Earth, Planets and Space

FULL PAPER

Open Access

High-resolution regional gravity field modeling in data-challenging regions for the realization of geopotential-based height systems

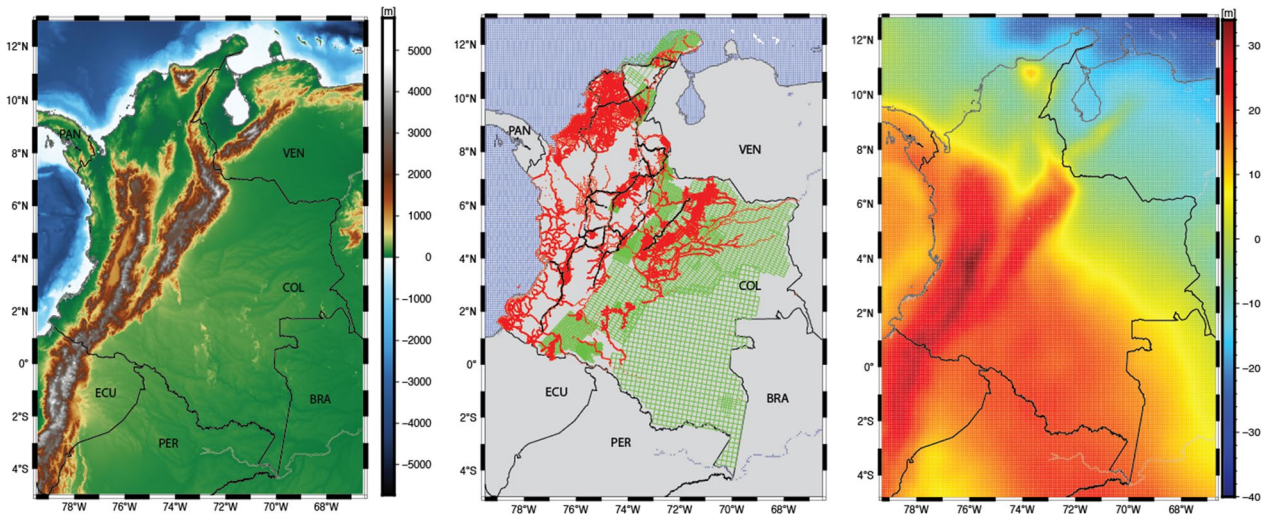
Qing Liu[1*], Michael Schmidt[1], Laura Sánchez[1], Leidy Moisés[2] and Diego Cortez[2]

Abstract Modern height systems are based on the combination of satellite positioning and gravity field models of high resolution. However, in many regions, especially developing or newly industrializing countries, there is no (reliable) regional gravity model at all, due to challenges such as limited data availability, unknown/low data quality, and missing metadata. This paper addresses this issue in a case study of Colombia, where eight decades of historical terrestrial and airborne gravity measurements are available but widely contain systematic errors, outliers, and biases. Correspondingly, processing strategies and structures are proposed and applied to validate and improve the quality of old gravity datasets. A novel method is developed based on spherical radial basis functions (SRBFs) for estimating biases, which are found in different airborne surveys with values exceeding 40 mGal. The validity of this bias estimation method is demonstrated both by a simulation test and by the evaluation of the airborne data in comparison to the SATOP (SAtellite-TOPography) model, which merges the satellite-only global gravity model GOCO06s with the Earth2014 topography model. The terrestrial and airborne data are then combined with a global gravity model (GGM), ultra-high-resolution topography models, as well as altimetry-derived gravity anomalies from DTU21GRA for the offshore areas. The results are presented in terms of height anomalies (QGeoidCOL2023), and they are thoroughly validated using GPS/leveling data both in the absolute and relative manner. The standard deviation in comparison to the GPS/leveling data after applying a correction surface to account for the datum inconsistencies amounts to 15.76 cm, which is 27% smaller compared to the mean standard deviation value given by five recent high-resolution GGMs, and 36% smaller than the one delivered by the latest South American quasi-geoid model QGEOID2021. The relative validation results show that QGeoidCOL2023 performs better, i.e., delivers lower RMS errors than the GGMs and QGEOID2021 in all the baseline length groups. These results indicate the validity and benefits of the developed methods and procedures, which can be used for other data-challenging areas to facilitate the realization of geopotential-based height systems.

Keywords Regional quasi-geoid modeling, Spherical radial basis functions, Heterogeneous gravity data combination, GPS/leveling validation, Geopotential-based height systems, International height reference system and frame

*Correspondence:
Qing Liu
qingqing.liu@tum.de
Full list of author information is available at the end of the article

Graphical Abstract



Introduction

The combination of satellite positioning techniques, such as the Global Positioning System (GPS), with high-resolution geoid or quasi-geoid models provides an accurate alternative to costly and time-consuming leveling for the determination of physical heights. In regions with advanced geodetic infrastructure, it has been demonstrated that physical heights derived from GPS and an accurate geoid or quasi-geoid model can achieve accuracies of 2–10 cm (e.g., Wang et al. 2021; Sánchez et al. 2021). In regions with less developed geodetic infrastructure, uncertainties can reach up to 40 cm, with extreme cases of about 1 m in areas with strong topographic gradients (e.g., Rummel et al. 2014; Gruber and Willberg 2019). The main accuracy limitation in these regions is the low availability or quality of surface, i.e., terrestrial, airborne, or shipborne gravity data. Given the economic constraints in some regions and the impossibility of systematically carrying out gravimetry to improve the coverage and distribution of gravity data, one of the current challenges is to recover as many existing gravity surveys as possible using modern mathematical methods that allow the evaluation and refinement of gravity data acquired long ago or lacking standard procedures and metadata.

Whereas in leveling-based height systems, the vertical datum, i.e., the zero-height level is realized by the mean sea level determined at an arbitrarily selected tide gauge, in a geopotential-based height system, the vertical datum is realized by a geoid or quasi-geoid model. Strictly speaking, the vertical coordinates are potential differences or geopotential numbers C with respect to a conventionally adopted W_0 value, the realization of which should be the geoid. Since the geoid and the quasi-geoid are practically identical in marine areas, the quasi-geoid is also widely accepted as the realization of a vertical datum. However, it is worth clarifying that the quasi-geoid is not an equipotential surface. In this study, we follow the conventions of the International Height Reference System (IHRS, Ihde et al. 2017). The IHRS is a geopotential-based reference system co-rotating with the Earth. The primary coordinates are the geopotential numbers $C = W_0 - W$ referring to the potential value $W_0 = 62636853.4\ \mathrm{m}^2/\mathrm{s}^2$ (Sánchez et al. 2016). The spatial location of the stations at which the geopotential numbers are calculated is defined by the coordinates (X, Y, Z) in the International Terrestrial Reference Frame (ITRF, e.g., Altamimi et al. 2023). The reference ellipsoid is the GRS80 (Geodetic Reference System 1980, Moritz (2000)). The realization of the IHRS is the International Height Reference Frame (IHRF), which corresponds to a global network of reference stations with precise reference coordinates specified in the IHRS (Sánchez et al. 2021). In this context, the objective of this work is to evaluate the feasibility of a precise realization of the IHRS in regions with challenging gravity data quality and distribution. The study area is Colombia, the only South American country with coastlines on both oceans, the Pacific and the Atlantic. The country features strong topographical gradients with elevation reaching more than 5000 m above mean sea level, and a large area of about 40% covered by the Amazon rainforest.

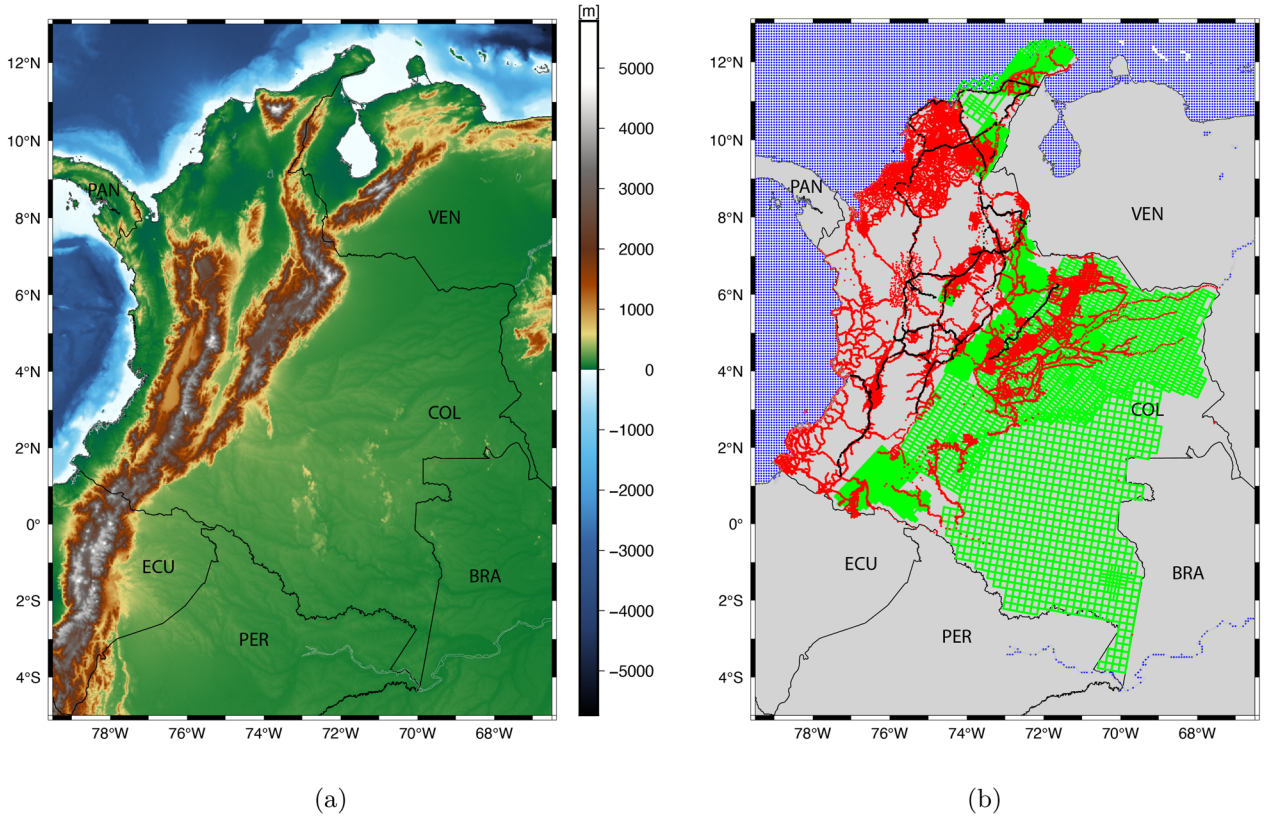


Fig. 1 **a** Terrain map of the study area; **b** available gravity data, including terrestrial (red points), airborne (green flight tracks), and altimetry (blue points) data, as well as the GPS/leveling data (black points) for validation purpose

Our research is based on the quasi-geoid determination using the method of spherical radial basis functions (SRBFs, see Freeden et al. 1998 among others). Although it is a relatively new approach compared to, e.g., the Stokes' integral and the least-squares collocation (LSC), SRBFs have become of high interest and been widely applied in regional gravity field modeling during the last two decades (e.g., Schmidt et al. 2007; Wittwer 2009; Bentel 2013; Lieb et al. 2016; Wu et al. 2017). Further advancements have also been made by different research groups, e.g., strategies for a proper choice of the SRBF settings have been proposed (e.g., Klees et al. 2008; Bentel 2013), new regularization approaches have been developed (e.g., Eicker et al. 2014; Liu et al. 2020a), and methods for a multi-resolution representation based on SRBFs for spectral combination have been advanced (e.g., Schmidt et al. 2006; Liu et al. 2022). The major advantages of the SRBFs are that they are computationally easy to implement, and allow the handling of the gravity observations directly at their original positions without applying any griding or interpolation procedures (e.g., Wittwer 2009; Li 2018). In 2017, four scientific groups under the umbrella of the International Association of Geodesy (IAG) set up the '1 cm Geoid Experiment' (Wang et al. 2021) to asses the repeatability of IHRF coordinate determination using different gravity field modeling methods in a case study area in Colorado, USA. Among 14 participating groups worldwide utilizing different methodologies, the SRBF-based quasi-geoid model (Liu et al. 2020b) delivers a standard deviation (STD) of 2.64 cm in comparison to the GPS/leveling data, which is among the five models of the best agreement.

Due to the challenging data in Colombia, novel approaches are further developed to generate a quasi-geoid model with the highest possible accuracy. To be more specific, as the earliest terrestrial surveys date back to the time before GPS, positions of these observations were frequently read from maps. Thus, systematic errors and mistakes in the data records need to be checked. The airborne data consist of 17 flight campaigns, and the available datasets have already been post-processed, i.e., reduced to a reference surface separately. We conduct a data evaluation using a SAtellite-TOPography combined model (SATOP, Zingerle et al. 2019), and the results show that each airborne survey has a distinct mean difference compared to the SATOP, with the largest value reaching

more than 40 mGal in terms of gravity disturbance. It indicates that significant biases exist in different airborne campaigns. Therefore, it is crucial to remove such large biases before the combination procedure to prevent any undesired errors. Li (2018 and 2021) proposes a bias estimation model based on the SRBFs and demonstrates that they can be used to remove biases in different airborne flight lines. By adding artificially generated biases ranging between -1 mGal and 1 mGal to the simulated airborne data, results show that the estimated bias for each flight line agrees well with the added ones. However, Li's model is based on the assumption that the mean value of biases in all flight lines equals to 0, which is not the case for different airborne surveys, and it is only verified with small bias level, i.e., within ± 1 mGal. In Colombia, we are facing significant bias values exceeding 40 mGal, and the mean value of the biases contained in different surveys is not close to 0. Thus, in this study, we additionally develop procedures for estimating biases in different airborne surveys using the method of SRBFs, and their validity is demonstrated based on both simulated and real airborne observations. The terrestrial and airborne data are then combined with a global gravity model (GGM) and topography models, which play an important role in mountainous areas, within the remove–compute–restore (RCR) procedure. In the offshore area, satellite altimetry-derived gravity data are additionally incorporated, which are obtained from the latest release of the DTU (Technical University of Denmark) gravity anomaly grid, DTU-21GRA (Anderson, personal communication, 2022).

