify the optimal cap size and modification degree (as per Featherstone et al. 2018) for both the WG- and FEO-kernel. In a parameter

sweep, multiple computations are run with a range of different cap sizes and modification degrees followed by testing of the resulting geoid against GNSS-levelling data (Sects. 2.6 and 2.7). The solution with the lowest standard deviation (or L2-norm) of differences with GNSS-levelling data is considered optimal. The FEO-kernel was used for AUSGeoid2020 (Featherstone et al. 2018), while the WG-kernel was used for NZGeoid2017 (McCubbine et al. 2018), and both kernels have been used for numerous other geoid models. The two kernels were tested here to see whether one performed differently to the other in the mountainous Colorado experiment area. While many tests of modified kernels have been performed before (e.g. Forsberg and Featherstone 1998; Ellmann 2005a; Li and Wang 2011), optimal results vary by area (e.g. Featherstone 2003), and the parameter sweeps therefore need to be performed for the Colorado area.

The decision to use the FEO- and WG-kernels, and to determine the modification parameters through parameter sweeps, was made solely on the basis that this is a significant feature of the AUSGeoid (and NZGeoid) computation method, which we set out to apply and test. Many other Stokes kernel modification methods exist: see, for example, Featherstone (2013, Appendix A) for an overview. We acknowledge that stochastic kernel modification (e.g. Sjöberg 1981, 2003; Wenzel 1983), which takes into account accuracy estimates for the GGM coefficients and local gravity data, may provide better results in the Colorado area. It is, at least conceptually, desirable to use spectral information for gravity data and noise in the combination of satellite, airborne and terrestrial data (e.g. Kern et al. 2003). However, error spectra of airborne and terrestrial data are difficult to estimate and are typically based on simple covariance models (e.g. Ellmann 2005b) subject to stationarity and isotropy assumptions (Featherstone 2013). We also acknowledge that parameter sweeps may not result in the optimal kernel, especially when the GNSS-levelling data are of poor or ambiguous quality. The results of the parameter sweeps are known to potentially be influenced by the treatment of outliers and by distortions in the height datum (e.g. Featherstone et al. 2018).

2.6 Geoid and potential computation

A zero-degree geoid term was estimated to take into account the geocentric gravitational constant selected for the project ($\mathrm{GM} = 3.986004415 \cdot 10^{14}\,\mathrm{m}^3\,\mathrm{s}^{-2}$) and the conventional reference potential value used in the IHRS ($W_0 = 62636853.4\,\mathrm{m}^2\,\mathrm{s}^{-2}$, Sánchez et al. 2016; Sánchez and Sideris 2017). This zero-degree term was added to the restored height anomalies to provide a final quasigeoid model.

As the AUSGeoid models are quasigeoid models, estimation of the geoid-to-quasigeoid separation is not included in the AUSGeoid computation procedure. In adherence to the basic requirements of the IHRS Colorado experiment, the geoid-to-quasigeoid separation was here approximated using the simple equation (e.g. Heiskanen and Moritz 1967, section 8-13; Rapp 1997a)

$$N - \zeta = \frac{\Delta g_{\mathrm{B}} H}{\gamma} \tag{8}$$

where N is the geoid height, ζ is the quasigeoid height (or height anomaly), Δg_{B} is the Bouguer gravity anomaly, H is the orthometric height, and γ is the magnitude of reference gravity. We acknowledge that there are more sophisticated methods for the computation of the geoid-to-quasigeoid separation (e.g. Flury and Rummel 2009; Sjöberg 2010; Tenzer et al. 2015), but these have not been applied here.

To create a grid of potential values, the magnitude of the GRS80 normal gravity γ and normal potential U were computed rigorously at the topography using Eqs. (2-62) and (2-72) in Heiskanen and Moritz (1967). The generalised Bruns's formula was used to compute the disturbing potential T, with the gravity potential W at the topography obtained by adding the normal potential U at the topography to the disturbing potential.

2.7 GNSS-levelling comparisons

As data from the Colorado Geoid Slope Validation Survey 2017 (GSVS17; Van Westrum 2019) were not yet publicly available at the time of writing, results were validated against an older data set of 509 GNSS-levelling points of unknown precision over the computation area. The levelling data refer to the NAVD88 (Zilkoski et al. 1992). The NAVD88 has long been known to contain systematic errors (e.g. Smith and Milbert 1999), most recently estimated as ranging from -20 cm to +130 cm across the USA (Li 2018). Therefore, systematic error was mitigated by estimating and removing a bias and tilt. The removal of a bias and tilt will also account for most systematic error resulting from different tide systems.

To avoid the influence of edge effects in the geoid computation, a 0.5° buffer was removed, so the validation area is (35.5°–39.5°N, 102.5°–109.5°W). This reduces the number of available GNSS-levelling points to 309. In this data set, 4 outliers were identified based on a 3σ-criterion. These outliers were removed, so a total of 305 points were used in all comparisons.

The geoid values used in the GNSS-levelling comparisons were prepared in two different ways, so as to test whether the interpolation method affected these comparisons. In the first method, herein called the N-interpolation method, the height anomalies were first converted to geoid heights (Eq. 8) and then bi-cubically interpolated to the horizontal coordinates of the GNSS-levelling marks. In the second method, herein called the ζ-interpolation method, the height anomalies were bi-cubically interpolated to the horizontal coordinates of the

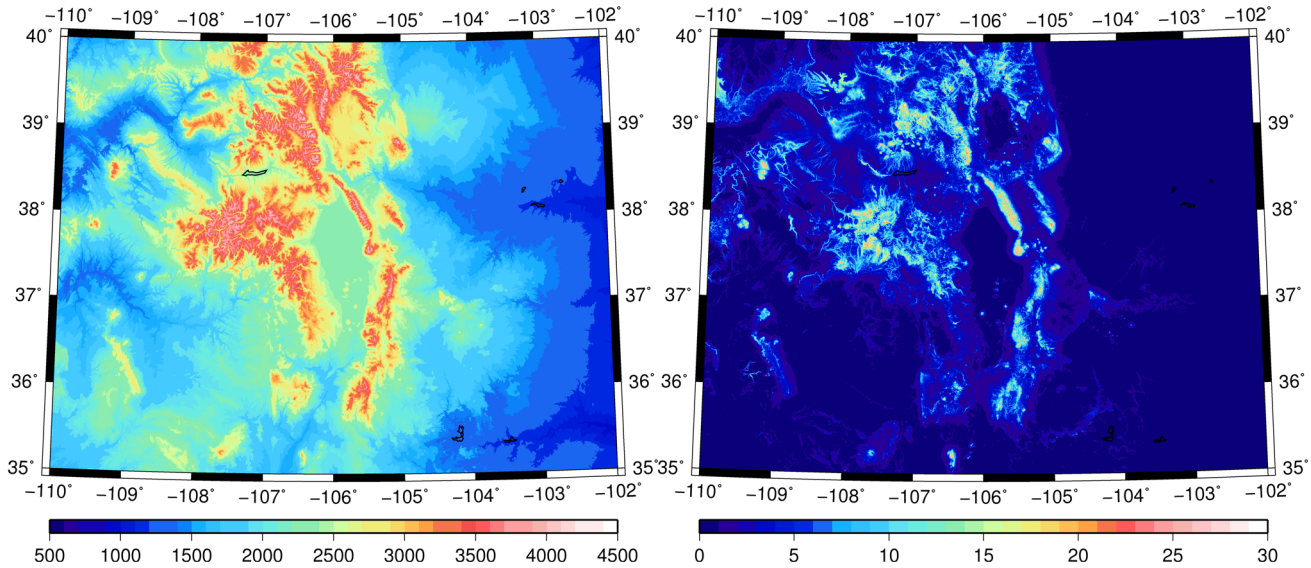

Fig. 3 Left: DEM terrain heights over the computation area (units in m; max: 4385 m, min: 932 m, mean: 2017 m, rms: 2109 m); right: terrain corrections according to Goyal (2019) (units in mGal; max: 56.1 mGal, min: 0.0 mGal, mean: 2.0 mGal, rms: 3.6 mGal)

GNSS-levelling marks and then converted to geoid heights. The latter was achieved by bi-cubic interpolation of the Bouguer anomalies to the marks, followed by application of Eq. (8) using the orthometric height of the GNSS-levelling point.

3 Results excluding airborne gravity data

Results of "terrestrial-only" geoid solutions are presented in this section, and results from the combination of terrestrial with airborne data are presented in Sect. 4. Note that all results are presented to the nearest 0.1 mGal and 0.001 m, except for standard deviations of GNSS-levelling comparisons which are shown with one additional decimal. The number of decimals shown should not be taken as an indication of precision, but to differentiate between similar results.

3.1 Terrain corrections

The DEM and terrain corrections following Goyal et al. (2019) over the $5° \times 8°$ computation area (35°–40°N, 102°–110°W) are shown in Fig. 3. For the computation of the terrain corrections, the full extent of the DEM (33°–42°N, 100°–112°W) was used. The differences with the terrain corrections following McCubbine et al. (2017), using software used for the computation of AUSGeoid2020, reach a maximum of 11.2 mGal, which translates into a maximum difference in the final geoid model of ~ 0.038 m. These differences are primarily attributed to the truncation of the binomial expansion in the method of McCubbine et al. (2017).

3.2 Terrestrial gravity gridding

As explained in Sect. 2.2, two gravity gridding methods were compared: (1) the simple gridding method and (2) the reconstruction method of Featherstone and Kirby (2000). The latter reduces aliasing which may arise when more gravity observations are taken in lowland areas. For the Colorado test area, the histograms in Fig. 4 indicate relatively few gravity observations are taken at low altitudes (~ 1000–1500 m altitude), but this is due to low observation density in the plains in the Eastern part of the computation area (Fig. 1). Most observations are taken between 1500 and 2200 m, although there are still many observed up to ~ 3000 m. Relatively few observations are taken at high altitudes between ~ 3000 and ~ 3500 m.