The significance of this study lies in four aspects: (1) previous to this work, the only available regional quasi-geoid model for Colombia was calculated in 2003 (see the International Service for Geoid - ISG[1]), even before the accessibility of satellite gravity observations from GRACE (Gravity Recovery And Climate Experiment, Tapley et al. (2004)) and GOCE (Gravity field and steady-state Ocean Circulation Explorer, Rummel et al. (2002)). Consequently, its accuracy was reported to be 1.33 m in terms of STD compared to the GPS/leveling data (Sánchez and Sideris 2017), which falls short of the requirement for the realization of a geopotential-based height system or any meaningful application nowadays. Therefore, there is a pressing need for the computation of a new model for both scientific and practical purposes. (2) Colombia is a challenging study area, known for its varying topography (see Fig. 1a). The Andes runs along the western part of the country, with high elevation reaching more than 5000 m and rugged topography; the southeastern part of the country is covered by the Amazon rainforest, which is nearly impossible for any ground gravity survey to take place. (3) We process and make use of all types of gravity data available currently, collected during eight decades. Satellite gravimetry and satellite altimetry data are combined with regional terrestrial and airborne data, and the very short wavelength parts are considered by including up-to-date topography models. Furthermore, the computed quasi-geoid model is thoroughly validated with independent GPS/levelling data. (4) The computation procedures developed in this study to deal with poor- and unknown-quality gravity data have significant practical implications and can be applied to other developing regions for the realization of the IHRS or any geopotential-based height system.

This work is organized as follows: in 'Study area and data processing' Section, we present the available data within the study area and explain the different data preprocessing procedures applied to the terrestrial and airborne observations in detail. Section 'Methodology' is dedicated to the methodologies, where the fundamentals of the SRBFs and the parameter estimation are introduced briefly. Furthermore, the developed bias estimation method based on the SRBFs is explained. Section 'Computation configuration' documents the computation procedure, including the airborne data bias estimation, the RCR procedure, and the model configuration. Section 'Results, validation, and discussion' presents the computed quasi-geoid model for Colombia and provides a thorough validation of it in comparison to both the GPS/leveling data and recent high-resolution GGMs. Finally, Section 'Conclusion' summarizes the findings and provides some conclusions and recommendation.

Study area and data processing

The study area covers Colombia and parts of its surrounding onshore and offshore areas, i.e., between $-79.5°$ and $-66.5°$ longitude and between $-5°$ and $13°$ latitude. Figure 1a displays the terrain map of the study area; the eastern part is characterized by relatively flat terrain. In the western part, the Andes spans across the entire study area in a north–south direction near the coast, contributing to the complex terrain, which features depths reaching around 5000 m below sea level and elevations rising to more than 5000 m above it. Such rugged terrain and thus, varying gravity field makes the quasi-geoid determination challenging. Figure 1b visualizes the available gravity data; terrestrial gravity observations since the year 1940, 17 airborne gravity surveys since 2000, as well as GPS/leveing data are provided by the National Mapping Agency (Instituto Geográfico Agustín Codazzi—IGAC), the Colombian Geological Survey, and the oil company Ecopetrol.

[1] International Service for Geoid (ISG): https://www.isgeoid.polimi.it/Geoid/America/Colombia/colombia04_g.html.

Terrestrial gravity data

Terrestrial data (red points in Fig. 1b) in Colombia stem from 101 gravity surveys with altogether 65,763 observation points that were conducted between the years 1941 and 2000 for oil and gas exploration as well as geophysical studies. These surveys took place mainly in the western part of Colombia and their distribution is not homogeneous. Terrestrial data are provided in terms of absolute gravity values along with latitude φ_{local} and longitude λ_{local} referring to the local geodetic datum, and physical height H referring to the local vertical datum. Although the exact measurement accuracy is unknown due to the unavailability of metadata, it is not expected to be high considering that the terrestrial data were collected by different surveys throughout six decades in the last century. The horizontal and vertical positions of older observations were frequently read from maps. Thus, systematic errors and wrong records need to be checked. The following pre-processing steps are conducted to the terrestrial data:

(1) The horizontal coordinates of the terrestrial data collected in Colombia refer to the local geodetic datum (Datum Bogotá), which presents displacements of up to 500 m w.r.t. the ITRF. This difference can cause a vertical dislocation of several hundred meters in the Andes, and correspondingly, the topographic effects may be misrepresented by large systematic effects (Sánchez et al. 2021). Therefore, it is necessary to perform a coordinate transformation from the local datum (φ_{local}, λ_{local}) to the ITRF (φ, λ)

$$\varphi = \varphi_{\text{local}} - 0.00278^{\circ} \tag{1a}$$

$$\lambda = \lambda_{\text{local}} - 0.00333^{\circ}, \tag{1b}$$

see IGAC (2004) for more details.

(2) Duplicate values in the terrestrial data, i.e., several gravity observations located at the same position are checked (Liu et al. 2020b). To be more specific, in such cases, if the observations with the same position differ less than 2 mGal from each other, only one record of these observations is kept. Otherwise, if the differences reach more than 2 mGal, all of them are removed. This step results in a deletion of 2692 points.

(3) Since a large portion of the measurements date back to the time before GPS and their horizontal and vertical positions are manually taken from maps, there might be misread records. To check for potential mistakes in the given height values as well as mismatches between the vertical and the horizontal positions, the physical heights H are compared to the heights obtained from the high-resolution terrain model SRTM (Hirt et al. 2014) at each observation point. Large height differences $\Delta H = H - H_{\text{SRTM}}$ indicate that either the given height values are problematic or they do not match with the horizontal coordinates, and thus, such observation points should be excluded. Data points are grouped into two categories, namely mountainous area (with elevation $\geq$ 2000 m) and moderate area (with elevation $<$ 2000 m). For each category, the three-sigma rule using the mean and STD (σ) of the ΔH is applied to check for anomaly values in the height differences. This step results in a deletion of 1177 data points.

(4) Transformation of the physical heights H in the terrestrial data to ellipsoidal heights h using the latest South American quasi-geoid model QGEOID2021 (Matos et al. 2021) available at the ISG. The reason for using a local geoid model instead of a high-resolution GGM for the height transformation is to keep consistency with the physical heights, which refer to the local mean sea level.

(5) As the terrestrial measurements were conducted in the Postdam gravity datum, a constant of 14 mGal is subtracted from the gravity value to transform them into the International Gravity Standardization Net 1971 (IGSN71) datum (Morelli et al. 1971).

(6) Transfer the observations in terms of absolute gravity g to gravity disturbance δg by subtracting the normal gravity γ at the observation site referring to the ellipsoid GRS80 (Moritz 2000), i.e., $\delta g = g - \gamma$ (Heiskanen and Moritz 1967).

(7) Outlier detection: after applying the RCR procedure (see Sect. 'Remove–compute–restore' for more details) with a GGM and a topography model, the remaining parts of the terrestrial gravity disturbances $\Delta\delta g$ are checked for outliers following the procedures: (a) first, the three-sigma rule using the median (Mdn) and normalized median of absolute deviations (NMAD) of the $\Delta\delta g$ is applied for identifying potential outliers. To be more specific, data points that have $\Delta\delta g$ larger than $[\text{Mdn} + 3 \cdot \text{NMAD}]$ or smaller than $[\text{Mdn} - 3 \cdot \text{NMAD}]$ are marked as potential outliers, and all the other points are identified as non-outliers. The reason for using the median instead of the mean is that it is less sensitive to outliers and more suitable for non-normalized datasets such as gravity observations (see Varga et al. 2021 for more details). (b) For those data points that are marked as potential outliers, their gravity disturbance values $\Delta\delta g$ are checked using the non-outlier observa-

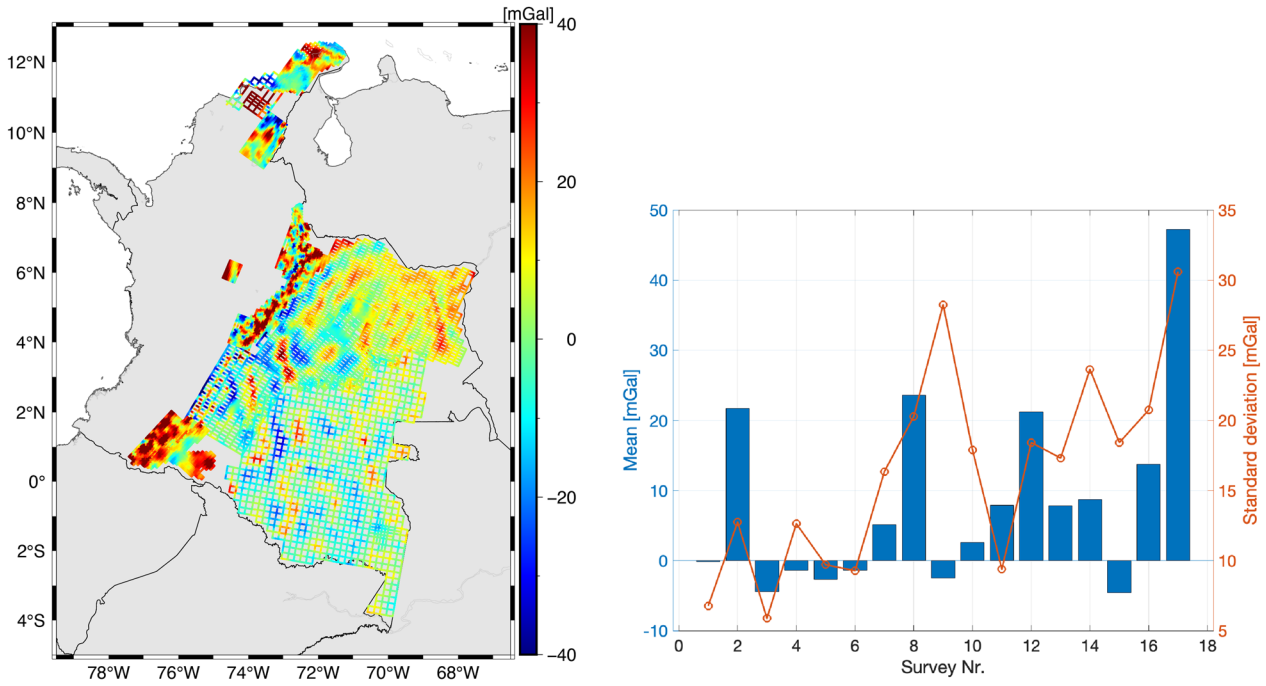


Fig. 2 **a** Differences between the airborne gravity data and the SATOP model in terms of gravity disturbance; **b** the mean values and standard deviations of the differences for each airborne survey

tions. In particular, for each potential outlier data point, a window of 10 km × 10 km is applied, and all the non-outlier points within this window are used to interpolate $\Delta\delta g^{\text{intp}}$ at this point, following the inverse distance weighting (IDW) interpolation algorithm. After this step, an interpolated value $\Delta\delta g^{\text{intp}}$ is obtained for each potential outlier. It is worth mentioning that there are other methods for outlier detection as well, e.g., interpolating the gravity values from the neighboring data records using the LSC (see e.g., Vergos et al. 2005), which is suitable especially in study cases where a proper estimation of the measurement errors is possible. (c) The value of $\Delta\delta g^{\text{intp}}$ is compared with $\Delta\delta g$ at each potential outlier point, and those with differences larger than a threshold ε are finally flagged as outliers and removed from the dataset. The threshold value ε is selected as the 95% percentiles of the cumulative distribution function (CDF) of the absolute values of all $\Delta\delta g$ in the dataset (Varga et al. 2021), which amounts to 19 mGal for the terrestrial data. A total amount of 570 observation points is flagged as outliers and excluded.

Airborne gravity data

The airborne gravity data (green flight tracks in Fig. 1b) consist of 17 airborne campaigns conducted between 2005 and 2010, with mean flight altitudes ranging from 1200 m to 6400 m. They mainly cover the eastern part of Colombia, which compensate the absence of terrestrial data in this region, especially in the Amazon rainforest. The airborne surveys also fill parts of the data gaps caused by the terrestrial campaigns in the northern mountainous area, i.e., between latitude 9° N and around 12° N. The survey above the Amazon rainforest has a lower along-track spatial resolution of around 100 m, compared to the other 16 surveys, which all reach around 50 m. The cross-track resolution ranges from 1 km to 20 km, but averagely amounts about 10 km. Airborne gravity data are given in terms of free-air anomalies along with latitude and longitude in the ITRF, ellipsoidal height, and the applied free-air correction F. As the airborne data have been post-processed and the corresponding metadata are not available, inconsistencies might have been introduced in different campaigns. Thus, knowledge about the accuracy of individual airborne surveys as well as possible systematic effects is required before using the airborne data for the quasi-geoid computation (Zingerle

et al. 2019). Accordingly, the following pre-processing steps are made:

(1) Recover gravity values g at the observation points from the free-air anomaly Δg_F by

$$g = \Delta g_F + \gamma_0 - F, \tag{2}$$

where γ_0 is the normal gravity at the reference ellipsoid (Heiskanen and Moritz 1967). Then the gravity value g is transferred to gravity disturbance δg by subtracting the GRS80 normal gravity γ at the observation height.

(2) To have an insight into the data accuracy, a cross-over analysis (e.g., Forsberg and Olesen 2010) is performed to each airborne survey. The gravity differences at the crossover points of the intersecting flight tracks reflect the data quality of airborne surveys (Smith et al. 2013). In this study, the RMS error of the crossover discrepancies obtained from the 17 surveys is estimated to range from 0.23 mGal to 2.40 mGal with a mean value of 0.67 mGal, which suggests that the 17 airborne surveys have different data accuracy.

(3) An along-track Gaussian low-pass filter (see Willberg et al. 2020) is applied to the airborne data to reduce high-frequency noise, which is a standard procedure in airborne data processing. Forsberg and Olesen (2010) point out that all types of airborne observations need filtering, and low-pass filters should be applied because in the low-frequency part of the airborne measurements, gravity signal dominates the noise (Childers et al. 1999).

(4) To check for possible systematic errors, the airborne data are then evaluated in comparison to the SATOP model (Zingerle 2022). It merges the satellite-only global gravity model GOCO06s (Kvas et al. 2021), which is very accurate in the low-frequency part and the Earth2014 topography model (Rexer et al. 2016). SATOP is chosen for evaluation due to two reasons: 1. as pointed out by Zingerle et al. (2019), the model used for data evaluation has to be truly independent from any of the observations being evaluated; 2. with a total amount of 1,775,140 airborne observations to be evaluated, it is time-consuming to compute gravity values directly from a GGM. In contrast, SATOP allows for dealing with large data sizes efficiently. Figure 2a shows the differences between the airborne gravity data and the SATOP model (limited to d/o 719 as suggested by Zingerle et al. (2019)). The differences have a mean value of 8.38 mGal and an STD of 16.70 mGal. The airborne gravity data, especially some campaigns, show large mean differences in comparison to SATOP, which reveals long-wavelength errors, i.e., biases in these airborne surveys. Figure 2b takes a closer look into the differences of each airborne survey w.r.t. SATOP and lists their mean values and standard deviations. Note that the survey number is sorted by the elevation of each flight area in an ascending order. Most of the surveys show standard deviations of less than 20 mGal w.r.t. the SATOP model, which is acceptable. Furthermore, the standard deviation shows a pattern of increasing with increased elevation. This is reasonable because the SATOP model above d/o 300 relies purely on the topography model using a constant mass density value, which cannot represent the true high-frequency gravity signal accurately, especially in mountainous areas. However, the mean values of the differences are pretty large, ranging from −4.54 mGal to 47.32 mGal for different surveys. It suggests that large biases exist in the airborne surveys, since the satellite-only gravity model is accurate in the long-wavelength part. A possible reason for these large biases could be the post-processing of each airborne campaign by different organizations at different times. These biases in the airborne data are expected to have a significant impact on the determined height anomaly. For example, Varga et al. (2021) show that a bias of 3 mGal in the airborne data can cause a bias of up to 30 cm in the calculated geoid. Thus, these biases need to be handled and removed to obtain meaningful modeling results. A bias estimation method based on the SRBFs is presented in 'Bias estimation using SRBFs' section, and the corresponding results are reported in 'Airborne data bias' section.

(5) As the airborne data have a very dense distribution with 1,775,140 observation points, using the whole dataset will result in a design matrix with a size of 399 GB (see 'Estimation model' section). Furthermore, consecutive airborne observations are highly correlated after low-pass filtering, which allows a significant reduction of the sampling frequency (Willberg et al. 2020). Thus, we down-sample the airborne data to an average along-track spatial resolution of approximately 5 km, which is a good balance between maintaining gravity information and improving computational efficiency, and it is a common choice in quasi-geoid modeling (see e.g., Wang et al. 2021). For the Amazon survey, to obtain a spatial resolution of 5 km from 100 m, only one observation of a fifty-observation block is kept, i.e., the sampling interval is reduced from 1 Hz to 1/50

Hz. For the other 16 surveys, one observation of a hundred-observation block is kept to get a spatial resolution of 5 km from 50 m.

(6) Outlier detection: after conducting bias correction and down-sampling to the airborne data, an outlier detection is also done in the same manner as for the terrestrial data, i.e., as explained in Step 7 of the terrestrial data pre-processing (see 'Terrestrial gravity data' section). A total amount of 335 observation points is flagged as outliers and excluded correspondingly.

Altimetry data

In the offshore area, gravity data derived from altimetry missions, namely the latest release of the DTU gravity anomaly model DTU21GRA (Anderson, personal communication, 2022), are used. It additionally includes 5 years of Sentinel-3A, 3 years of Sentinel-3B as well as reprocessed CryoSat-2 data, compared to its previous release DTU17GRA, which was developed by including satellite altimetry data over the ocean from the missions ERS-1/2, Envisat, CryoSat-2, Jason-1, and SARAL/AltiKa. For the DTU gravity anomaly series, the newer releases typically show better accuracy compared to the earlier versions when comparing to regional airborne or shipborne gravity data (Andersen and Knudsen 2020). Wu et al. (2022) demonstrate that the inclusion of the Sentinal-3A/B SAR altimetry data in DTU21GRA improves the accuracy in marine quasi-geoid modeling, especially over the regions close to continents and island. Their results further show that this recently published altimetric gravity model contains additional signals that were unresolved in the currently available high-resolution GGMs. Thus, the $2' \times 2'$ DTU21GRA gravity anomaly grid (blue points in Fig. 1b) is used as gravity data for the offshore area in this study.

Validation data

GPS/leveling data in Colombia are used for the quasi-geoid validation, which consist of 3209 GPS/leveling benchmarks (black points in Fig. 1b). These data are majorly located in the western part of Colombia, coinciding with the distribution of the terrestrial observations along the main roads of the country. Given measurement quantities include ellipsoidal heights h, normal heights $H^\star$ referring to the local height reference system, and correspondingly height anomalies $\zeta = h - H^\star$. These data were measured in the last decades and their accuracy is unknown. Thus, outliers in the height anomaly values are checked through a three-sigma rule using the median, i.e., in analogy to Step 7 of the terrestrial data pre-processing, by comparing to QGeoid2021. A remaining of 3025 data points are used for validation.

Methodology

Spherical radial basis function

SRBFs are an appropriate tool for regional gravity field modeling to consider the heterogeneity of different data sources, due to their localizing features. In general, an SRBF $B(\boldsymbol{x}, \boldsymbol{x}_k)$ between an observation point P and a grid point P_k on a sphere Ω_R with radius R is defined by the Legendre series

$$B(\boldsymbol{x}, \boldsymbol{x}_k) = \sum_{n=0}^{n_{max}} \frac{2n+1}{4\pi} \left(\frac{R}{r}\right)^{n+1} B_n P_n(\boldsymbol{r}^T \boldsymbol{r}_k) \tag{3}$$

(Freeden et al. 1998; Schmidt et al. 2007), where $\boldsymbol{x} = r \cdot \boldsymbol{r}$ is the position vector of the observation point $P(\varphi, \lambda, r)$, with $\boldsymbol{r} = [\cos\varphi \cos\lambda, \cos\varphi \sin\lambda, \sin\varphi]^T$ being the corresponding unit vector. $\boldsymbol{x}_k = R \cdot \boldsymbol{r}_k$ is the position vector of the grid point P_k. P_n is the Legendre polynomial of degree n, and n_{max} is the maximum degree of the expansion. B_n are the Legendre coefficients which specify the shape of the SRBFs. In case of $B_n = 1$ for all degree values $n = 0, 1, \ldots, n_{max}$, the SRBF is the Shannon kernel.

A gravity observation $y(\boldsymbol{x})$ can be represented as a series expansion of the SRBFs

$$y(\boldsymbol{x}) + e(\boldsymbol{x}) = \sum_{k=1}^{K} d_k B(\boldsymbol{x}, \boldsymbol{x}_k), \tag{4}$$

where K and d_k are the number of basis functions and the corresponding series coefficients, respectively. $e(\boldsymbol{x})$ includes the observation error and truncation error. For describing different gravitational functionals, e.g., the gravity disturbance δg and the gravity anomaly Δg, the general expression of the basis functions $B(\boldsymbol{x}, \boldsymbol{x}_k)$, Eq. (3), needs to be adapted. A list of adapted basis functions can be found in, e.g., Koop (1993) and Liu et al. (2020a).

Estimation model

Based on Eq. (4), a Gauss–Markov model can be set up to estimate the unknown coefficients

$$\boldsymbol{y}_p + \boldsymbol{e}_p = \boldsymbol{A}_p \boldsymbol{d} \;\text{ with }\; \Sigma_{y_p} = \sigma_p^2 \boldsymbol{P}_p^{-1}, \tag{5}$$

where $\boldsymbol{y}_p$ is the $N_p \times 1$ observation vector of the p^{th} gravity dataset with $p = 1, 2, \ldots, P$, and $\boldsymbol{A}_p$ is the $N_p \times K$ design matrix, which contains the corresponding (adapted) scaling functions. $\boldsymbol{d} = [d_1, d_2, \ldots, d_K]^T$ is the $K \times 1$ vector of the unknown coefficients. Σ_{y_p} is the $N_p \times N_p$ covariance matrix of the observation vector $\boldsymbol{y}_p$, with σ_p^2 being the unknown variance factor and $\boldsymbol{P}_p$ being its positive definite weight matrix.

However, the associated normal equation system is usually ill-posed (see, e.g., Liu (2023) for more details). To solve this problem, the Tikhonov regularization is

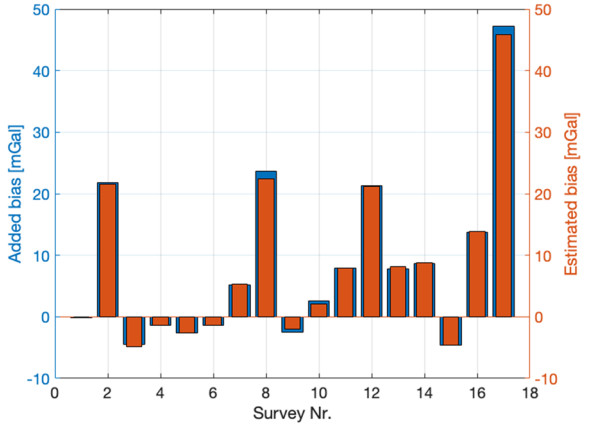


Fig. 3 Comparison between the added bias and the estimated bias for each airborne survey in a simulation test

applied. The expectation vector $\boldsymbol{\mu}_d$ of the $K \times 1$ coefficient vector $\boldsymbol{d}$ is introduced as prior information, and an additional linear model can be formulated as

$$\boldsymbol{\mu}_d + \boldsymbol{e}_d = \boldsymbol{d} \quad \text{with} \quad \Sigma_{\mu_d} = \sigma_\mu^2 \boldsymbol{P}_\mu^{-1}, \tag{6}$$

Σ_{μ_d} is the $K \times K$ covariance matrix of the prior information, with σ_μ^2 being the corresponding unknown variance factor and $\boldsymbol{P}_\mu$ being its given positive definite weight matrix.

The coefficient vector $\boldsymbol{d}$ and its covariance matrix can then be estimated as

$$\widehat{\boldsymbol{d}} = \left(\sum_{p=1}^{P} \left(\frac{1}{\sigma_p^2} \boldsymbol{A}_p^T \boldsymbol{P}_p \boldsymbol{A}_p \right) + \frac{1}{\sigma_\mu^2} \boldsymbol{P}_\mu \right)^{-1} \left(\sum_{p=1}^{P} \left(\frac{1}{\sigma_p^2} \boldsymbol{A}_p^T \boldsymbol{P}_p \boldsymbol{y}_p \right) + \frac{1}{\sigma_\mu^2} \boldsymbol{P}_\mu \boldsymbol{\mu}_d \right) \tag{7a}$$

$$\widehat{\Sigma}_d = \left(\sum_{p=1}^{P} \left(\frac{1}{\sigma_p^2} \boldsymbol{A}_p^T \boldsymbol{P}_p \boldsymbol{A}_p \right) + \frac{1}{\sigma_\mu^2} \boldsymbol{P}_\mu \right)^{-1}. \tag{7b}$$

The variance factors σ_p^2 and σ_μ^2 are determined by the variance component estimation (VCE, Koch and Kusche (2002)). The diagonal elements of the covariance matrix $\widehat{\Sigma}_d$, i.e., the variances, give information about the uncertainty of the estimated coefficients.