A potential disadvantage of the reconstruction method is that interpolation errors may be larger, because the simple Bouguer anomalies that are interpolated are less smooth than refined Bouguer anomalies. This is especially so in the most mountainous terrain and therefore more of a concern in Colorado than in Australia. In addition, the $3'' \times 3''$ grid resolution of the DEM over Colorado is coarser than the $1'' \times 1''$ grid resolution used over Australia (Featherstone et al. 2018), which may also lead to larger interpolation errors.

The differences between the final $1' \times 1'$ grid of Faye anomalies (simple gridding versus the reconstruction method of Featherstone and Kirby 2000), and the effect on height anomalies is shown in Fig. 5. Maximum differences are in excess of 20 mGal, resulting in height anomaly differences up to 0.085 m in magnitude. However, when both models are validated against available GNSS-levelling data, the standard deviation of the differences is very small. The simple gridding method gives a standard deviation of differences

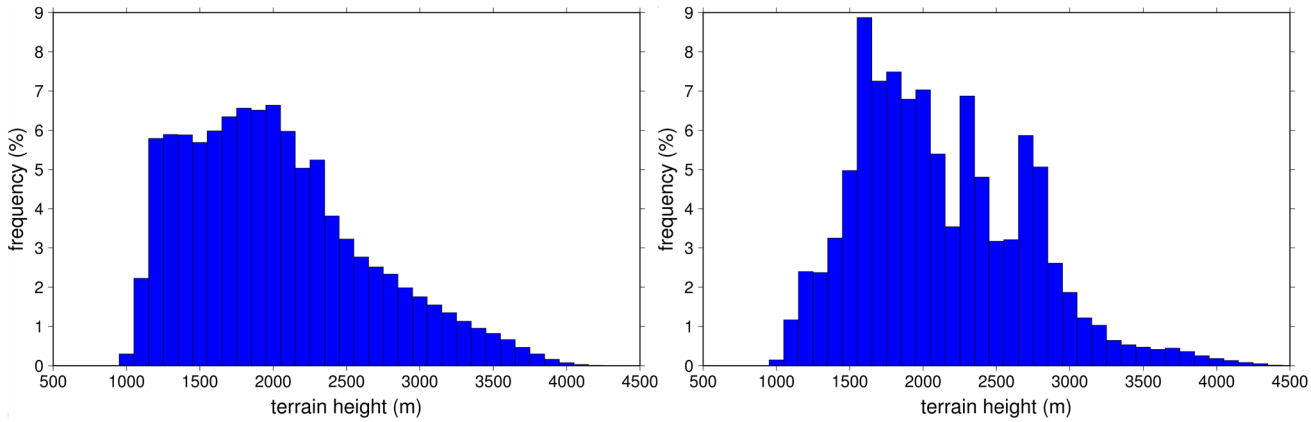

Fig. 4 Left: histogram of orthometric heights in the 3″ × 3″ DEM over the computation area (57,615,601 points; mean: 2013 m); right: histogram of orthometric heights of gravity observations over the computation area (59,303 points; mean: 2111 m)

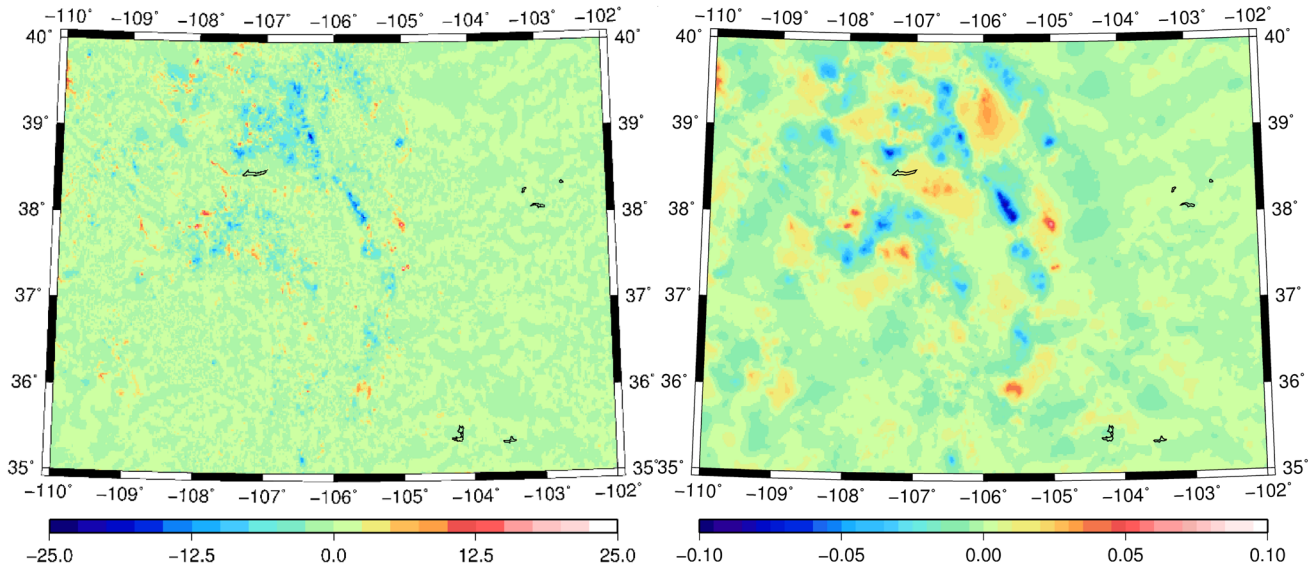

Fig. 5 Left: differences between Faye anomalies from the simple gridding method and the reconstruction method (units in mGal; max: 20.5 mGal, min: − 20.8 mGal, mean: − 0.2 mGal, rms: 1.5 mGal); right: differences between height anomalies (units in m; max: 0.065 m, min: − 0.085 m; mean: 0.001 m, rms: 0.009 m)

with GNSS-levelling of ± 0.0438 m, while the reconstruction method gives a standard deviation of ± 0.0441 m. We investigated correlation of differences with terrain height and with terrain gradients for both methods, but no significant difference between the two gridding methods was found. The simple gridding method was selected for subsequent computations based on the slightly smaller (although not significant) standard deviation of differences with GNSS-levelling.

3.3 Stokes integration

Figure 6 shows the residual Faye anomalies resulting from the simple gridding method (gridding of refined Bouguer anomalies). Before Stokes integration and before application of any filtering of the residual Faye anomalies, a surface spherical harmonic series of the residual Faye anomalies was computed using the procedure described in Sect. 2.5. The power spectrum of this series is shown in Fig. 7. It shows that degree variances below degree 2159 are substantially smaller than those above degree 2159, as expected, but a peak in degrees below 100 can also be seen. This peak will result in a long-wavelength signal in the residual height anomalies if it is not filtered out.

Unmodified Stokes integration of these residual anomalies results in a large and dominant long-wavelength effect in the residual height anomalies. This is undesirable, as the long-wavelength height anomaly/geoid signal is better modelled by the GGM, which includes high-quality GRACE and

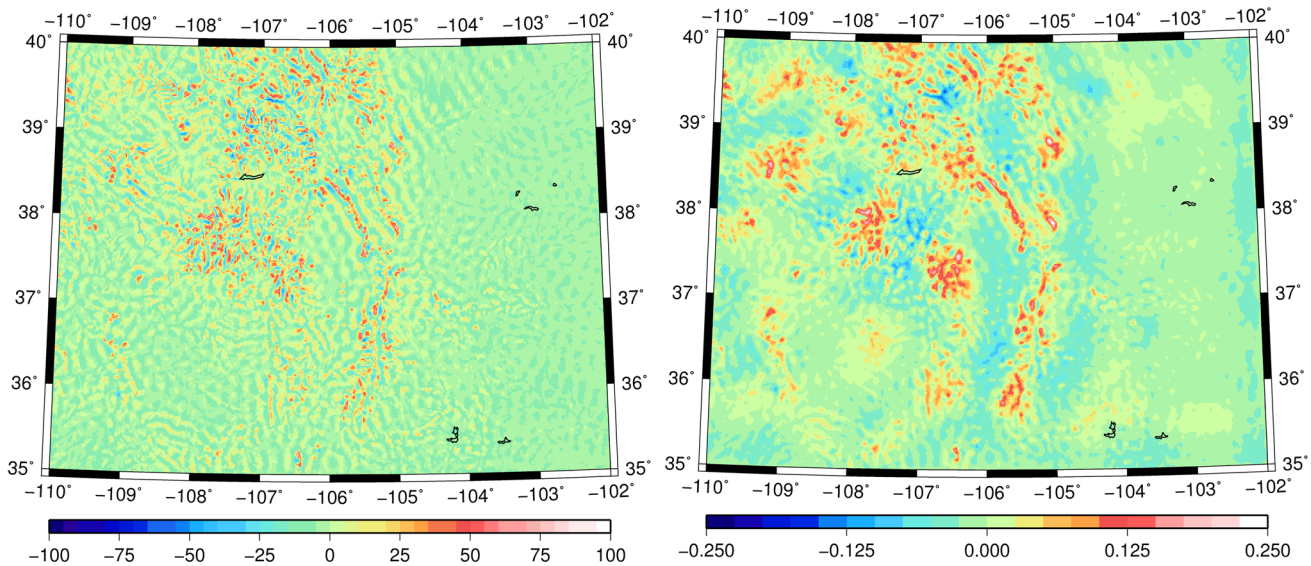

Fig. 6 Left: residual Faye anomalies (units in mGal; max: 92.2 mGal, min: − 64.4 mGal, mean: − 1.2 mGal, rms: 10.5 mGal); right: residual height anomalies using WG-kernel with $\psi_0 = 1.4°$ and $M = 280$ (units in m; max: 0.248 m, min: − 0.155 m; mean: − 0.005 m, rms: 0.032 m)

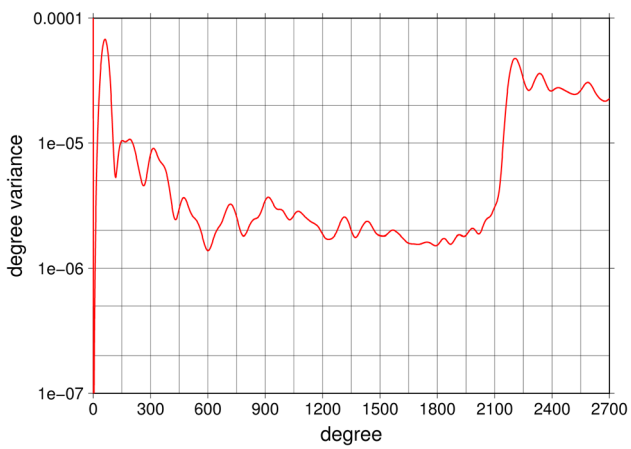

Fig. 7 Spherical harmonic spectrum of residual Faye anomalies, showing a peak in degrees below 100 attributed to errors in the Faye anomaly grid

GOCE data. However, integration with a modified kernel can reduce this long-wavelength effect (e.g. Vaníček and Featherstone 1998). A parameter sweep was performed to determine the optimal Stokes kernel modification type, cap size and modification degree, as the values used in the computation of AUSGeoid2020 (FEO-kernel with $\psi_0 = 0.5°$ and $M = 40$) may not be suitable for the Colorado area.