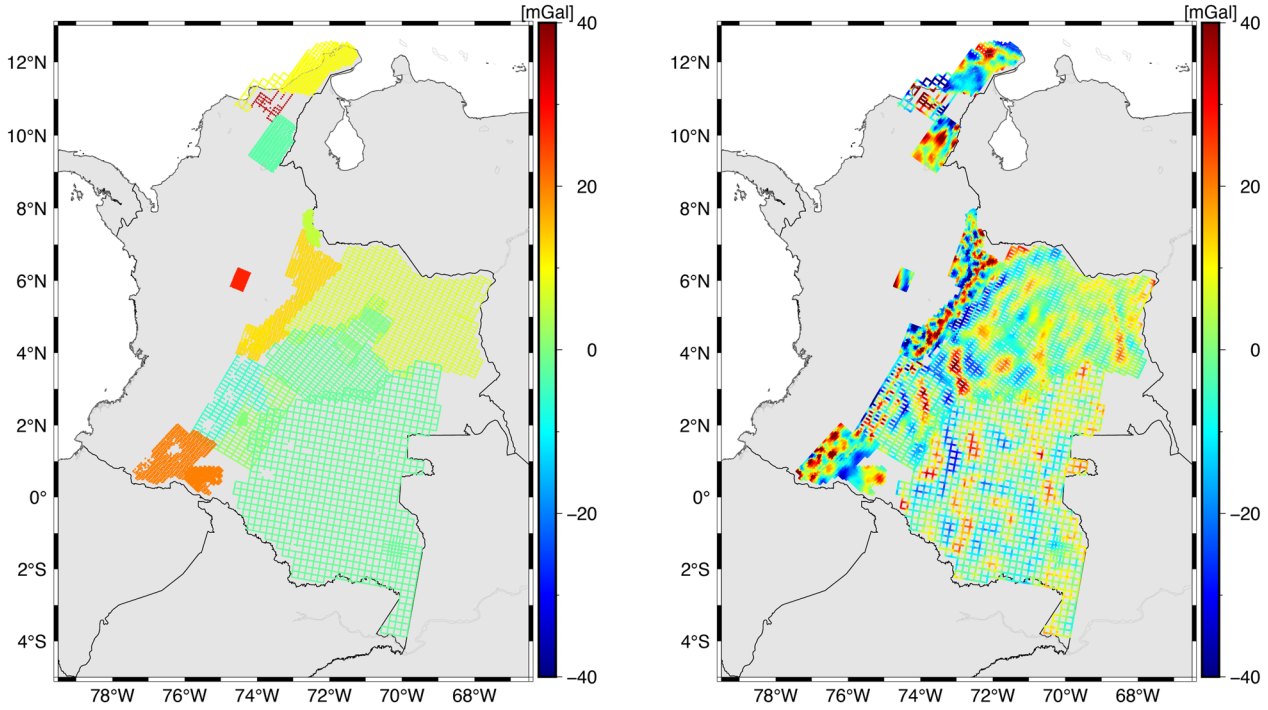


Fig. 4 **a** The estimated bias for each airborne survey, i.e., the differences before and after removing the estimated biases in each campaign; **b** differences between the airborne gravity data after removing the estimated bias in each survey and the SATOP model

Bias estimation using SRBFs

As shown in Sect. 'Airborne gravity data', there are varying biases in the airborne surveys. Nevertheless, the first survey (Nr. 1) shows a very small mean difference of only −0.14 mGal w.r.t. SATOP (see Fig. 2b). Furthermore, this survey took place in a flat area, with a maximum elevation of 230 m. As a consequence, the airborne measurements as well as the SATOP model both agree quite well with different high-resolution GGMs in terms of mean gravity disturbance values. Thus, this survey is assumed to be bias-free and set to be the reference for estimating the biases of the other 16 surveys. Correspondingly, the bias estimation model of each airborne survey q is set up

$$\begin{cases} \boldsymbol{y}_1 + \boldsymbol{e}_1 = \boldsymbol{A}_1 \boldsymbol{d} & \text{with } \Sigma_{y_1} = \sigma_1^2 \boldsymbol{P}_1^{-1} \\ \boldsymbol{y}_q + \boldsymbol{e}_q = \boldsymbol{A}_q \boldsymbol{d} + \mathbf{1}_q \Delta_{q|1} (q = 2, 3, \ldots, Q) & \text{with } \Sigma_{y_q} = \sigma_q^2 \boldsymbol{P}_q^{-1} \end{cases} \tag{8}$$

with $\mathbf{1}_q = [1, 1, \ldots, 1]^T$ being the $N_q \times 1$ vector of ones. The bias term $\Delta_{q|1}$ can be estimated together with the coefficient vector $\boldsymbol{d}$ by setting up a corresponding Gauss–Markov model

$$\begin{aligned} &\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \boldsymbol{y}_3 \\ \vdots \\ \boldsymbol{y}_Q \\ \boldsymbol{\mu}_d \end{bmatrix} + \begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \boldsymbol{e}_3 \\ \vdots \\ \boldsymbol{e}_Q \\ \boldsymbol{e}_\mu \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{A}_1 & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ \boldsymbol{A}_2 & \mathbf{1}_2 & \mathbf{0} & \ldots & \mathbf{0} \\ \boldsymbol{A}_3 & \mathbf{0} & \mathbf{1}_3 & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{A}_Q & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{1}_Q \\ \boldsymbol{I} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{d} \\ \Delta_{2|1} \\ \Delta_{3|1} \\ \vdots \\ \Delta_{Q|1} \end{bmatrix} \text{ with } D\left(\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \boldsymbol{y}_3 \\ \vdots \\ \boldsymbol{y}_Q \\ \boldsymbol{\mu}_d \end{bmatrix} \right) \\ &= \begin{bmatrix} \sigma_1^2 \boldsymbol{P}_1^{-1} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_2^2 \boldsymbol{P}_2^{-1} & \mathbf{0} & \vdots & \vdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_3^2 \boldsymbol{P}_3^{-1} & \vdots & \vdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \sigma_Q^2 \boldsymbol{P}_Q^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \sigma_\mu^2 \boldsymbol{P}_\mu^{-1} \end{bmatrix}. \end{aligned} \tag{9}$$

Set $\boldsymbol{A}_1^\star = [\boldsymbol{A}_1, \mathbf{0}, \mathbf{0}, \ldots, \mathbf{0}]$, $\boldsymbol{A}_2^\star = [\boldsymbol{A}_2, \mathbf{1}_2, \mathbf{0}, \ldots, \mathbf{0}]$, ..., $\boldsymbol{A}_Q^\star = [\boldsymbol{A}_Q, \mathbf{0}, \mathbf{0}, \ldots, \mathbf{1}_Q]$ to be the $N_q \times (K + Q - 1)$ new design matrix for the q^{th} airborne survey, and $\boldsymbol{A}_\mu = [\boldsymbol{I}, \mathbf{0}, \mathbf{0}, \ldots, \mathbf{0}]$ to be the $K \times (K + Q - 1)$ design matrix for the prior information $\boldsymbol{\mu}_d$, which contains a $K \times K$ identity matrix $\boldsymbol{I}$ and a $K \times (Q - 1)$ zero matrix $\mathbf{0}$, the $(K + Q - 1) \times 1$ unknown vector $\boldsymbol{c} = [\boldsymbol{d}^T, \Delta_{2|1}, \Delta_{3|1}, \ldots, \Delta_{Q|1}]^T$ can be estimated as

$$\widehat{\boldsymbol{c}} = \left(\sum_{q=1}^{Q} \left(\frac{1}{\sigma_q^2} \boldsymbol{A}_q^{\star T} \boldsymbol{P}_q \boldsymbol{A}_q^\star \right) + \frac{1}{\sigma_\mu^2} \boldsymbol{A}_\mu^T \boldsymbol{P}_\mu \boldsymbol{A}_\mu \right)^{-1} \left(\sum_{q=1}^{Q} \left(\frac{1}{\sigma_q^2} \boldsymbol{A}_q^{\star T} \boldsymbol{P}_q \boldsymbol{y}_q \right) + \frac{1}{\sigma_\mu^2} \boldsymbol{A}_\mu^T \boldsymbol{P}_\mu \boldsymbol{\mu}_d \right) \tag{10a}$$

$$\widehat{\Sigma}_c = \left(\sum_{q=1}^{Q} \left(\frac{1}{\sigma_q^2} \boldsymbol{A}_q^{\star T} \boldsymbol{P}_q \boldsymbol{A}_q^\star \right) + \frac{1}{\sigma_\mu^2} \boldsymbol{A}_\mu^T \boldsymbol{P}_\mu \boldsymbol{A}_\mu \right)^{-1}. \tag{10b}$$

Again, the variance factors σ_q^2 and σ_μ^2 can be determined by VCE.

Computation configuration

Airborne data bias

The performance of the bias estimation model using SRBFs, i.e., Eq. (9) is first tested using simulated data. For this purpose, airborne gravity disturbances are calculated from XGM2019 (Zingerle et al. 2020) up to d/o 5480 at the actual position of the 17 airborne surveys in Colombia, and 1 mGal random noise is added. An artificial bias with the value being the mean difference between the airborne observations and the SATOP model (Fig. 2b) is added to each airborne survey. Then, Eq. (9) is applied to estimate the bias. Figure 3 shows the difference between the added and the estimated bias of each survey. It is clear that the estimated bias values match the added ones very well. The RMS error of the differences between the added and estimated biases is 0.51 mGal, which is rather small considering the very large added bias values with a mean of 8.39 mGal and an STD of 13.69 mGal. These results demonstrate the validity of the developed bias estimation procedure using the SRBFs.

After evaluating the performance of the estimation model, it is applied to the real airborne measurements in Colombia for quantifying the bias in each survey. Figure 4a illustrates the biases estimated for each airborne survey, the values range from −5.67 mGal to 35.27 mGal with a mean of 8.74 mGal and an STD of 11.50 mGal. These biases are then removed from the airborne data before they are combined with the terrestrial and altimetry data for the quasi-geoid modeling. Figure 4b shows the comparison between the airborne gravity data and the SATOP model in terms of gravity disturbance values, after removing the estimated biases. The differences have a mean value of 0.31 mGal and an STD of 14.65 mGal. Compared to Fig. 2, the differences w.r.t. SATOP become much smaller, which demonstrate the importance of this bias estimation procedure. However, it is worth mentioning that the differences between the gravity data and the

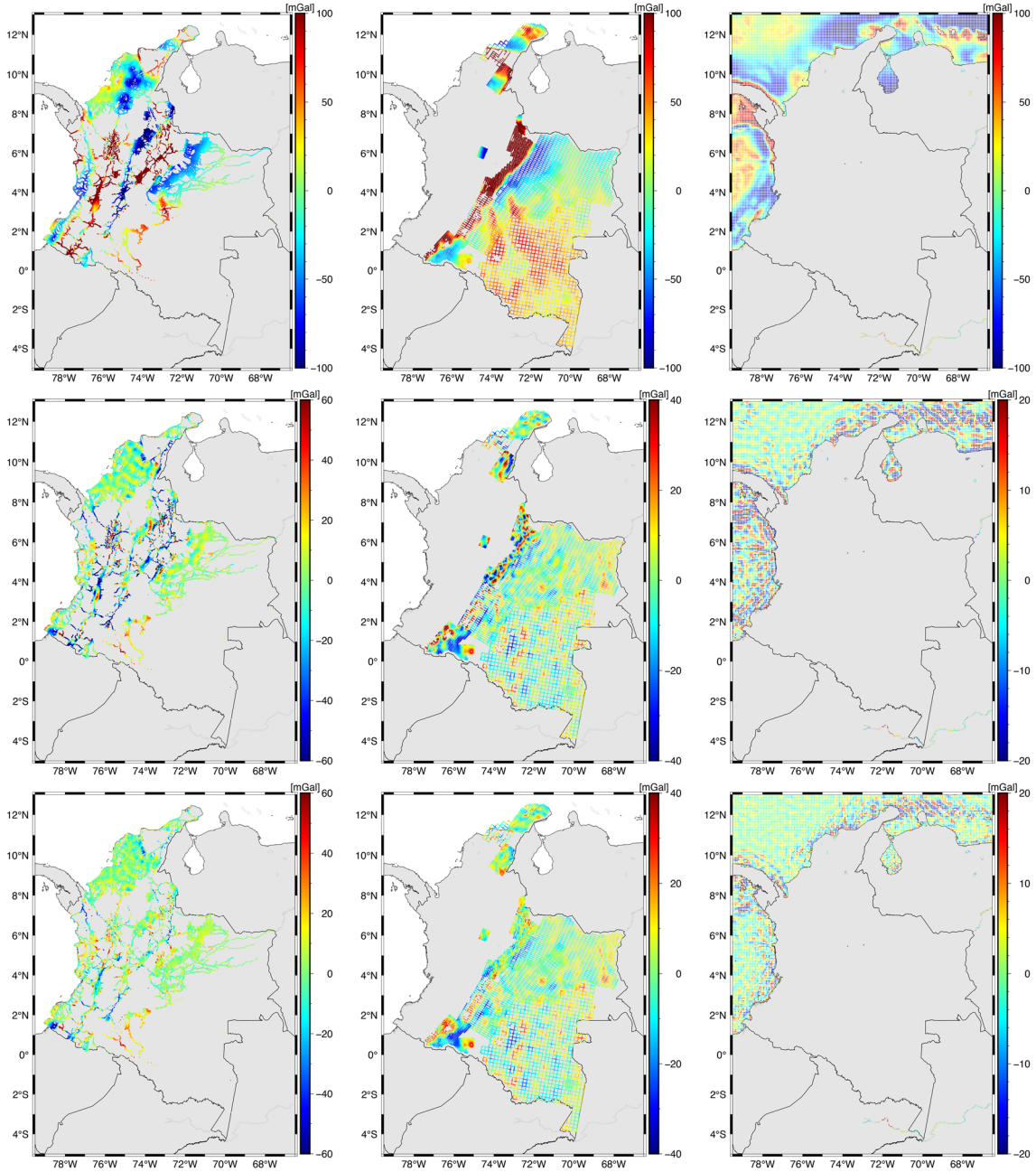


Fig. 5 The observations δg (first row), the remaining part after removing the GGM $\delta g - \delta g_{\mathrm{GGM}}$ (second row), and after removing both the GGM and topography model $\delta g - \delta g_{\mathrm{GGM}} - \delta g_{\mathrm{topo}}$ (third row), for the terrestrial data (first column), airborne (second column), and altimetry (right column) data. Note for the last two rows, different colorbars are used

SATOP model also contain the omission error of SATOP, which is spectral-limited to d/o 719, as well as additional signals in the gravity data that cannot be represented by Earth2014, which assumes the topographic masses to have constant density and does not represent the true gravity signal accurately (Hirt et al. 2010; Bucha et al. 2016). Thus, our goal is not to have the airborne data as close to SATOP as possible, but to use it for checking the long-wavelength bias in the airborne data.

Remove–compute–restore

The remove–compute–restore (RCR) procedure

$$\Delta\delta g = \delta g - \delta g_{\mathrm{GGM}} - \delta g_{\mathrm{topo}}, \tag{11}$$

Table 1 The statistics of the observations, the remaining part after removing the GGM, and after removing both the GGM and topography model for each data type (unit [mGal])

		Min	Max	Mean	STD
Terrestrial	δg	− 295.02	370.49	− 1.25	71.38
	$\delta g - \delta g_{\mathrm{GGM}}$	− 222.36	141.28	− 3.91	23.82
	$\delta g - \delta g_{\mathrm{GGM}} - \delta g_{\mathrm{topo}}$	− 205.77	128.24	0.70	13.27
Airborne	δg	− 105.81	523.90	32.43	66.58
	$\delta g - \delta g_{\mathrm{GGM}}$	− 58.79	84.45	− 0.08	12.98
	$\delta g - \delta g_{\mathrm{GGM}} - \delta g_{\mathrm{topo}}$	− 34.25	34.35	− 0.01	9.96
Altimetry	Δg	− 208.52	257.34	− 15.12	53.04
	$\Delta g - \Delta g_{\mathrm{GGM}}$	− 77.66	78.96	− 0.14	11.23
	$\Delta g - \Delta g_{\mathrm{GGM}} - \Delta g_{\mathrm{topo}}$	− 47.83	51.60	0.01	7.24

is applied, where δg is the gravity observations, δg_{GGM} is the long-wavelength component from a GGM, and δg_{topo} represents the high-frequency topographic effects from a topography model. $\Delta\delta g$ is the remaining part after the remove step, which serve as input for the estimation model, Eq. (5). Note that in case of the altimetry data (DTU21GRA), the gravity observable is given in terms of gravity anomaly Δg instead of the gravity disturbance δg. In this study, the long-wavelength component δg_{GGM} is computed from XGM2019 (Zingerle et al. 2020) up to d/o 719 for all three types of gravity observations, namely the terrestrial, airborne, and altimetry data. Regarding δg_{topo}, the topography model dV_ELL_Earth2014 (Rexer et al. 2016) from d/o 720 to 2159 and a residual terrain model ERTM2160 (Hirt et al. 2014) from degree 2160 to around 80,000, which corresponds to a spatial resolution of 250 m, are used for the terrestrial data; the dV_ELL_Earth2014 from degree 720 to 5480 is used for the airborne and altimetry data. Two different topography models are used above degree 2160 because they are calculated using the same original data and contain the same signal (Hirt et al. 2014; Rexer et al. 2016), but the ERTM2160 is only available as a grid on the Earth surface for land areas; see Liu et al. (2020b) for more details.

Figure 5 visualizes the remove step, i.e., δg (first row), $\delta g - \delta g_{\mathrm{GGM}}$ (second row), and $\delta g - \delta g_{\mathrm{GGM}} - \delta g_{\mathrm{topo}}$ (third row), for the three types of gravity data, respectively. The corresponding statistics are listed in Table 1. The gravity observations show large variations in this study area, which are caused by the diverse topography in Colombia. Removing the GGM smooths the gravity observations by 67%, 81%, and 79% in terms of STD for the terrestrial, airborne, and altimetry data, respectively. After removing the GGM contribution, the gravity field (second row in Fig. 5) is clearly dominated by topographic effects, i.e., a correlation to the terrain map (Fig. 1a) can be seen. After removing the topography effects, the gravity observations are further smoothed especially in mountainous areas. An interesting finding is that the topographic effects are also strong in the coastal region of the offshore area, due to its rugged bathymetry where the depth changes from 0 to -5000 m in a short distance. The smoothing effect by topography model is the largest with 44% for the terrestrial data and the smallest for the airborne data with 23%, which is reasonable as the terrestrial observations are given on the Earth's surface, but the airborne data are observed at the flight height and thus, less sensitive to the high-frequency gravity signal.

Model configuration

There are four factors of SRBFs that influence the modeling result, and thus, need to be specified, namely (1) the bandwidth, i.e., the maximum degree of expansion, (2) the locations of the SRBFs, (3) the type of the SRBFs, and (4) the extensions of the data zone for reducing the edge effects. These four factors are chosen following Liu et al. (2020b). To be more specific, the maximum degree of expansion n_{max} depends on the average spatial resolution of all available observations, and for this study area $n_{max} = 2190$ is chosen. The Reuter grid (Reuter 1982) is used, which generates a homogenous coverage of grid points on the sphere. The non-smoothing Shannon function is used in the analysis step for estimating the unknown coefficients to avoid the loss of spectral information. The Cubic Polynomial (CuP) function, which has smoothing features, is applied in the synthesis step for calculating the output gravity functionals to reduce erroneous systematic effects (Lieb et al. 2016). To minimize edge effects, the computation area $\partial\Omega_C$, where the SRBFs are located, should be larger than the observation area $\partial\Omega_O$, where the observations are given (see Fig. 1b). And $\partial\Omega_O$ should be larger than the investigation area $\partial\Omega_I$, where the final quasi-geoid model is computed, i.e., $\partial\Omega_I \subset \partial\Omega_O \subset \partial\Omega_C$. The margin size η between the three areas is determined by the maximum degree of expansion and the maximum latitude value of the investigation area (Liu et al. 2020b), and it is chosen as $\eta = 0.2°$.

The terrestrial, airborne, and altimetry data are combined using the Gauss–Markov model, i.e., Eq. (5), and the unknown coefficients are then estimated by Eq. (7) with the relative weights determined by VCE. The expectation vector $\boldsymbol{\mu}_d$ is set to the zero vector since the background model, i.e., the models δg_{GGM} and δg_{topo} removed within the RCR procedure serve as the prior information. The weight matrix $\boldsymbol{P}_d$ of the prior information is set to the identity matrix, i.e., $\boldsymbol{P}_d = \boldsymbol{I}$, assuming that the unknown coefficients are not correlated and have the same accuracy (Lieb et al. 2016). Among regional gravity field modeling publications (e.g., Wu et al. 2017; Slobbe et al. 2019), the weight matrix $\boldsymbol{P}_p$ of different observation

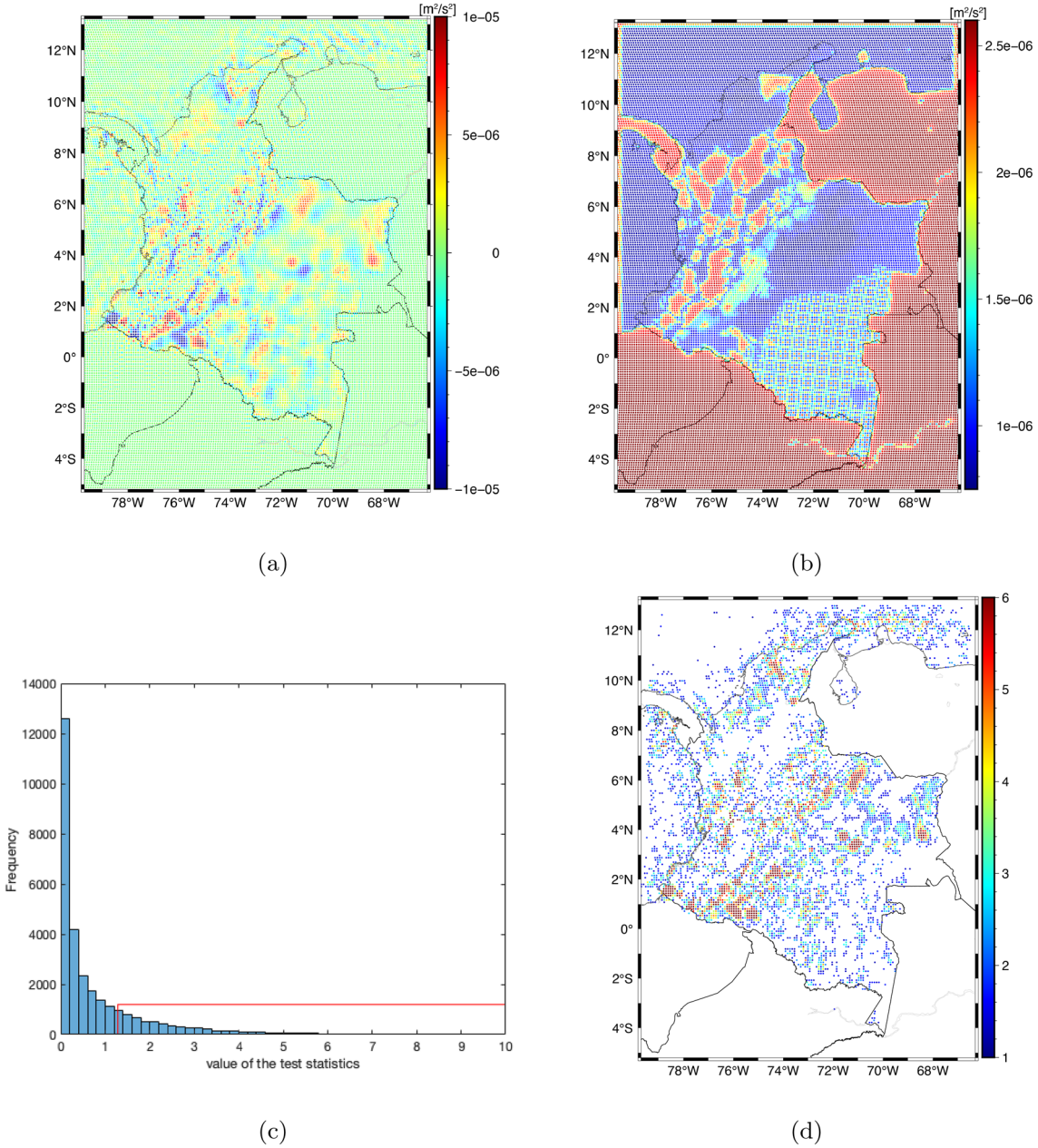


Fig. 6 **a** The estimated coefficients, **b** their standard deviations, **c** the histogram of the test statistic $|\widehat{d}|/\widehat{\sigma}_{\widehat{d}}$, and **d** the corresponding test statistic value of the significant coefficients

types is also commonly set to the identity matrix, i.e., $\boldsymbol{P}_p = \boldsymbol{I}$, by assuming that the same type of measurements has the same accuracy and is uncorrelated. The reason is that the exact data accuracy or quality of real gravity measurements is usually unknown, and consequently, it is difficult to set up a realistic full error variance–covariance matrix. In this study, the weight matrices for the terrestrial and altimetry data are both set to the identity matrix. For the airborne measurements, both the crossover analysis results and the comparison with the SATOP model suggest that different surveys have significantly different accuracy, and thus, it is not optimal to use the identity matrix as weight matrix. In this case, we take advantage of the conducted bias estimation (See Sect. 'Bias estimation using SRBFs') to set up the weight matrix. Namely, in the airborne bias estimation model Eq. (10), variance factors σ_q^2 $(q = 1, 2, \ldots, Q)$ are estimated for the airborne surveys using VCE, which

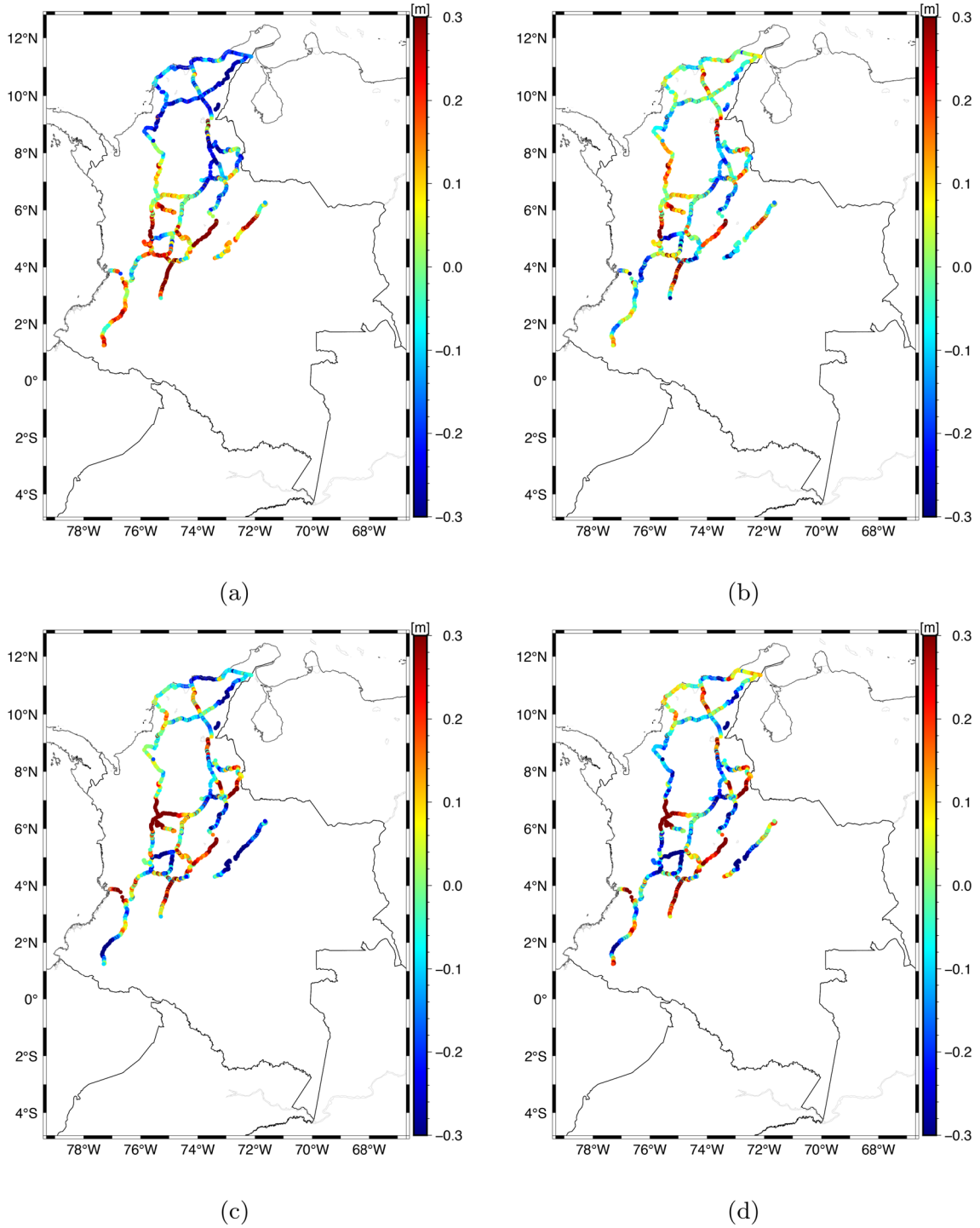


Fig. 7 Differences between QGeoidCOL2023 and the GPS/leveling data **a** before and **b** after applying the 7-parameter correction model, and the differences between QGEOID2021 and the GPS/leveling data **c** before and **d** after applying the 7-parameter correction model. Note that for the plots before applying the correction surface, the mean values have been removed

indicate the relative weights of the $Q = 17$ airborne surveys. These variance factors are then used for setting up the weight matrix of the airborne observations, i.e., $\boldsymbol{P}_{\text{airborne}} = \text{diag}(\frac{1}{\sigma_1^2}\boldsymbol{I}_1, \frac{1}{\sigma_2^2}\boldsymbol{I}_2, \ldots, \frac{1}{\sigma_Q^2}\boldsymbol{I}_Q)$.