Stokes integration was performed for cap sizes from $\psi_0 = 0°$ to $\psi_0 = 2°$ with a step size of 0.1° and for modification degrees from $M = 40$ to $M = 360$ with a step size of 40. This was done for both the FEO-kernel (Featherstone et al. 1998) and the WG-kernel (Wong and Gore 1969). Resulting gravimetric geoid models were compared to the available GNSS-levelling data (Sect. 3.5).

Standard deviations of the differences between gravimetric geoid heights and GNSS-levelling-derived geoid heights are plotted in Fig. 8. It can be seen that the WG-kernel filters the long-wavelength errors more rapidly with increasing modification degree. Furthermore, the WG-kernel remains stable with increasing cap size, while the FEO-kernel shows irregular behaviour at high modification degrees. The optimal parameters from this parameter sweep are a WG-kernel with cap size $\psi_0 = 1.4°$ and modification degree $M = 280$. The resulting model is herein called model T ("T" for "terrestrial only").

A pragmatic alternative to reliance on kernel modification for signal filtering is to apply a spherical harmonic high-pass filter up to a specific degree n to the residual Faye anomalies before Stokes integration. This adds another parameter to the search space, as the value of n can be optimised. For instance, an extreme choice is to filter out all signals up to $n = 2160$. This would result in a geoid model where the height anomalies up to $n = 2160$ are completely determined by the GGM, and the Stokes integration only adds higher-frequency information.

Figure 9 shows parameter sweep results for this case, again with both the FEO- and WG-kernel. When compared to Fig. 8, it shows that due to the spherical harmonic filter, the fit to GNSS-levelling is much less sensitive to modification degree and cap size. Higher modification degrees still improve the fit for both kernels, but the FEO-kernel again shows unstable behaviour at high modification degrees for large cap sizes.

The optimal FEO-kernel result is achieved with $\psi_0 = 0.1°$ with any modification degree. The optimal WG-kernel result is achieved with $\psi_0 = 0.5°$ and $M = 360$. Both solu-

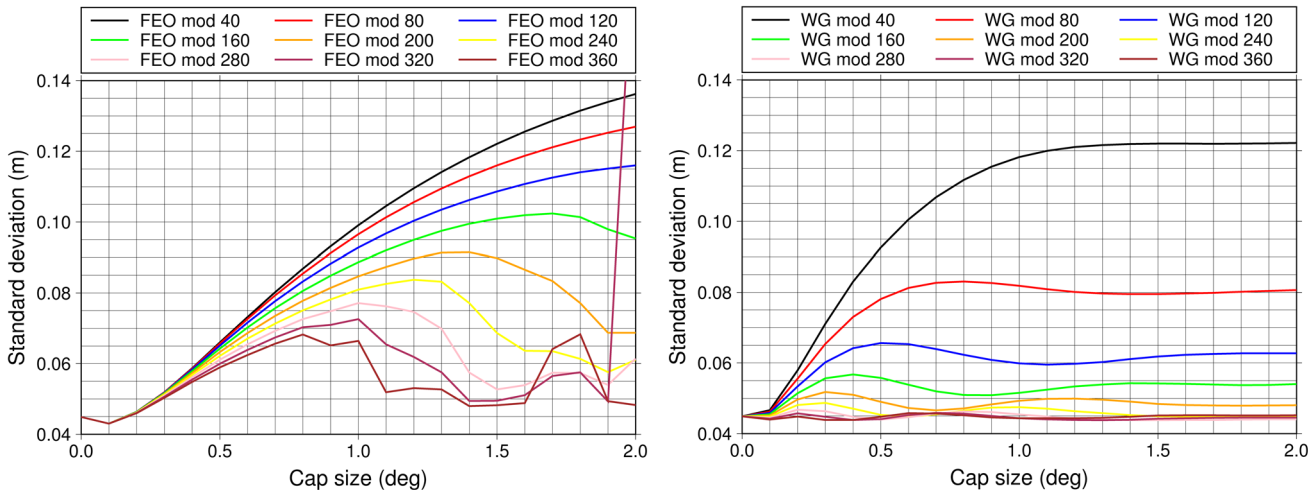

Fig. 8 Standard deviations from GNSS-levelling comparisons for different cap sizes and modification degrees using the FEO-kernel (left) and the WG-kernel (right)

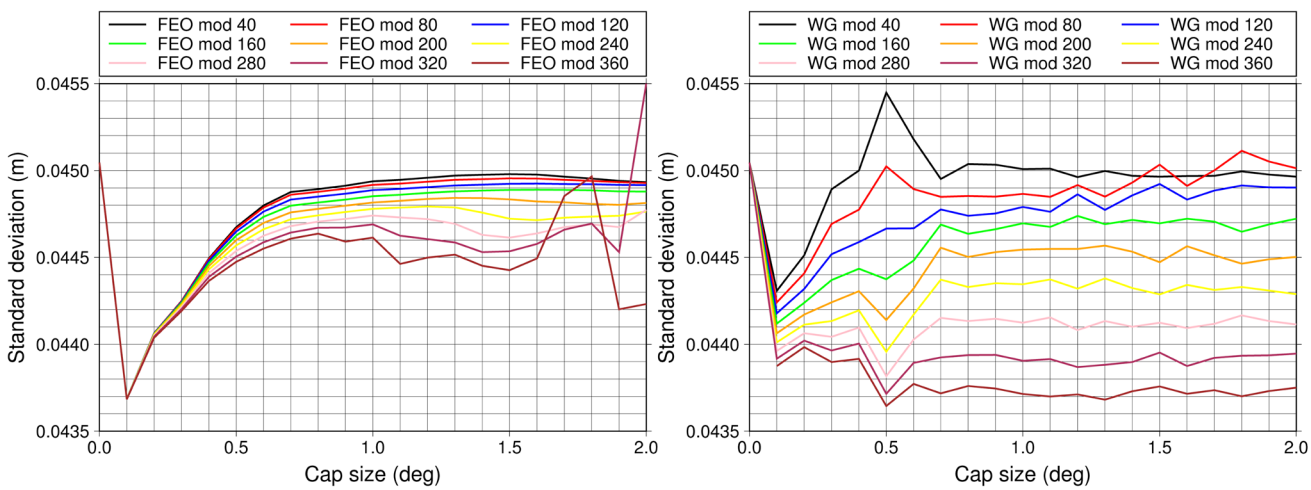

Fig. 9 Standard deviations from GNSS-levelling comparisons after spherical harmonic filtering to $n = 2160$ for different cap sizes and modification degrees using the FEO-kernel (left) and the WG-kernel (right)

tions give a similar standard deviation in GNSS-levelling differences of ± 0.0437 m. This is almost identical to the standard deviation obtained without spherical harmonic filtering (± 0.0438 m, Fig. 8), and it is therefore difficult to ascertain which solution is optimal. However, extended analysis of the GNSS-levelling comparisons (Sect. 3.5) provides further insight. The model using the FEO-kernel with modification parameters $\psi_0 = 0.1°$ and $M = 40$ is herein called model T-SH2160.

3.4 Geoid and potential computation

The geoid-to-quasigeoid separation computed according to Eq. (8) reaches a maximum magnitude of 1.370 m (Fig. 10). It is largely dependent on terrain height and hence has significant power over short wavelengths. When added to the relatively smooth quasigeoid resulting from Stokes integration, it results in a geoid model that is 'less smooth', i.e. has more power over short wavelengths than the quasigeoid model. However, from theoretical considerations, the geoid, being an equipotential surface of the Earth's gravity field, should be smoother than the quasigeoid (e.g. Sjöberg 2013). This indicates that the quasigeoid models computed using the AUSGeoid approach may be too smooth, i.e. lacking in power over short wavelengths. This is not a significant issue over flat terrain, but is more significant over the mountainous terrain of Colorado.