Table 2 Comparison between the regional gravimetric quasi-geoid model and the GPS/leveling data (unit [cm])

	QGeoidCOL2023				QGEOID2021			
	Min	**Max**	**Mean**	**STD**	**Min**	**Max**	**Mean**	**STD**
Before correction	− 46.14	87.74	19.21	21.27	− 57.92	84.37	13.89	27.22
planar correction	− 76.65	58.70	0.00	16.54	− 80.84	73.50	0.00	26.63
polynomial correction	− 72.44	60.94	0.00	16.08	− 83.48	72.90	0.00	24.57
4-parameter correction	− 76.33	58.40	0.00	16.55	− 81.78	70.55	0.00	24.98
5-parameter correction	− 73.25	58.09	0.00	16.35	− 82.12	71.24	0.00	24.93
7-parameter correction	− 76.06	65.21	0.00	15.76	− 85.37	72.93	0.00	24.51

We follow the standards and conventions for the realization of the IHRS recommended by Sánchez et al. (2021) throughout the computation. The computation is conducted in the zero-tide system, with GRS80 being the reference ellipsoid. The height anomaly ζ is determined at the Earth's surface, and the zero-degree correction term

$$\zeta_0 = \frac{(GM_{\mathrm{GGM}} - GM_{\mathrm{GRS80}})}{r_P \cdot \gamma_Q} - \frac{W_0 - U_0}{\gamma_Q} \tag{12}$$

(Heiskanen and Moritz 1967) is added, where $GM_{\mathrm{GGM}} = 3.986004415 \cdot 10^{14}\ \mathrm{m}^3/\mathrm{s}^2$ and $GM_{\mathrm{GRS80}} = 3.986005 \cdot 10^{14}\ \mathrm{m}^3/\mathrm{s}^2$ are the GM constant value of the GGM and the reference ellipsoid GRS80, respectively. $W_0 = 62636853.4\ \mathrm{m}^2/\mathrm{s}^2$ is the conventional reference potential value adopted for the IHRS (IAG Resolution 1, Drewes et al. (2016)), and $U_0 = 62636860.850\ \mathrm{m}^2/\mathrm{s}^2$ is the potential value on the reference ellipsoid, i.e., GRS80. r_P is the geocentric radial distance of the points P at the Earth surface, where the quasi-geoid model is calculated, and γ_Q is the normal gravity at the corresponding point mapped onto the telluroid.

Results, validation, and discussion

Estimated coefficients

The estimated coefficients $\widehat{d}$ and their standard deviations $\widehat{\sigma}_d$ from Eqs. (7a) and (7b) are plotted in Fig. 6. The coefficients reflect the energy of the recovered gravity field at their locations due to the spatial localizing features of the SRBFs (Lieb 2017). Thus, the estimated coefficients with large absolute values, i.e., those not close to zero, indicate that additional gravity signals are captured with respect to the background model. In Fig. 6a, larger absolute values are observed in areas where the gravity data are located, i.e., a clear correlation between the estimated coefficients and the data distribution (Fig. 1b) can be seen, which demonstrates that additional gravity signals have been obtained from the gravity measurements. Furthermore, the largest absolute values are mainly located in areas with high elevation (see Fig. 1a), which shows the notable contribution of the terrestrial and airborne observations in rugged terrain, although topography models have already been introduced in the background model. In the offshore region, larger absolute values in the estimated coefficients are seen in coastal areas, representing the extra gravity information introduced by DTU21GRA in comparison to the older version DTU13GRA, which is used for developing XGM2019. This coincides with the conclusion of Wu et al. (2022), who state that the inclusion of the Sentinal-3A/B SAR altimetry data in DTU21GRA improves the marine quasi-geoid accuracy compared to the earlier releases especially over coastal regions. The estimated standard deviations (Fig. 6b) are about 1 order of magnitude smaller than the estimated coefficients, which indicates a well-balanced combination of the datasets (Lieb et al. 2016). They are larger in areas without gravity observations, which is also reasonable.

The majority of the coefficients in Fig. 6a, i.e., mainly over marine areas and the surrounding countries, are close to zero, which means that they do not contain extra gravity information w.r.t. the background model. A hypothesis t-test is then conducted to check whether a coefficient is significant, i.e., significantly different from zero, using the test statistic $|\widehat{d}|/\widehat{\sigma}_d$ (Bentel 2013). An individual coefficient $\widehat{d}$ is regarded to be significant if its corresponding test statistic $|\widehat{d}|/\widehat{\sigma}_d$ is larger than a critical value $t_{nu,p}$, see Koch (1999) and Liu et al. (2020b) for more details. Figure 6c shows the histogram of the test statistics, with the red rectangular marking those with value larger than $t_{nu,p}$, and Fig. 6d visualizes the corresponding coefficients which are significant. As we can see, only around 21% of the coefficients are regarded as significantly different from zero according to the t-test. The non-significant coefficients do not have influences on the modeling result and could be removed. This procedure would considerably save the computation effort, and it is especially beneficial when calculating high-resolution grid models, as the large design matrix size would be significantly reduced.

Validation with GPS/leveling data

The calculated quasi-geoid model, denoted as QGeoidCOL2023, is validated using the 3025 GPS/leveling data points (Fig. 1b) in both absolute and relative sense. In absolute evaluation, the calculated height anomaly values are directly compared with those from the GPS/leveling data point-by-point. In an ideal case, the relation between quasi-geoid, ellipsoidal, and normal heights should fulfill $\zeta = h - H^{\star}$. However, in reality, this equality is not satisfied due to different reasons (Fotopoulos et al. 1999), including (1) datum inconsistencies, i.e., the deviation between gravimetric quasi-geoid and the reference surface of the leveling datum, (2) random noise, errors, or systematic distortions in the ellipsoidal and physical heights, (3) geodynamic effects, such as surface deformation due to earthquakes, especially when the GPS/leveling data are collected during a large time span. Such systematic effects can be seen in Fig. 7a, i.e., the differences between the calculated gravimetric height anomaly $\zeta^{\text{gravimetric}}$ and the GPS/leveling-derived height anomaly $\zeta^{\text{GPS/leveling}}$ show a systematic trend from the north (negative values) to the south (positive values). To model these deviations, a correction surface could be applied to the GPS/leveling-based height anomaly, i.e.,

$$\zeta_i^{\text{GPS/leveling}} - \zeta_i^{\text{gravimetric}} = \boldsymbol{a}_i^T \boldsymbol{t} + v_i \tag{13}$$

(Heiskanen and Moritz 1967), where $\boldsymbol{t}$ is the vector of n unknown parameters, $\boldsymbol{a}_i^T \boldsymbol{t}$ is expected to describe the systematic effects and datum inconsistencies, and v_i is the error term, which is used as a measure of the height anomaly accuracy. $\boldsymbol{a}_i$ is a $n \times 1$ vector depending on the type of the applied correction surface, and the commonly used ones include the planar correction surface, the polynomial correction model (second order), the 4-parameter, 5-parameter, and 7-parameter correction model, respectively. Correspondingly, the $\boldsymbol{a}_i^T$ vector can be set up as

$$\boldsymbol{a}_i^T = [1, \varphi_i, \lambda_i]; \tag{14a}$$

$$\boldsymbol{a}_i^T = [1, \varphi_i, \lambda_i, \varphi_i^2, \lambda_i^2, \varphi_i \lambda_i]; \tag{14b}$$

$$\boldsymbol{a}_i^T = [1, \cos\varphi_i \cos\lambda_i, \cos\varphi_i \sin\lambda_i, \sin\varphi_i]; \tag{14c}$$

$$\boldsymbol{a}_i^T = [1, \cos\varphi_i \cos\lambda_i, \cos\varphi_i \sin\lambda_i, \sin\varphi_i, \sin^2\varphi_i]; \tag{14d}$$

$$\boldsymbol{a}_i^T = \left[\cos\varphi_i \cos\lambda_i, \cos\varphi_i \sin\lambda_i, \sin\varphi_i, \frac{\sin\varphi_i \cos\varphi_i \sin\lambda_i}{W_i}, \frac{\sin\varphi_i \cos\varphi_i \cos\lambda_i}{W_i}, \frac{1 - f^2 \sin^2\varphi_i}{W_i}, \frac{\sin^2\varphi_i}{W_i}\right]; \tag{14e}$$

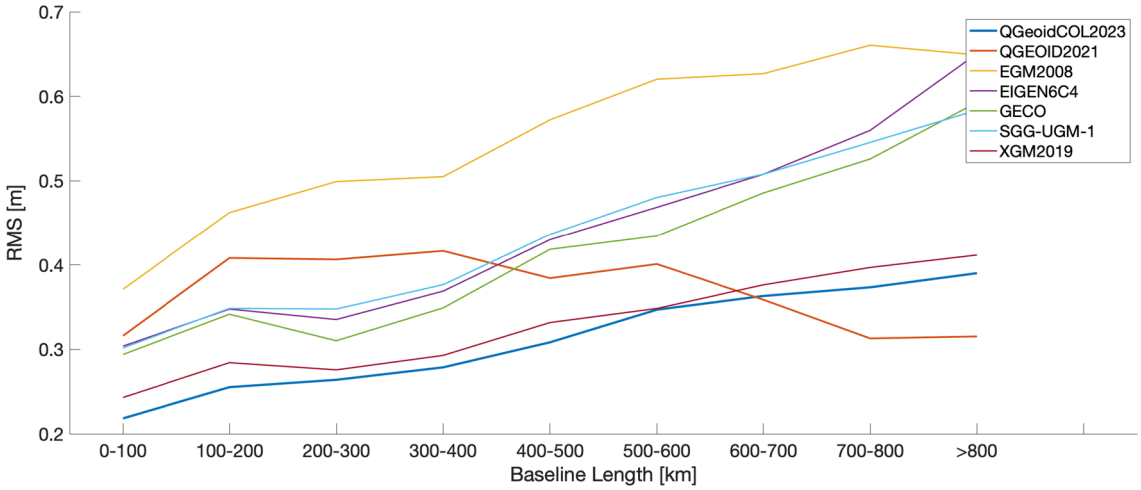
respectively (Fotopoulos et al. 1999; Kotsakis and Sideris 1999; Soycan and Soycan 2003), where φ_i and λ_i are the latitude and longitude of the i^{th} GPS/leveling point, $W_i = \sqrt{(1 - e^2 \sin^2\varphi_i)}$ with e^2 being the eccentricity, and f is the flattening of the reference ellipsoid, i.e., the GRS80. The parameter vector $\boldsymbol{t}$ as well as v_i can then be estimated through a least-squares adjustment. The estimation of a correction surface does not change the calculated gravimetric quasi-geoid model. It is used in this study to identify possible patterns in the discrepancies between the gravimetric height anomaly $\zeta_i^{\text{gravimetric}}$ and the GPS/leveling-derived height anomaly $\zeta_i^{\text{GPS/leveling}}$. The identification of such patterns can provide a more meaningful and reliable accuracy assessment for the gravimetric quasi-geoid model. The choice of the correction surface model depends on each study area, e.g., the observed discrepancies and the geological complexity. A model with less parameters is simple but might not be sufficient enough to fit the systematic effects, especially in mountainous areas. The comparison between the calculated quasi-geoid model QGeoidCOL2023 and the GPS/leveling data before and after applying different types of correction surfaces is shown in Table 2. The STD value of the differences decreases after applying the correction surface, and the 7-parameter model delivers the best result, i.e., the smallest STD value. This is plausible because the varying topography of Colombia demands a more detailed correction surface. Figure 7b presents the differences between the calculated height anomaly and the GPS/leveling data after applying the 7-parameter model. The systematic deviations in Fig. 7a are not visible anymore in Fig. 7b, and the STD value decreases from 21.27 cm to 15.76 cm.

It is worth mentioning that the differences between QGeoidCOL2023 and the GPS/leveling data contain not only the computation error of the gravimetric quasi-geoid height but also errors in the ellipsoidal and leveling height. Furthermore, the GPS/leveling data are mainly distributed in the mountainous area over the western part of Colombia, where no airborne gravity data are available and large data gaps exist in the terrestrial data. Thus, considering the limited distribution of the GPS/leveling dataset, an STD of 15.76 cm does not imply a unsatisfying accuracy of the calculated quasi-geoid model. As a reference, the South American quasi-geoid model QGEOID2021 is also validated using the same GPS/leveling data, and their differences are

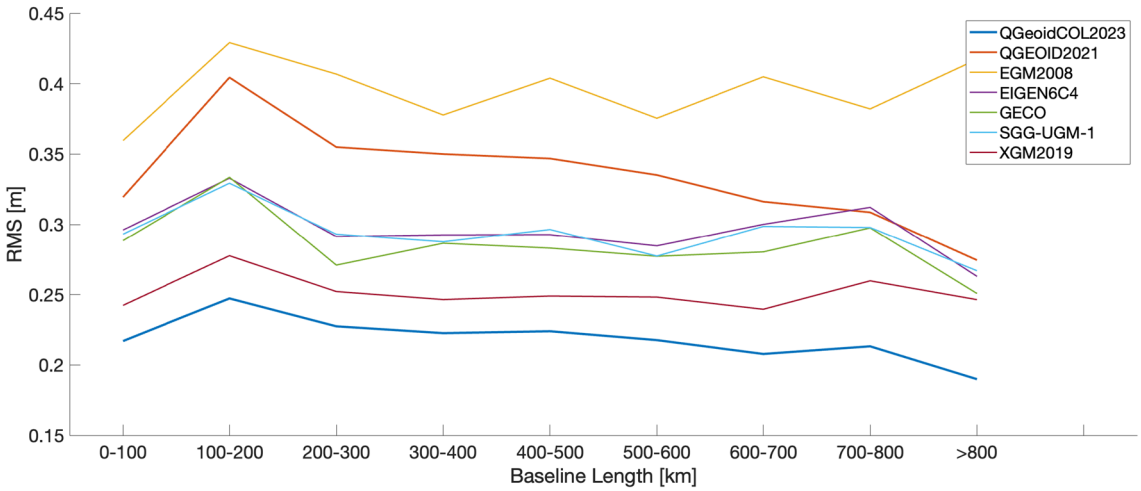
Table 3 Comparison between different high-resolution GGMs and the GPS/leveling data after applying the 7-parameter correction model (unit [cm])

	Min	Max	Mean	STD
EGM2008	− 80.73	83.85	0.00	28.09
EIGEN6C4	− 93.29	66.80	0.00	21.10
GECO	− 110.82	60.37	0.00	20.39
SGG-UGM-1	− 71.03	68.28	0.00	20.93
XGM2019	− 75.93	80.06	0.00	17.86

shown in Fig. 7c. Less systematic effects are observed compared to Fig. 7a, which reflects smaller datum inconsistency between QGEOID2021 and the GPS/leveling data. This is explainable as QGEOID2021 and the GPS/leveling data both refer to the local height reference system but QGeoidCOL2023 refers to the IHRS. Nevertheless, QGEOID2021 delivers an STD of 27.22 cm (Table 2), which is 28% larger than the one given by QGeoidCOL2023. By applying a correction surface to the GPS/leveling data, the STD values delivered by QGEOID2021 also decrease (see Table 2), and again, the smallest STD is obtained when the



(a)



(b)

Fig. 8 Relative validation in terms of RMS error between the gravimetric quasi-geoid models, i.e., QGeoidCOL2023, QGEOID2021, and different GGMs, and the GPS/leveling data depending on the baseline length **a** before and **b** after applying the 7-parameter correction model

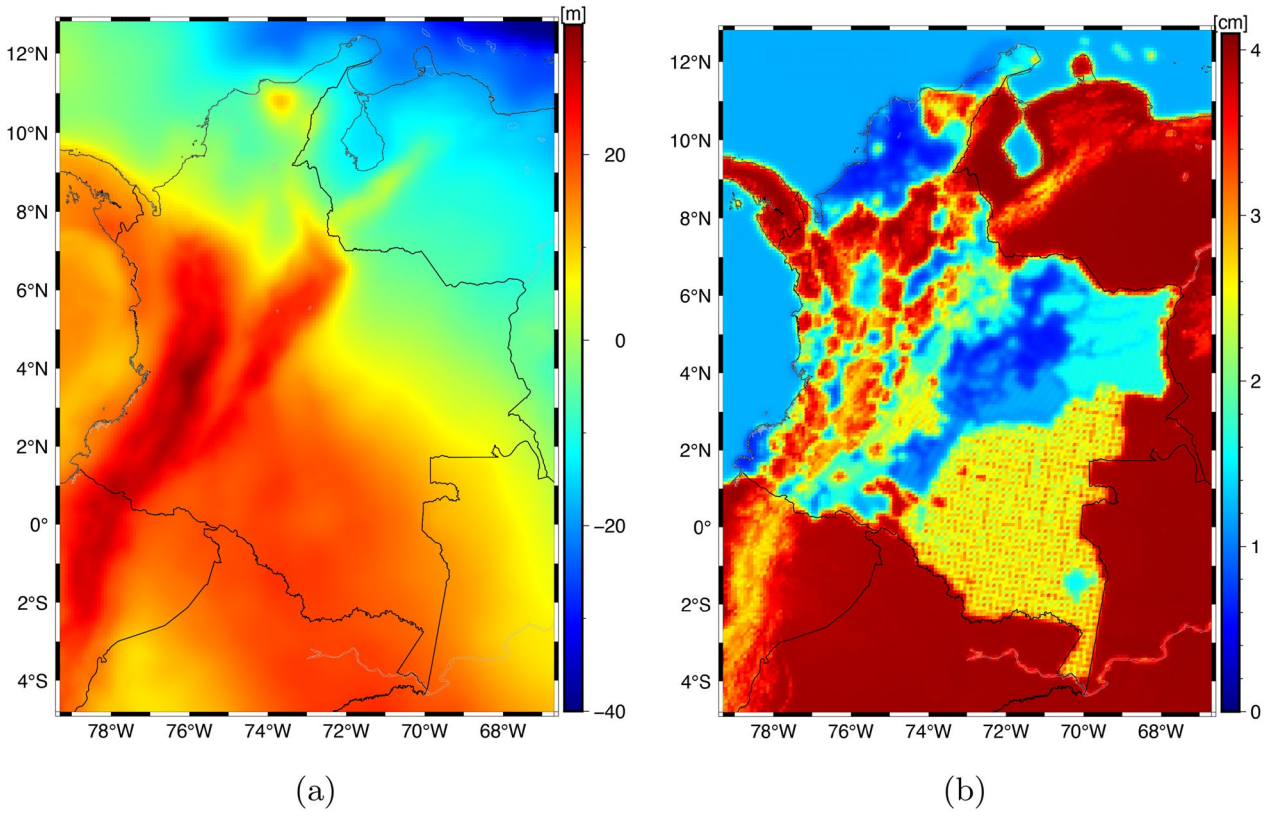


Fig. 9 **a** Quasi-geoid model QGeoidCOL2023; and **b** its standard deviation map

7-parameter model is applied. The corresponding differences between QGEOID2021 and the GPS/leveling data after applying the 7-parameter correction model are shown in Fig. 7d, giving an STD of 24.51 cm, which is 55% larger than that obtained by QGeoidCOL2023 (Fig. 7b). The 7-parameter model best presents the systematic discrepancies between the gravimetric model and the GPS/leveling data. These discrepancies are expected being related to the challenging topography in the region and the fact that the terrestrial gravity data and leveling are deployed along roads, which usually run along the slopes of the mountains.

Besides regional gravity field refinement, another approach for the determination of the IHRF coordinates is to use high-resolution GGMs (Sánchez et al. 2021). To test the validity of this approach, the performance of different high-resolution GGMs, namely EGM2008 (Pavlis et al. 2012), EIGEN6C4 (Förste et al. 2014), GECO (Gilardoni et al. 2016), SGG-UGM-1 (Liang et al. 2018), and XGM2019 (Zingerle et al. 2020), is evaluated using the GPS/leveling data. Each GGM is truncated at degree 2159, and the topography model ERTM2160 is added to count for the very high-frequency parts from degree 2160 to around 80,000. Many recent publications have been dedicated to the evaluation and comparison of different GGMs in specific regions or countries, e.g., Yilmaz et al. (2017) for Turkey, Foroughi et al. (2017) for Iran, Pham et al. (2023) for Vietnam, among many others. To our knowledge, such evaluation has never been made for Colombia, and thus, is of high interest. Table 3 shows the comparison between different GGMs and the GPS/leveling data; note that the 7-parameter correction model has been applied to count for the systematic inconsistencies. EGM2008 delivers the largest STD value of 28.09 cm. EIGEN6C4, GECO, and SGG-UGM-1, which are all developed based on EGM2008, show similar performance and significant improvements (around 25%) compared to EGM2008. This can be explained by the inclusion of GOCE data in these three models and reveals the contribution of GOCE data to the gravity model. XGM2019 is the best performing GGM for Colombia, which could be due to the usage of an improved ground gravity anomaly dataset provided by the National Geospatial-intelligence Agency (NGA). As stated by Pail et al. (2018), this improved dataset results in significant upgrades especially in continental areas such as South America, Africa, parts of Asia,

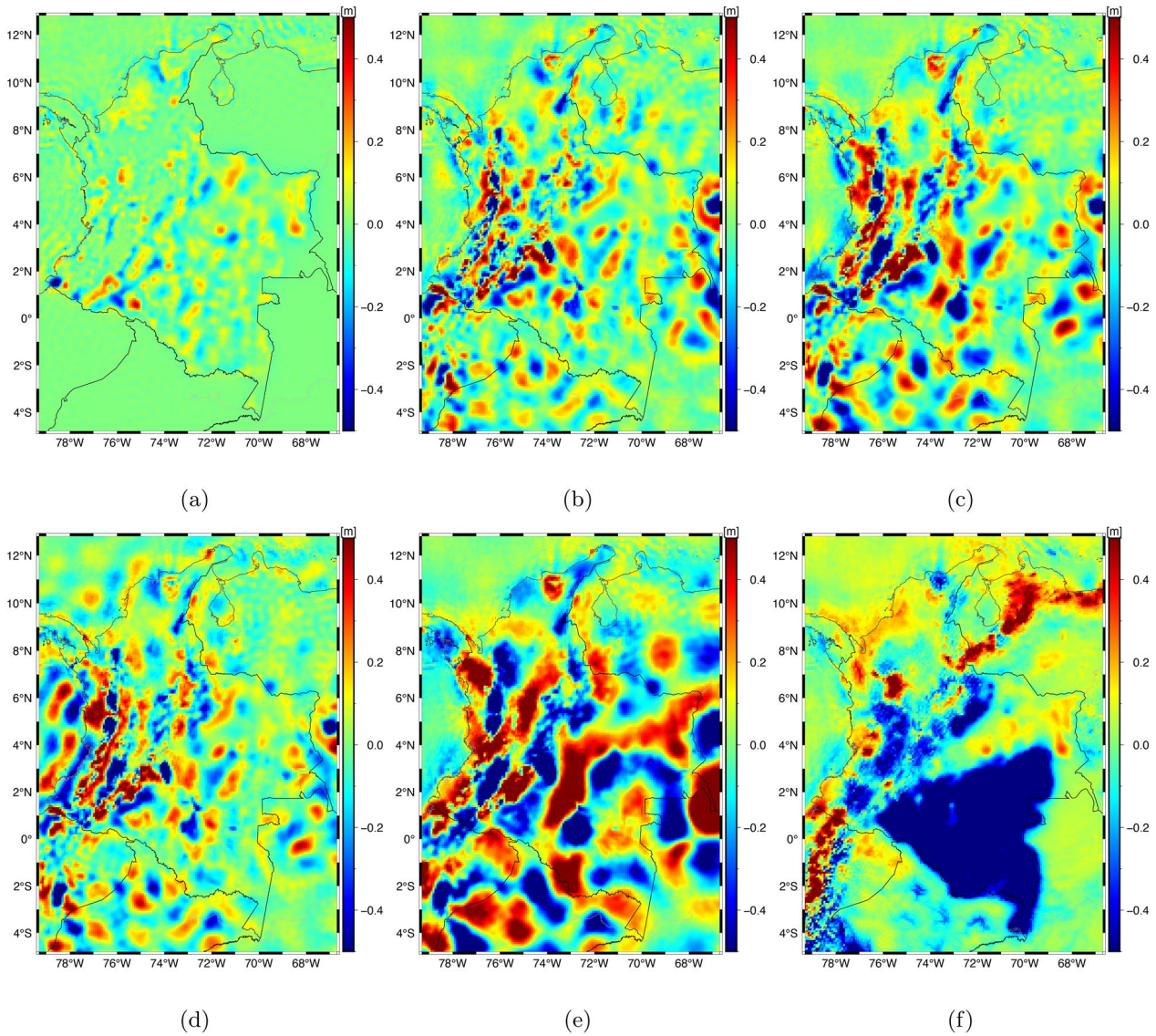


Fig. 10 Quasi-geoid differences between QGeoidCOL2023 and **a** XGM2019, **b** GECO, **c** SGG-UGM-1, **d** EIGEN6C4, **e** EGM2008, and **f** QGEOID2021

Table 4 Statistics of the differences between the quasi-geoid model QGeoidCOL2023 and the five different high-resolution GGMs as well as QGEOID2021 (unit [cm])

	Min	Max	Mean	STD
XGM2019	− 70.76	46.48	0.04	5.24
GECO	− 275.53	150.49	− 0.03	16.15
SGG-UGM-1	− 298.41	164.64	− 0.01	20.11
EIGEN6C4	− 273.38	145.25	0.00	18.41
EGM2008	− 356.00	210.66	0.65	35.32
QGEOID2021	− 130.94	154.57	− 8.80	30.28

and Antarctica. Nevertheless, our regional quasi-geoid model QGeoidCOL2023 outperforms all the evaluated high-resolution GGMs, delivering improvements from 12% (compared to XGM2019) to 44% (compared to EGM2008), w.r.t. the GPS/leveling data. It demonstrates the necessity and importance of regional gravity field refinement, and it is preferred for the determination of the IHRF coordinates, compared to using current high-resolution GGMs in combination with topography models directly.

A relative validation with GPS/leveling data is also conducted, both before and after applying the correction surface. The advantage of relative validations is that vertical datum-related errors can be canceled to a large

extent, especially over short baseline distances (Featherstone 2001). It is realized by computing the differential height anomaly value between each pair of the GPS/leveling benchmarks (see, e.g., Sánchez et al. (2021))

$$\Delta\zeta_{ij} = \left(\zeta_i^{\text{GPS/leveling}} - \zeta_j^{\text{GPS/leveling}}\right) - \left(\zeta_i^{\text{gravimetric}} - \zeta_j^{\text{gravimetric}}\right), \tag{15}$$

where i and j are the indexes of the pair of GPS/leveling data points. For l number of GPS/leveling points, a total number of $l(l-1)/2$ baselines or pairs can be formulated. In our case, with $l = 3025$, altogether 4,573,800 pairs of $\Delta\zeta_{ij}$ are calculated. These pairs are grouped based on the baseline length d_{ij}, which is calculated as the geodestic distance, i.e., the shortest distance between two points i and j on the reference ellipsoid (Featherstone 2001). The RMS error of $\Delta\zeta_{ij}$ within each group before and after applying the 7-parameter correction model is visualized in Fig. 8a and b, respectively.