Models of the geoid and the gravity potential were computed based on all of the quasigeoid solutions described above (with different modifications and parameter settings). These are not shown here, because there is little visual differ-

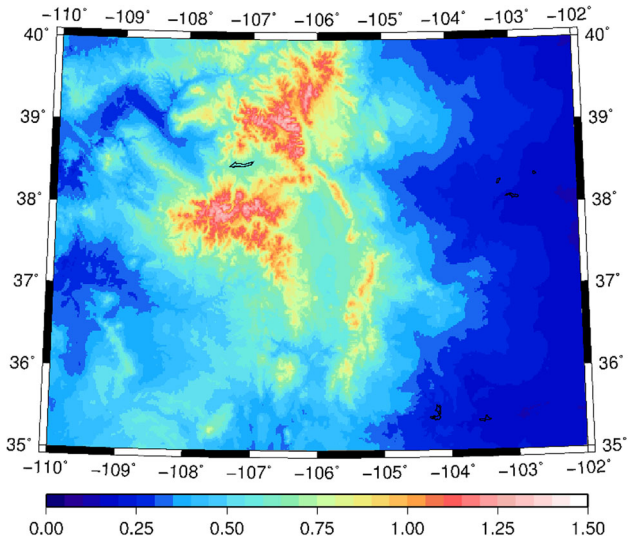

Fig. 10 Geoid-to-quasigeoid separation using Eq. (8) (units in m; max: 1.370 m, min: 0.130 m; mean: 0.462 m, rms: 0.514 m)

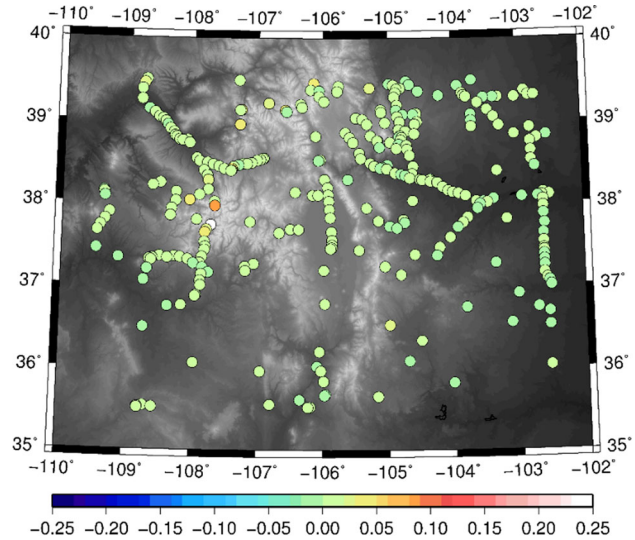

Fig. 11 Differences between geoid heights at GNSS-levelling points computed using the N-interpolation and ζ-interpolation methods (units in m; max: 0.291 m, min: − 0.012 m; mean: 0.005 m, rms: 0.020 m)

ence to the combined terrestrial and airborne gravity models shown in Sect. 4.2 (Fig. 16).

3.5 GNSS-levelling comparison

As per Sect. 2.7, the geoid models computed in this research were validated against 305 GNSS-levelling points. The differences between geoid heights at the GNSS-levelling points from the N-interpolation method and ζ-interpolation method are shown in Fig. 11. For most points, the difference is under 0.020 m, but there is a single point (107.7°W, 37.7°N) with a difference of 0.291 m, indicating that the choice of interpolation method can significantly affect the results, especially in regions of rugged mountainous topography.

An example of the impact of the differences shown in Fig. 11 can be seen upon comparison of two different geoid models. Model T is the geoid model computed using a WG-kernel with $\psi_0 = 1.4°$ and $M = 280$. Model T-SH2160 is the geoid model computed using a FEO-kernel with $\psi_0 = 0.1°$ and $M = 40$ after spherical harmonic filtering to $n = 2160$. Their standard deviation of differences with GNSS-levelling is almost identical using the ζ-interpolation method (after removal of bias and tilt). This is the method that was used in the parameter sweeps in Sect. 3.3. However, the results are different using the N-interpolation method (Table 1). The standard deviation of model T is lower using N-interpolation, while the standard deviation of model T-SH2160 is larger. This suggests model T is the superior model, particularly when used with the N-interpolation method. It is noticeable that the WG model appears more compatible with the N-interpolation, while the FEO model with spherical harmonic filtering is more suited to the ζ-interpolation. The differences in standard deviations shown in Table 1 are caused primarily by only a few GNSS-levelling stations located in regions of steep terrain gradients that are highly sensitive to the interpolation method, resulting in large magnitude differences (>0.05 m). The spatial distribution of the differences using the ζ-interpolation method is shown in Fig. 12.

4 Results including airborne gravity data

4.1 3D least-squares collocation

When airborne gravity data are added, the gridding of observational data is performed in a completely different manner, as explained in Sect. 2.4. The combination of terrestrial and airborne data is performed using 3D least-squares collocation based on residual gravity anomalies Δg_{res}. Both the terrestrial and airborne gravity data are reduced to residual gravity anomalies. The choice of low-degree DEM in the computation of these residual gravity anomalies has a significant influence on the final geoid model. Figure 13 shows the low-degree DEM generated using a 2D Gaussian filter (left) and the low-degree DEM generated using a spherical harmonic box filter (right). Note that the spherical harmonic box filter results in a DEM with steeper slopes.

Figure 14 shows the terrestrial (left) and airborne (right) residual gravity anomalies, resulting from the application of a 2D Gaussian filter to construct the low-degree DEM. As expected, the airborne anomalies are smoother than the terrestrial anomalies, but visually there is a good agreement between the main features of both data sets.

The impact of the different low-degree DEMs can be seen in Fig. 15. The left panel shows the differences between

Table 1 Statistics of the GNSS-levelling comparison using two different geoid interpolation methods

Model	Interpolation method	Maximum (m)	Minimum (m)	Standard deviation (m)
T (WG)	ζ-interpolation	0.140	− 0.228	0.0438
T (WG)	N-interpolation	0.153	− 0.120	0.0425
T-SH2160 (FEO)	ζ-interpolation	0.181	− 0.116	0.0437
T-SH2160 (FEO)	N-interpolation	0.254	− 0.109	0.0464

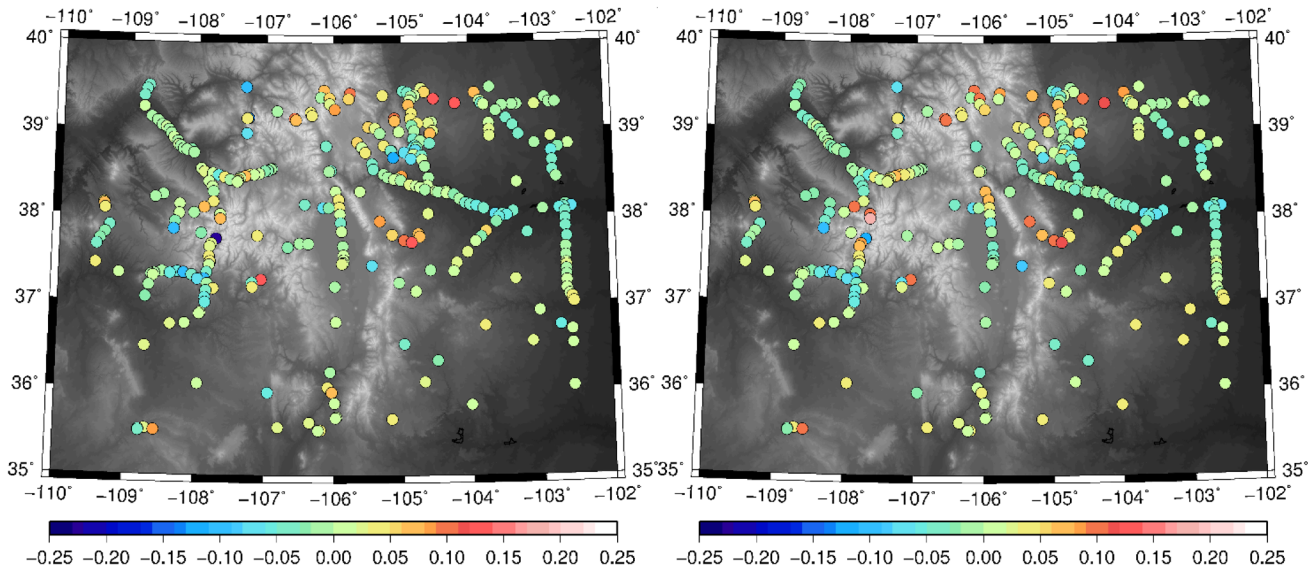

Fig. 12 Left: GNSS-levelling differences after removal of bias and tilt for geoid model T using WG-kernel with $\psi_0 = 1.4°$ and $M = 280$ (units in m; max: 0.140 m, min: − 0.228 m; mean: 0.000 m, rms: 0.044 m); right: GNSS-levelling differences after removal of bias and tilt for geoid model T-2160 using FEO-kernel with $\psi_0 = 0.1°$ and $M = 40$ after spherical harmonic filter to $n = 2160$ (units in m; max: 0.181 m, min: − 0.116 m; mean: 0.000 m, rms: 0.044 m); both have used the ζ-interpolation method

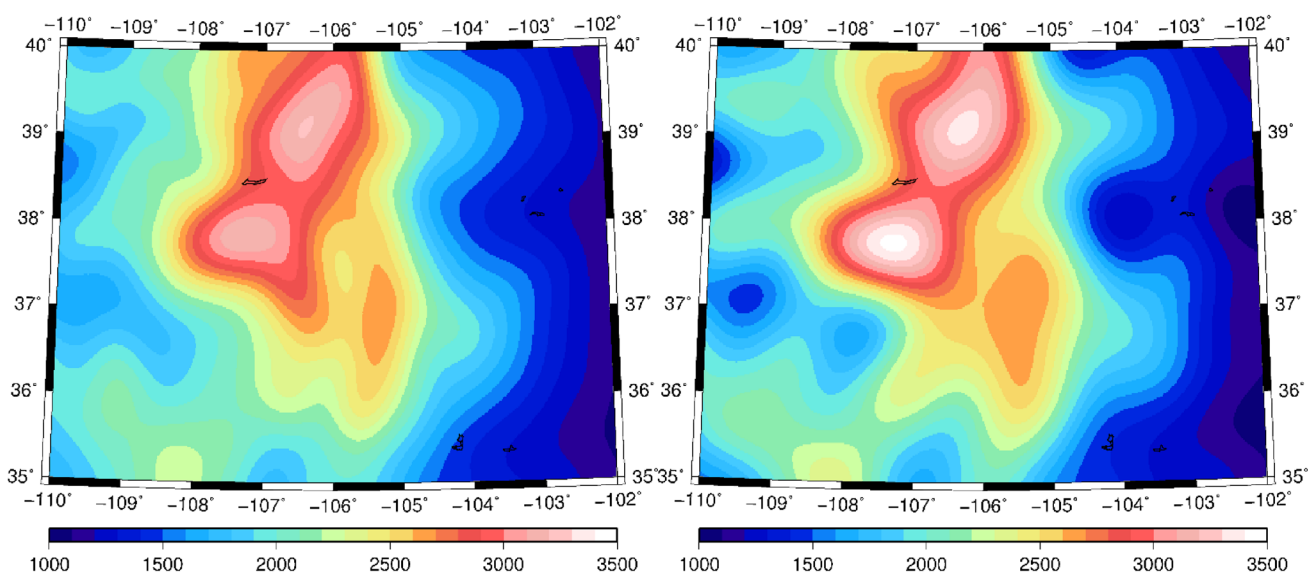

Fig. 13 Left: low-degree DEM from 2D Gaussian filter (units in m; max: 3210.204 m, min: 1076.610 m; mean: 2010.018 m, rms: 2077.737 m); right: low-degree DEM from spherical harmonic box filter (units in m; max: 3434.994 m, min: 1049.561 m; mean: 2014.540 m, rms: 2090.629 m)

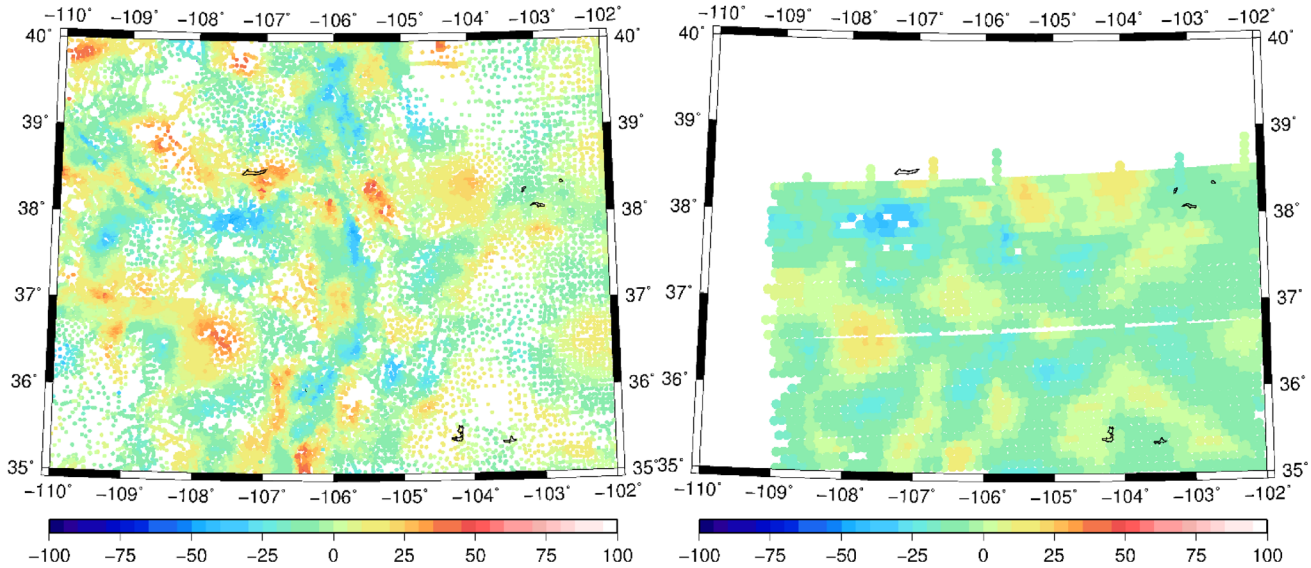

Fig. 14 Left: residual anomalies from terrestrial observations (units in mGal; max: 116.2 mGal, min: − 136.7 mGal; mean: − 1.7 mGal, rms: 15.0 mGal), right: residual anomalies from airborne observations (units in mGal; max: 24.2 mGal, min: − 40.8 mGal; mean: − 6.5 mGal, rms: 11.2 mGal) both using the 2D Gaussian filter

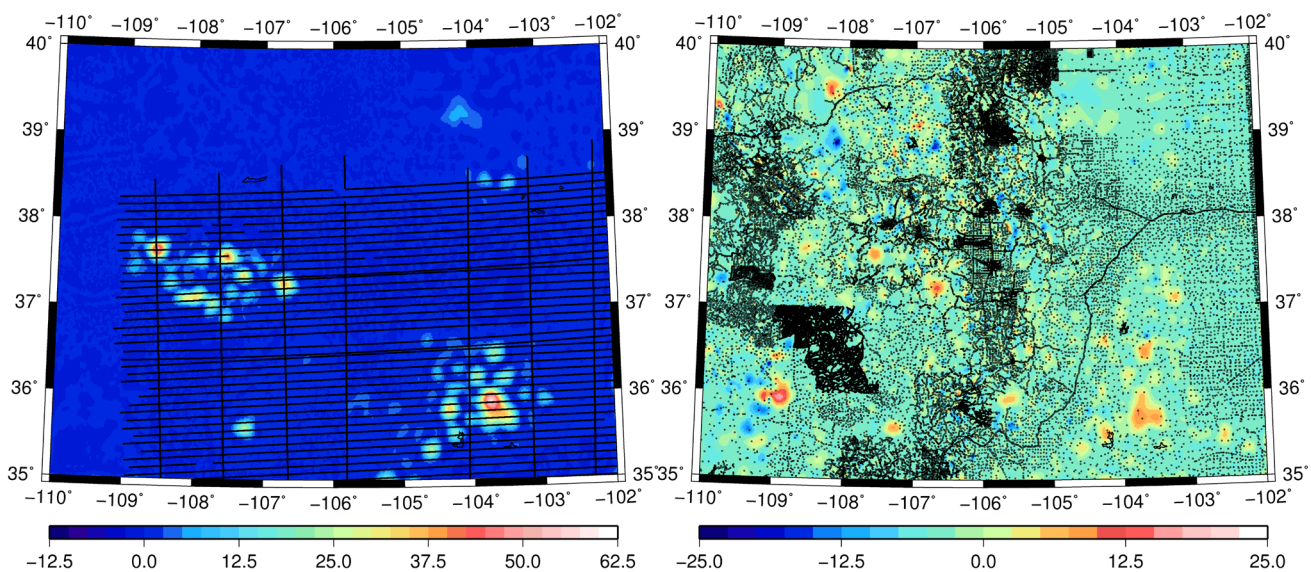

Fig. 15 Left: differences between Faye anomaly grids based on combined terrestrial and airborne observations using spherical harmonic box filter and 2D Gaussian filter, with locations of airborne gravity observations overlaid (units in mGal; max: 55.8 mGal, min: − 3.4 mGal; mean: 0.8 mGal, rms: 3.5 mGal); right: differences between Faye anomaly grids based on terrestrial observations only and combined terrestrial and airborne observations with 2D Gaussian filter used, with locations of terrestrial gravity observations overlaid (units in mGal; max: 21.4 mGal, min: − 36.4 mGal; mean: − 3.3 mGal, rms: 4.0 mGal)

reconstituted Faye anomalies after combination of terrestrial and airborne observations through 3D LSC. The differences are largest in locations where there are gaps in the terrestrial observations. The 2D Gaussian filter leads to combined terrestrial and airborne Faye anomalies that are closer to the terrestrial-only Faye anomalies. For this reason, results achieved with the 2D Gaussian filter are used in the remainder of this paper. The right panel in Fig. 15 shows the differences between the Faye anomaly grids based on terrestrial observations only and based on combined terrestrial and airborne observations. Part of the differences visible in the right panel in Fig. 15 are due to inclusion of airborne data, but there are also differences due to the different processing methods used, in particular the gridding of the data using either tensioned splines (terrestrial gravity) or 3D LSC (airborne gravity). These two effects are investigated further in Sect. 4.2. It can

be seen that the largest differences occur in areas where the terrestrial gravity observations are sparse.

4.2 Stokes integration and model comparison

A parameter sweep was performed to determine the optimal modified kernel, cap size and modification degree, as in Sect. 3.3. The lowest standard deviation of differences with GNSS-levelling data was achieved for the WG-kernel with a cap size $\psi_0 = 0.3°$ and $M = 360$. The standard deviations obtained were ± 0.0435 m (ζ-interpolation) and ± 0.0421 m (N-interpolation). These are the lowest of all models created in this study, cautiously suggesting that the addition of airborne data has improved the geoid model. This model is herein called model TA.

We have also tested spherical harmonic high-pass filtering of the residual Faye anomalies from the combined terrestrial and airborne gravity, as in Sect. 3.3 for the terrestrial-only gravity. The signal up to $n = 200$ was removed, so that the geoid up to this d/o is completely determined by the GGM. The choice for $n = 200$ was based on the evaluations of GOCE-based GGMs, indicating that cumulative height anomaly error up to this d/o for recent models is under 0.01 m (e.g. Rexer et al. 2014; Voigt and Denker 2015), which is expected to be superior to the terrestrial and airborne gravity data in this spectral range. However, standard deviations of differences to GNSS/levelling data for this solution were slightly larger than the solution without spherical harmonic high-pass filtering (± 0.0436 m with ζ-interpolation and ± 0.0439 m with N-interpolation). This indicates that there may be systematic errors in the long-wavelengths of the GNSS-levelling data.

The final geoid model for the solution without spherical harmonic high-pass filtering (model TA) is shown in Fig. 16 (left). The gravity potential at terrain height resulting from this solution is shown in Fig. 16 (right). Figure 17 (top left) shows the differences between the geoid models computed without and with inclusion of airborne gravity data. This compares the terrestrial-only model T with WG-kernel, $\psi_0 = 1.4°$ and $M = 280$ to the combined terrestrial and airborne model TA with WG-kernel, $\psi_0 = 0.3°$ and $M = 360$. As expected, the differences are largest over areas where the airborne gravity fills in gaps in the terrestrial data, but there are also substantial differences in areas without airborne gravity coverage (roughly north of 38°N and west of 109°W; cf. Figure 2). This is due to the different gridding techniques employed and the different kernel modification parameters used for models T and TA.

Figure 17 (top right) shows the impact of the different gridding techniques. For this, the terrestrial data only were gridded using the 3D LSC gridding technique and then processed identically to the terrestrial-only solution with the same kernel modification parameters. This shows that the gridding technique is responsible for some of the differences seen in Fig. 17 (top left), but the remaining differences are due to the influence of airborne observations on the geoid model (and the resulting different kernel modification parameters used).

The bottom two panels in Fig. 17 show the differences between model T and model T-SH2160 (left) and model T-SH2160 and model TA (right). The left of these panels shows that spherical harmonic high-pass filtering of the terrestrial gravity results in significant medium-wavelength differences

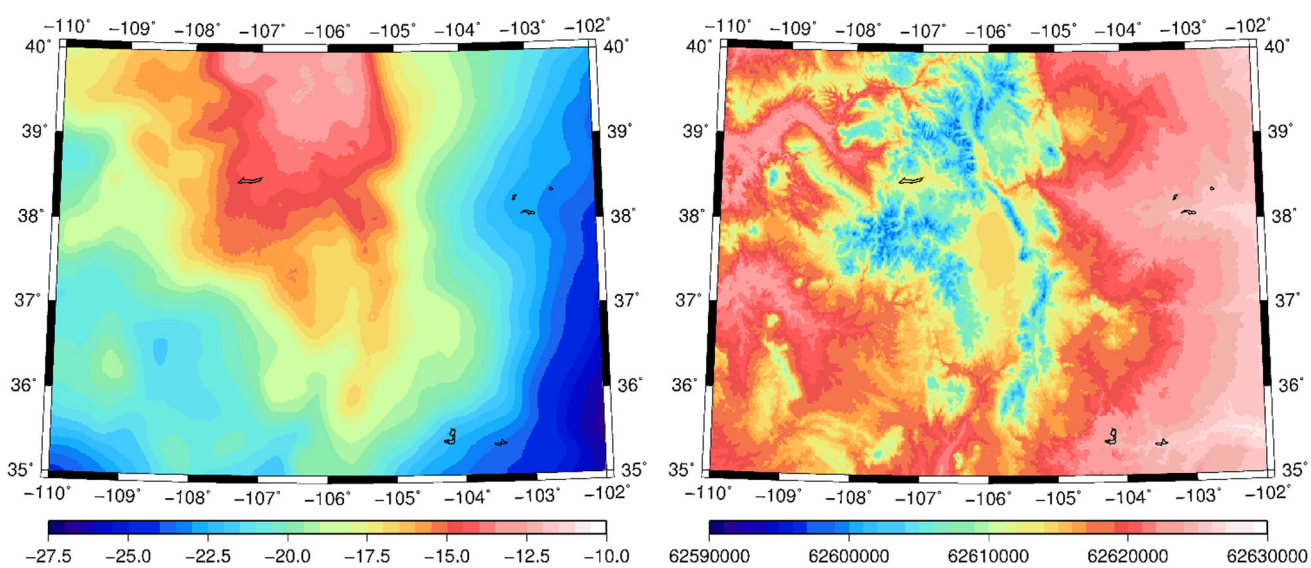

Fig. 16 Left: final geoid model (units in m; max: − 12.241 m, min: − 26.717 m; mean: − 19.163 m, rms: 19.413 m), right: final gravity potential at terrain height (units in m; max: 62,627,536.470 m^2/s^2, min: 62,596,746.447 m^2/s^2; mean: 62,617,112.968 m^2/s^2, rms: 62,617,113.255 m^2/s^2)

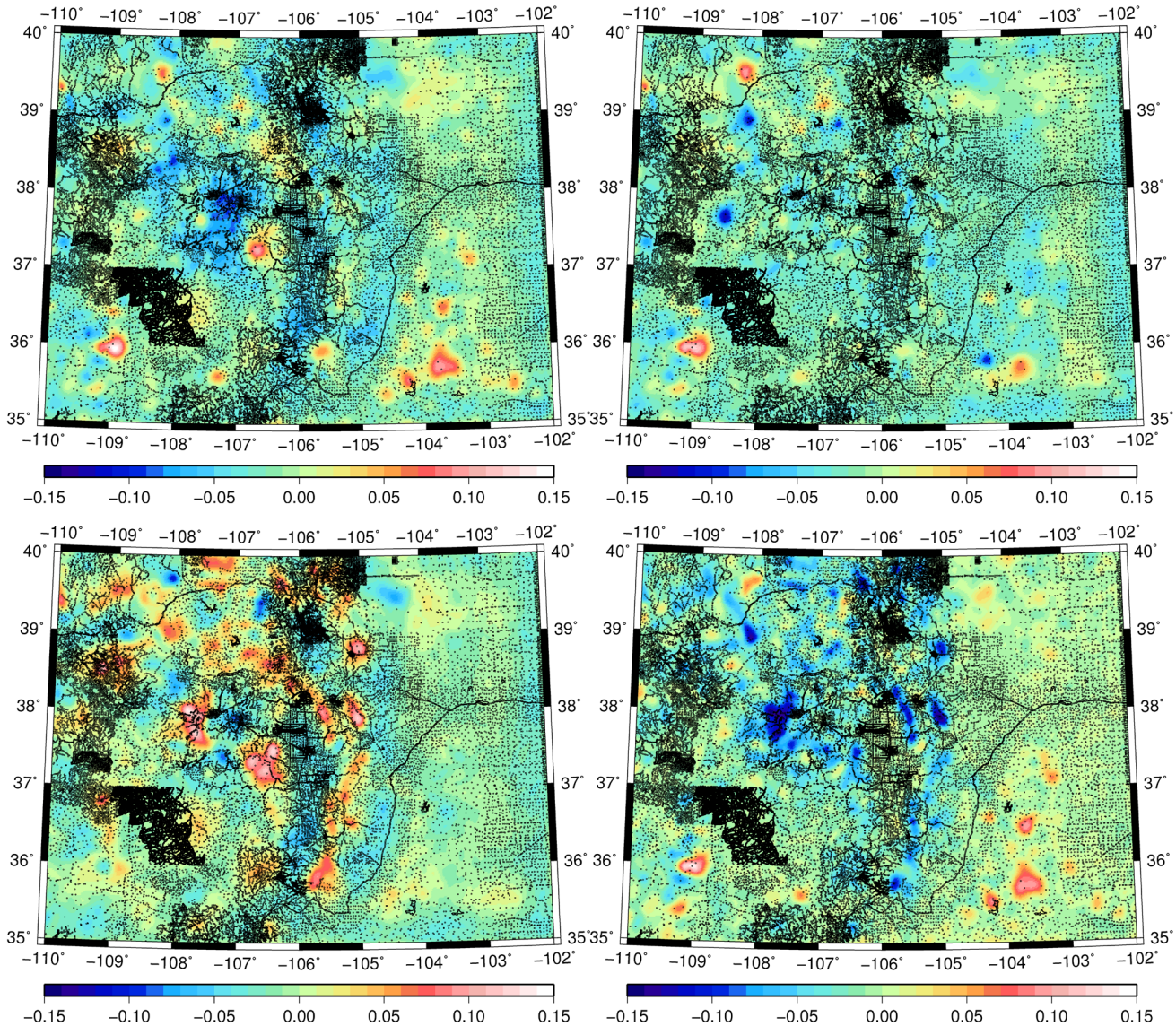

Fig. 17 Top left: differences between geoid models computed using terrestrial data only and terrestrial and airborne data (T vs TA) (units in m; max: 0.147 m, min: − 0.107 m; mean: − 0.018 m, rms: 0.028 m); top right: differences between geoid models computed using terrestrial data only using simple gridding and LSC gridding (units in m; max: 0.123 m, min: − 0.122 m; mean: − 0.018 m, rms: 0.023 m); bottom left: difference between geoid models computed using terrestrial data without and with spherical harmonic high-pass filtering to d/o 2160 (T vs T-SH2160) (units in m; max: 0.136 m, min: − 0.106 m; mean: − 0.005 m, rms: 0.026 m); bottom right: differences between geoid models computed using terrestrial data with spherical harmonic high-pass filtering to d/o 2160 and terrestrial and airborne data without spherical harmonic filtering (T-SH2160 vs TA) (units in m; max: 0.141 m, min: − 0.151 m; mean: − 0.013 m, rms: 0.027 m). All figures have the locations of terrestrial gravity observations overlaid

that are roughly correlated with the topography (largest differences over the most mountainous terrain).

5 Discussion

The final results shown in Table 2 are interesting in that the overall standard deviations of each model were no more than 0.3 mm different, suggesting that the AUSGeoid computation method is robust when different methods or data preparation is used. However, we should be cautious when interpreting these results, because the differences between the models are in excess of 0.02 m (Table 3), and therefore of significance in the quest for the 1 cm geoid (e.g. Rapp 1997b, Foroughi et al. 2019). Furthermore, the GNSS-levelling heights are related to the NAVD88, and may contain levelling errors, GNSS observation errors, and errors in the data and method used in the Helmert height correction for the NAVD88 heights,

Table 2 GNSS-levelling comparisons among geoids from terrestrial (T), terrestrial with spherical harmonic high-pass filtering to d/o 2160 (T-SH2160) and combined terrestrial and airborne (TA) gravity. All have used the ζ-interpolation method

Model	Kernel	Cap size (°)	Modification degree	Minimum (m)	Maximum (m)	Standard deviation (m)
T	WG	1.4	280	− 0.228	0.140	0.0438
T-SH2160	FEO	0.1	40	− 0.116	0.181	0.0437
TA	WG	0.3	360	− 0.215	0.136	0.0435

Table 3 Statistics of differences between geoids from terrestrial (T), terrestrial with spherical harmonic high-pass filtering to d/o 2160 (T-SH2160) and combined terrestrial and airborne (TA) gravity

Models	Minimum (m)	Maximum (m)	Mean (m)	RMS (m)
T vs T-SH2160	− 0.106	0.136	− 0.005	0.026
TA vs T-SH2160	− 0.141	0.151	0.013	0.027
T vs TA	− 0.107	0.147	− 0.018	0.028

which can all contribute to these data not necessarily being more accurate than the geoid models they are testing. In fact, our tests with spherical harmonic high-pass filtering of residual Faye anomalies suggest a potential issue with the long-wavelength behaviour of the GNSS-levelling data. Hence, it is problematic to suggest one geoid method is 'better' than the others. Another point to consider is that the GNSS-levelling points may be observed at specific height bands that may correlate with a particular geoid model's best performing heights, thus causing a bias in the results.

We also tested AUSGeoid methods with alternatives to help determine whether methods tailored to Australian conditions may be applied to a rugged mountainous area like that in the IHRS Colorado experiment, or whether these may need to be adapted. We tested the Goyal et al. (2019) and McCubbine et al. (2017) terrain corrections, finding a maximum of 0.038 m difference in geoid height which may be significant considering the GNSS-levelling differences of ~ 0.044 m (Table 2). We used the Goyal et al. (2019) terrain correction for the Colorado experiment as it is more rigorous in higher mountains.

Differences were also found between the AUSGeoid method of terrestrial gravity reconstruction (Featherstone and Kirby 2000) compared to the gridding of refined Bouguer anomalies. These differences ranged from 20.5 to − 20.8 mGal, or 0.065 m to − 0.085 m in the geoid, so are significant. We chose to use the simple gridding method (gridding of refined Bouguer anomalies), because the GNSS-levelling comparisons did not indicate an improved agreement from the reconstruction method, which is tailored for Australian terrestrial gravity.

The presence of long-wavelength signals in the residual Faye anomalies led us to test different filtering methods. The Stokes integration with the modified FEO-kernel has been standard as a filtering method in AUSGeoid computations (e.g. Featherstone et al. 2011, 2018), using parameter sweeps to optimise the cap size and modification degrees. For the Colorado experiment study region, we found the FEO-kernel to be unstable at modification degrees above 250, whereas the WG-kernel remained stable, with the optimal parameters of cap size $\psi_0 = 1.4°$ and modification $M = 280$. The instability of the FEO-kernel has also been noted in other computation areas by Featherstone (2003), Li and Wang (2011) and McCubbine et al. (2018).

We also experimented with a spherical harmonic high-pass filter prior to the Stokes integration. This efficiently removed the long-wavelength signal from the residual Faye anomalies and hence avoids long-wavelength errors in the geoid model. However, comparison with the GNSS-levelling produced mixed results, and the level of improvement, if any, could not be demonstrated conclusively. When high-quality validation data from the GSVS17 (Van Westrum 2019) become publicly available, this may be investigated further, as data from the earlier GSVS11 (Smith et al. 2013) and GSVS14 (Wang et al. 2015) surveys have been used to successfully confirm geoid accuracy at the 1 cm level.

To compare the geoid models computed in this study to the GNSS-levelling data at benchmarks, interpolation of the geoid height to the benchmarks is required. This is because the AUSGeoid method requires a grid format. We tested two different methods, again with mixed results, but it is important to realise that the two interpolation methods greatly affect the GNSS-levelling comparisons. From this study, it appears that the N-interpolation works best with the WG-kernel geoid, while the ζ-interpolation works reasonably well with both WG and FEO-kernels, but does not agree as well with the GNSS-levelling data.

The inclusion of the airborne gravity resulted in a maximum change in geoid height of 0.147 m, but a significant portion of this change is due to the different gridding technique employed when incorporating the airborne data (3D LSC vs tensioned spline gridding). The airborne gravity

appeared to improve the quality of the geoid model, although some of this improvement may also come from better coverage rather than better quality gravity data. However, the quality of the GNSS-levelling is not sufficient to provide a conclusive validation for the geoid models. This was also a limiting factor in geoid validation in Featherstone et al. (2018), even when tilts and biases were removed and when the Australian national levelling network (ANLN) was readjusted with corrections at tide gauges for the ocean's mean dynamic topography (MDT) (Filmer et al. 2014; Featherstone and Filmer 2012).

6 Conclusions

The object of this study was to use AUSGeoid data preparation and computation methods (Featherstone et al. 2018) to compute a geoid model over the limited-extent Colorado experiment study area for comparison with other research groups. Through this process, we have used alternative methods to adapt the AUSGeoid methods to more mountainous terrain than is encountered in Australia. Comparisons between these methods have indicated differences that are likely to be significant, although cannot be properly validated by the existing GNSS-levelling data provided for the experiment.

We used a different (1) terrain correction method (Goyal et al. 2019), (2) method for gridding the gravity data and (3) Stokes integration kernel modification. We also experimented with spherical harmonic high-pass filtering of residual anomalies before Stokes integration. Airborne gravity data were included using 3D least-squares collocation, which was something that has not previously been done in the AUSGeoid method, and experimented with different DEM filters. All modifications tested resulted in geoid differences in excess of 0.02 m (rms). However, it cannot definitively be demonstrated which modifications are improvements, and how much of an improvement the inclusion of the airborne gravity provides, because the quality of the GNSS-levelling is not sufficient to show this.

The comparison of the final three geoids computed showed a standard deviation of differences with the GNSS-levelling of ~ ±0.044 m following the removal of the tilt and bias in the NAVD88. This indicates that the AUSGeoid methods are suitable for computing accurate geoid models in mountainous regions outside Australia, albeit with some modifications. The current Colorado GNSS-levelling data are not adequate to fully assess these modifications. It is hoped that GSVS17 slope validation data will be sufficiently accurate to properly validate our experimental models.

Acknowledgements We acknowledge NGS/NOAA for making available all data sets used in the IHRS Colorado experiment. We also thank W.E. Featherstone and J.C. McCubbine for provision of software and advice on its use. In addition, thanks go to R. Goyal for assistance with the computation of terrain corrections. Finally, we acknowledge the thorough reviews of an earlier version of this manuscript by R. Pail and two anonymous reviewers. All figures were created using the Generic Mapping Tools (GMT) software (Wessel et al. 2013).

Author contributions Sten Claessens and Mick Filmer designed the research, analysed the data and wrote the paper. Sten Claessens performed the computations.

Data availability The data sets generated during and/or analysed during the current study are available from the corresponding author on reasonable request. The final terrestrial and airborne quasigeoid and geoid models will be submitted to the geoid repository of the International Service for the Geoid (IGS) for inclusion on their website: www.isgeoid.polimi.it.

References

Barnes D, Factor JK, Holmes SA, Ingalls S, Presicci MR, Beale J, Fecher T (2015) Earth gravitational model 2020. Paper presented to the AGU Fall Meeting 2015, San Francisco, USA

Brown NJ, McCubbine JC, Featherstone WE, Gowans N, Woods A, Baran I (2018) AUSGeoid2020 combined gravimetric–geometric model: location-specific uncertainties and baseline-length-dependent error decorrelation. J Geodesy 92:1457–1465. https://doi.org/10.1007/s00190-018-1202-7

Bucha B, Janák J (2014) A MATLAB-based graphical user interface program for computing functionals of the geopotential up to ultra-high degrees and orders: efficient computation at irregular surfaces. Comput Geosci 66:219–227. https://doi.org/10.1016/j.cageo.2014.02.005

Claessens SJ, Hirt C, Amos MJ, Featherstone WE, Kirby JF (2011) The NZGeoid09 model of New Zealand. Surv Rev 43:2–15. https://doi.org/10.1179/003962610X12747001420780

Driscoll JR, Healy DMJ (1994) Computing Fourier transforms and convolutions on the sphere. Adv Appl Math 15:202–250. https://doi.org/10.1006/aama.1994.1008

Ekman M (1989) Impacts of geodynamic phenomena on systems for height and gravity. Bull Géod 63:281–296

Ellmann A (2005a) Two deterministic and three stochastic modifications of Stokes's formula: a case study for the Baltic countries. J Geodesy 79:11–23. https://doi.org/10.1007/s00190-005-0438-1

Ellmann A (2005b) Computation of three stochastic modifications of Stokes's formula for regional geoid determinations. Comput Geosci 31(6):742–755. https://doi.org/10.1016/j.cageo.2005.01.008

Featherstone WE (2003) Software for computing five existing types of deterministically modified integration kernel for gravimetric geoid determination. Comput Geosci 29:183–193. https://doi.org/10.1016/S0098-3004(02)00074-2

Featherstone WE (2013) Deterministic, stochastic, hybrid and band-limited modifications of Hotine's integral. J Geodesy 87:487–500. https://doi.org/10.1007/s00190-013-0612-9

Featherstone WE, Filmer MS (2012) The north-south tilt in the Australian Height Datum is explained by the ocean's mean dynamic topography. J Geophys Res 117:C08035. https://doi.org/10.1029/2012JC007974

Featherstone WE, Kirby JF (2000) The reduction of aliasing in gravity anomalies and geoid heights using digital terrain data. Geophys J Int 141:204–212. https://doi.org/10.1046/j.1365-246X.2000.00082.x

Featherstone WE, Evans JD, Olliver JG (1998) A Meissl-modified Vaníček and Kleusberg kernel to reduce the truncation error in gravimetric geoid computations. J Geodesy 72:154–160. https://doi.org/10.1007/s001900050157

Featherstone WE, Kirby JF, Hirt C, Filmer MS, Claessens SJ, Brown NJ, Hu G, Johnston GM (2011) The AUSGeoid09 model of the Australian Height Datum. J Geodesy 85:133–150. https://doi.org/10.1007/s00190-010-0422-2

Featherstone WE, McCubbine JC, Brown NJ, Claessens SJ, Filmer MS, Kirby JF (2018) The first Australian gravimetric quasigeoid model with location-specific uncertainty estimates. J Geodesy 92:149–168. https://doi.org/10.1007/s00190-017-1053-7

Filmer MS, Featherstone WE, Kuhn M (2010) The effect of EGM2008-based normal, normal-orthometric and Helmert orthometric height systems on the Australian level network. J Geodesy 84:501–513. https://doi.org/10.1007/s00190-010-0388-0

Filmer MS, Featherstone WE, Claessens SJ (2014) Variance component estimation uncertainty for unbalanced data: application to a continent-wide vertical datum. J Geodesy 88:1081–1093. https://doi.org/10.1007/s00190-014-0744-6

Flury J, Rummel R (2009) On the geoid–quasigeoid separation in mountain areas. J Geodesy 83:829–847. https://doi.org/10.1007/s00190-009-0302-9

Foroughi I, Vaníček P, Kingdon RW, Goli M, Sheng M, Afrasteh Y, Novák P, Santos MC (2019) Sub-centimetre geoid. J Geodesy 93:849–868. https://doi.org/10.1007/s00190-018-1208-1

Forsberg R (1987) A new covariance model for inertial gravimetry and gradiometry. J Geophys Res 92(B2):1305–1310. https://doi.org/10.1029/JB092iB02p01305

Forsberg R (2002) Downward continuation of airborne gravity data. The 3rd meeting of the International Gravity and Geoid Commission 'Gravity and Geoid 2002', Thessaloniki, Greece

Forsberg R, Featherstone W (1998) Geoids and cap sizes. In: Forsberg R, Feissel M, Dietrich R (eds) Geodesy on the move. International Association of Geodesy Symposia, vol 119. Springer, Berlin

Forsberg R, Olesen AV, Einarsson I, Manandhar N, Shreshta K (2014) Geoid of Nepal from airborne gravity survey. In: Earth on the edge: science for a sustainable planet. Springer, Berlin Heidelberg, pp 521–527. https://doi.org/10.1007/978-3-642-37222-3_69

Förste C, Abrykosov O, Bruinsma S, Dahle C, König R, Lemoine J-M (2019) ESA's Release 6 GOCE gravity field model by means of the direct approach based on improved filtering of the reprocessed gradients of the entire mission. GFZ Data Serv. https://doi.org/10.5880/ICGEM.2019.004

Goyal R, Featherstone WE, Claessens SJ, Devaraju B, Balasubramania N, Dikshit O (2019) A numerical approach to the mass-prism integration for fast determination of terrain corrections. The 27th IUGG General Assembly, Montréal, Québec, Canada

GRAV-D Team (2017) Block MS05 (Mountain South 05); GRAV-D Airborne Gravity Data user manual. Monica A. Youngman and Jeffery A Johnson. BETA Available DATE. http://www.ngs.noaa.gov/GRAV-D/data_MS05.shtml

Hackney RI, Featherstone WE (2003) Geodetic versus geophysical perspectives of the 'gravity anomaly'. Geophys J Int 154:35–43. https://doi.org/10.1046/j.1365-246X.2003.01941.x

Heiskanen WA, Moritz H (1967) Physical geodesy. WH Freeman and Co, San Francisco

Huang J, Holmes SA, Zhong D, Véronneau M, Wang Y, Crowley JW, Li X, Forsberg R (2017) Analysis of the GRAV-D Airborne Gravity Data for geoid modelling. In: Vergos G, Pail R, Barzaghi R (eds) International symposium on gravity, geoid and height systems 2016. International Association of Geodesy Symposia, vol 148. Springer, Cham. https://doi.org/10.1007/1345_2017_23

Ihde J, Sánchez L, Barzaghi R, Drewes H, Förste C, Gruber T, Liebsch G, Marti U, Pail R, Sideris M (2017) Definition and proposed realization of the International Height Reference System (IHRS). Surv Geophys 38:549–570. https://doi.org/10.1007/s10712-017-9409-3

Jarvis A, Guevara E, Reuter HI, Nelson AD (2008) Hole-filled SRTM for the globe: version 4: data grid. Web publication/site, CGIAR Consortium for Spatial Information. http://srtm.csi.cgiar.org/

Jekeli C (2000) Heights, the geopotential and vertical datums. Report No. 459. The Ohio State University, Columbus, USA

Kern M, Schwarz K-P, Sneeuw N (2003) A study on the combination of satellite, airborne, and terrestrial gravity data. J Geodesy 77(3–4):217–225. https://doi.org/10.1007/s00190-003-0313-x

Li X (2018) Modeling the North American vertical datum of 1988 errors in the conterminous United States. J Geod Sci 8:1–13. https://doi.org/10.1515/jogs-2018-0001

Li X, Wang YM (2011) Comparisons of geoid models over Alaska computed with different Stokes' kernel modifications. J Geod Sci 1:136–142. https://doi.org/10.2478/v10156-010-0016-1

McCubbine JC, Featherstone WE, Kirby JF (2017) Fast Fourier-based error propagation for the gravimetric terrain correction. Geophysics 82:G71–G76. https://doi.org/10.1190/GEO2016-0627.1

McCubbine JC, Amos MJ, Tontini FC, Smith E, Winefied R, Stagpoole V, Featherstone WE (2018) The New Zealand gravimetric quasigeoid model 2017 that incorporates nationwide airborne gravity. J Geodesy 92:923–937. https://doi.org/10.1007/s00190-017-1103-1

Moritz H (2000) Geodetic reference system 1980. J Geodesy 74:128–140. https://doi.org/10.1007/s00190050278

Nagy D, Papp G, Benedek J (2000) The gravitational potential and its derivatives for the prism. J Geodesy 74:552–560. https://doi.org/10.1007/s001900000116

Pavlis NK, Holmes SA, Kenyon SC, Factor JK (2012) The development and evaluation of the Earth Gravitational Model 2008 (EGM2008). J Geophys Res Solid Earth 117:B04406. https://doi.org/10.1029/2011JB008916

Rapp RH (1989) The treatment of permanent tidal effects in the analysis of satellite altimeter data for sea surface topography. Manuscr Geod 14:368–372

Rapp RH (1997a) Use of potential coefficient models for geoid undulation determinations using a spherical harmonic representation of the height anomaly/geoid undulation difference. J Geodesy 71:282–289. https://doi.org/10.1007/s001900050096

Rapp RH (1997b) Global models for the 1 cm geoid—present status and near term prospects. In: Sansò F, Rummel R (eds) Geodetic boundary value problems in view of the one centimeter geoid. Lecture notes in earth sciences, vol 65. Springer, Berlin. https://doi.org/10.1007/bfb0011708

Rexer M, Hirt C, Pail R, Claessens SJ (2014) Evaluation of the third- and fourth-generation GOCE Earth gravity field models with Australian terrestrial gravity data in spherical harmonics. J Geodesy 88:319–333. https://doi.org/10.1007/s00190-013-0680-x

Roelse A, Granger HW, Graham JW (1971, 2nd ed. 1975) The adjustment of the Australian levelling survey 1970–1971, Technical Report 12. Division of National Mapping, Canberra, Australia

Sánchez L, Sideris M (2017) Vertical datum unification for the International Height Reference System (IHRS). Geophys J Int 209:570–586. https://doi.org/10.1093/gji/ggx025

Sánchez L, Čunderlík R, Dayoub N, Mikula K, Minarechová Z, Šíma Z, Vatrt V, Vojtíšková M (2016) A conventional value for the geoid reference potential W_0. J Geodesy 90:815–835. https://doi.org/10.1007/s00190-016-0913-x

Sandiford M, Quigley M (2009) TOPO-OZ: insights into the various modes of intraplate deformation in the Australian continent. Tectonophysics 474:405–416. https://doi.org/10.1016/j.tecto.2009.01.028

Sjöberg LE (1981) Least squares combination of satellite and terrestrial data in physical geodesy. Ann Geophys 37(1):25–30

Sjöberg LE (2003) A general model of modifying Stokes' formula and its least-squares solution. J Geodesy 77(7–8):459–464. https://doi.org/10.1007/s00190-003-0346-1

Sjöberg LE (2010) A strict formula for geoid-to-quasigeoid separation. J Geodesy 84:699–702. https://doi.org/10.1007/s00190-010-0407-1

Sjöberg LE (2013) The geoid or quasigeoid—which reference surface should be preferred for a national height system? J Geod Sci 3:103–109. https://doi.org/10.2478/jogs-2013-0013

Smith D (2007) The GRAV-D project: gravity for the Redefinition of the American Vertical Datum. http://www.ngs.noaa.gov/GRAV-D/pubs/GRAV-D_v2007_12_19.pdf

Smith DA, Milbert DG (1999) The GEOID96 high resolution geoid height model for the United States. J Geodesy 73:219–236. https://doi.org/10.1007/s001900050239

Smith WHF, Wessel P (1990) Gridding with a continuous curvature surface in tension. Geophysics 55:293–305. https://doi.org/10.1190/1.1442837

Smith DA, Holmes SA, Li X, Guillaume S, Wang YM, Bürki B, Roman DR, Damiani TM (2013) Confirming regional 1 cm differential geoid accuracy from airborne gravimetry: the geoid slope validation survey of 2011. J Geodesy 87:885–907. https://doi.org/10.1007/s00190-013-0653-0

Tenzer R, Hirt C, Claessens S, Novák P (2015) Spatial and spectral representations of the geoid-to-quasigeoid correction. Surv Geophys 36:627–658. https://doi.org/10.1007/s10712-015-9337-z

Van Westrum D (2019) Field observation results from the 2017 Geoid Slope Validation Survey in Colorado, USA. The 27th IUGG General Assembly, Montréal, Québec, Canada

Vaníček P, Featherstone WE (1998) Performance of three types of Stokes's kernel in the combined solution for the geoid. J Geodesy 72:684–697. https://doi.org/10.1007/s001900050

Vaníček P, Kleusberg A (1987) The Canadian geoid—Stokesian approach. Manuscr Geod 12:86–98

Voigt C, Denker H (2015) Validation of GOCE gravity field models in Germany. Newton's Bull 5:37–48

Wang YM, Becker C, Breidenbach S, Geoghegan C, Martin D, Winester D, Hanson T, Mader GL, Eckl MC (2015) Results of analysis of the geoid slope validation survey 2014 in Iowa. 26th General Assembly of the International Union of Geodesy and Geophysics (IUGG), Prague, Czech Republic

Wenzel H-G (1983) Geoid computation by least squares spectral combination using integral kernels. In: Proceedings of the International Association of Geodesy General Meeting, Tokyo, Japan, pp 438–453

Wessel P, Smith WHF, Scharroo R, Luis JF, Wobbe F (2013) Generic mapping tools: improved version released. EOS Trans AGU 94:409–410. https://doi.org/10.1002/2013EO450001

Wieczorek MA, Meschede M (2018) SHTools—tools for working with spherical harmonics. Geochem Geophys Geosyst 19:2574–2592. https://doi.org/10.1029/2018GC007529

Wong L, Gore R (1969) Accuracy of geoid heights from modified Stokes kernels. Geophys J Roy Astron Soc 18:81–91. https://doi.org/10.1111/j.1365-246X.1969.tb00264.x

Wu Y, Abulaitijiang A, Featherstone WE, McCubbine JC, Andersen OB (2019) Coastal gravity field refinement by combining airborne and ground-based data. J Geodesy 93:2569–2584. https://doi.org/10.1007/s00190-019-01320-3

Zilkoski DB, Richards JH, Young GM (1992) Results of the general adjustment of the North American Vertical Datum of 1988. Surv Land Inf Syst 52:133–149