Before applying the correction surface, the RMS error of the calculated gravimetric model QGeoidCOL2023 w.r.t. the GPS/leveling data increases gradually when the baseline length increases. This is expected and can be explained by the systematic errors between the quasi-geoid model and the GPS/leveling data, as visualized in Fig. 7a. Such pattern is also observed in all the evaluated high-resolution GGMs. Nevertheless, QGeoidCOL2023 provides smaller RMS values than all the GGMs at each baseline length group. However, QGEOID2021 does not present this behavior, i.e., at baseline length 0–400 km, QGEOID2021 shows the second largest RMS error among all the evaluated models, but for baseline length larger than 700 km, it delivers the smallest RMS value. One possible reason could be that the effects of datum inconsistencies are smaller for this model as shown in Fig. 7c, because it refers to the local vertical datum. After reducing the systematic effects by a correction surface, no clear correlation between the RMS value and the baseline length is observed. QGeoidCOL2023 delivers the smallest RMS error at all baseline length. Among the GGMs, XGM2019 gives the best result at each baseline length, followed by GECO, SGG-UGM-1, and EIGEN6C4, which show very similar performance. EGM2008 and QGEOID2021 deliver the largest and second-largest RMS errors, which are almost twice the value given by QGeoidCOL2023.

Quasi-geoid model QGeiodCOL2023

The calculated $5' \times 5'$ quasi-geoid grid QGeoidCOL2023 is visualized in Fig. 9a. The height anomaly values range from −39.95 m to 33.42 m, which illustrates the high variation of the gravity field in this study area. In Colombia, the north-east part has height anomaly around 0 m, and the south-west part has a much larger value of more than 20 m. In the surrounding offshore area, the north (Caribbean Sea) shows negative height anomaly values but the west (Pacific Ocean) has positive values. The corresponding standard deviation map of the calculated quasi-geoid is shown in Fig. 9b, which ranges from few millimeters to around 4 cm. The values are smaller in areas where denser gravity observations are available, which is reasonable.

QGeoidCOL2023 is compared with the five high-resolution GGMs considered in Sect. 'Validation with GPS/leveling data', namely XGM2019, GECO, SGG-UGM-1, EIGEN6C4, and EGM2008, ordered according to their performance in the validation with the GPS/leveling data (see Table 3). Again, the GGMs are truncated at degree 2159, and the topography model ERTM2160 is added to count for the very high frequency parts. The differences between QGeoidCOL2023 and the five GGMs are shown in Fig. 10a–e, respectively. The corresponding statistics of these differences are listed in Table 4. In Fig. 10a, almost no differences can be seen in the surrounding countries of Colombia, i.e., where no gravity observations are considered (see Fig. 1b). This is expected because in these regions QGeoidCOL2023 relies on the background model, namely XGM2019 for degree 2 to 719, dV_ELL_Earth2014 for degree 720 to 2159, and ERTM2160, while beyond degree 719 XGM2019 is developed from Earth2014 as well. In Colombia, the differences reveal the additional contribution of the terrestrial and airborne data. The differences are larger in the eastern part where the airborne data are located, and the possible reasons are either these airborne data are not included in the NGA gravity anomaly grid that is used for calculating XGM2019 or our data processing methods, e.g., bias estimation and outlier detection brought improvements to the airborne gravity data. In the western part of Colombia, where the terrestrial data are located, large differences between QGeoidCOL2023 and XGM2019 are seen in the mountainous areas. This is explainable because although the terrestrial data have already been included in the NGA gravity anomaly grid, it has a sparse spatial resolution of $15' \times 15'$ and is spectrally limited to degree 719 (Zingerle et al. 2020), while the topography model Earth2014 is not able to represent the true high-frequency gravity signals accurately. Thus, QGeoidCOL2023 also shows improvement especially in the mountainous areas by using the terrestrial data to a much higher spectral degree. In the offshore area, the differences are larger in the coastal regions, which is due to the advances of DTU21GRA in comparison to DTU13GRA, as already explained

in Sect. 'Estimated coefficients'. The differences between QGeoidCOL2023 and GECO, SGG-UGM-1, and EIGEN6C4, i.e., Fig. 10b–d show similar pattern. Compared to Fig. 10a, larger differences are observed in these three GGMs in Colombia, which could be due to the fact that XGM2019 incorporates an improved gravity anomaly grid, but an older gravity grid is used for the other three models, as already discussed in Sect. 'Validation with GPS/leveling data'. In the offshore area, EIGEN6C4 shows much larger differences w.r.t. QGeoidCOL2023, compared to all the other GGMs, especially in the Pacific Ocean. GECO and SGG-UGM-1 are developed by directly combining GECO and EGM2008, but EIGEN6C4 combines EGM2008 over the continents only and DTU10 over the oceans. Thus, these large differences in the Pacific Ocean could either come from DTU10 or from the data combination strategy used for developing EIGEN6C4. The latter is supported by the results presented in Gilardoni et al. (2016), which show that the EIGEN6C4 combination is not optimal in the Weddell Sea. The exact reason for such behaviour of EIGEN6C4 in the ocean areas still needs further investigation. Nevertheless, it is an interesting result and indicates that special attention should be paid if one intends to use EIGEN6C4 in offshore study areas. Figure 10e presents the largest deviations among the five GGMs in comparison to QGeoidCOL2023. The differences show long-wavelength patterns, which can be attributed to the absence of GOCE data in EGM2008.

QGeoidCOL2023 is also compared with QGEOID2021, and their differences are visualized in Fig. 10f with the corresponding statistics shown in Table 4. Major differences are located in the Andes as well as in the Amazon rain forest (see Fig. 1a). As we do not have enough knowledge regarding how QGEOID2021 was computed and how the gravity data were processed, it is difficult to conclude the exact reasons for these differences. However, it is likely that they result from datum inconsistencies. In the Amazon region, the same set of airborne data (note that different data processing has been implemented in this study) is also used in QGEOID2021 (Matos et al. 2021), but the resulting quasi-geoid differences are around 40 cm. Such large but consistent differences are unlikely to be from computation error, and they are not visible in the comparison between QGeoidCOL2023 and different GGMs. It is possible that the bias in the airborne survey was not corrected in the computation of QGEOID2021 or it was reduced to a different reference surface. Large differences also show up in the Andes, both inside and outside Colombia, where QGeoidCOL2023 replies purely on XGM2019 and topography model. Thus, the possibility that these differences come from computation errors in QGeoidCOL2023 is ruled out.

Conclusion

The determination of geopotential-based height systems relies on regional gravity field refinement, but high-resolution regional gravity models are often missing in developing or newly industrializing areas due to challenges such as heterogeneous data distribution and quality. This study addresses this issue and computes the first high-resolution quasi-geoid model for Colombia since 2004. Robust data processing strategies and structures are proposed to mitigate systematic discrepancies, erroneous records, and outliers in the historical terrestrial gravity measurements spanning six decades. We undertake a thorough examination of 17 airborne gravity surveys, each measured at different times and post-processed using different procedures, before their inclusion into the quasi-geoid computation. A crossover analysis is conducted within each survey, and the results indicate heterogeneous data accuracy. An evaluation of the airborne data is carried out by comparing to the SATOP model, which reveals substantial and diverse biases exceeding 40 mGal in different surveys. To solve this problem, a bias estimation method is developed based on the SRBFs. Its validity is confirmed first by a simulation test and then by the evaluation w.r.t. SATOP after removing the estimated biases in the airborne data. Additionally, this bias estimation model provides relative weights between different airborne surveys, and offers insights for setting up the covariance matrix of the airborne data in the parameter estimation. Removing data bias is essential in quasi-geoid modeling to prevent any undesired errors, and it would be interesting for future work to explore the possibilities of implementing such a bias estimation based on other regional gravity field modeling methods, e.g., the LSC. To compensate for the absence of gravimetric observations in the offshore area, we take advantage of the altimetry-derived gravity anomalies DTU21GRA. Modeling results demonstrate improvements compared to using GGMs directly for the offshore region. The methods and procedures developed in this computation can be easily integrated into other study areas with undefined or challenging data quality, facilitating the realization of the IHRS or any geopotential-based height system.

The calculated quasi-geoid model QGeoidCOL2023 is thoroughly validated using the GPS/leveling data in both absolute and relative sense. To account for the systematic effects caused by datum inconsistencies, a

correction surface is applied to the GPS/leveling data. Among the different types of correction models examined, the 7-parameter model delivers the smallest STD value between the gravimetric quasi-geoid model and the GPS/leveling data, which is explainable due to the complex topography of the study area. The STD obtained by QGeoidCOL2023 is 15.76 cm in comparison to the GPS/leveling data. It is worth mentioning that this STD value also contains uncertainties and errors of the GPS/leveling dataset, which is not expected to be of high accuracy and mainly located in the mountainous area where large data gaps exist. Although 15.76 cm is rather large compared to those reported in other well-surveyed regions, it still represents a satisfying quasi-geoid accuracy for this study area. As a reference, the latest South American quasi-geoid model QGEOID2021 gives an STD of 24.51 cm w.r.t. the GPS/leveling data, which is 55% larger than that obtained by QGeoidCOL2023. The performance of five recent high-resolution GGMs is also evaluated using the GPS/leveling data. They deliver STD values ranging from 17.86 cm (XGM2019) to 28.09 cm (EGM2008), with a mean value of 21.67 cm, which is 37% higher than that of QGeoidCOL2023. In the relative GPS/leveling validation, QGeoidCOL2023 again outperforms QGEOID2021 and all the five GGMs, i.e., it delivers the smallest RMS errors in every baseline length group.

Comparisons are also made between QGeoidCOL2023 and QGEOID2021 as well as the five GGMs in a $5' \times 5'$ grid over the whole study area. This comparison reveals the contributions and improvements brought by the airborne gravity measurements and the DTU21GRA, which are not reflected in the validation with GPS/leveling data due to their limited distribution. EGM2008 is the most unreliable GGM for Colombia, which shows large long-wavelength errors due to the absence of GOCE data. The quasi-geoid models calculated from GECO, SGG-UGM-1, and EIGEN6C4 are similar, but an interesting finding is that EIGEN6C4 is not able to represent the gravity field in the offshore area properly, especially in the Pacific Ocean. QGEOID2021 shows a large bias of around 40 cm over the Amazon region compared to QGeoidCOL2023 and all the GGMs, which could be introduced by the uncorrected bias in the airborne data during its computation. In summary, both the GPS/leveling validation and the grid comparison results indicate the validity and benefits of our developed computation methods and data processing procedures. They also demonstrate why regional gravity field refinement is preferred for the determination of IHRS coordinates compared to using high-resolution GGMs in combination with topography models directly. Accordingly, QGeoidCOL2023 represents so far the most reliable quasi-geoid model in Colombia for the realization of the IHRS.

Acknowledgements
The authors would like to thank the German Research Foundation (DFG) for funding the project 'Enhanced geopotential field modelling as basis for the establishment of precise height systems (Geo-H)' (Grant number: SA3330/4-1). We thank Dr. Ole B. Andersen for kindly providing the DTU21GRA. We acknowledge the developers of the Generic Mapping Tool (GMT) mainly used for generating the figures in this work. The authors also thank the reviewers and editors for their valuable comments and suggestions.

Author contributions
QL Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Visualization, Writing - original draft, Writing - review and editing. MS: Methodology, Validation, Investigation, Writing - review and editing, Project Administration, Funding acquisition. LS: Conceptualization, Validation, Investigation, Data curation, Writing - review and editing, Project Administration, Funding acquisition. LM: Data curation, Writing - review and editing. DC: Data curation, Writing - review and editing.

Funding
Open Access funding enabled and organized by Projekt DEAL. This study was supported by the German Research Foundation (DFG). Open Access funding enabled and organized by Projekt DEAL.

Availability of data and materials
The terrestrial and airborne gravity observations, as well as the GPS/leveling data are provided by the National Mapping Agency IGAC, the Colombian Geological Survey, and the oil company Ecopetrol. The DTU21GRA is provided by Dr. Ole B. Andersen from Technical University of Denmark. The global gravity models used in this study can be publicly accessed through http://icgem.gfz-potsdam.de. The quasi-geoid model QGeoidCOL2023 obtained within this study will be available at the repository of the International Service for Geoid.

Declarations

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Deutsches Geodätisches Forschungsinstitut, Technical University of Munich (DGFI-TUM), Arcisstrasse 21, 80333 Munich, Germany. [2]Instituto Geográfico Agustín Codazzi, Cra. 30 No 48-51, Bogotá, Colombia.

Received: 2 November 2023 Accepted: 11 February 2024
Published online: 26 February 2024

References

Altamimi Z, Rebischung P, Collilieux X, Métivier L, Chanard K (2023) ITRF2020: an augmented reference frame refining the modeling of nonlinear station motions. J Geodesy 97(5):47. https://doi.org/10.1007/s00190-023-01738-w

Andersen O, Knudsen P (2020) The DTU17 global marine gravity field: First validation results. In: Fiducial Reference Measurements for Altimetry: Proceedings of the International Review Workshop on Satellite Altimetry Cal/Val Activities and Applications, Springer, pp 83–87, https://doi.org/10.1007/1345_2019_65

Bentel K (2013) Regional gravity modeling in spherical radial basis functions – on the role of the basis function and the combination of different observation types. PhD thesis, Norwegian University of Life Sciences

Bucha B, Janák J, Papčo J, Bezděk A (2016) High-resolution regional gravity field modelling in a mountainous area from terrestrial gravity data. Geophys J Int 207(2):949–966. https://doi.org/10.1093/gji/ggw311

Childers VA, Bell RE, Brozena JM (1999) Airborne gravimetry: an investigation of filtering. Geophysics 64:61–69. https://doi.org/10.1190/1.1444530

Drewes H, Kuglitsch F, Adám J, Rózsa S (2016) The geodesists' handbook 2016. J Geodesy 90:907–1205. https://doi.org/10.1007/s00190-016-0948-z

Eicker A, Schall J, Kusche J (2014) Regional gravity modelling from spaceborne data: case studies with GOCE. Geophys J Int 196(3):1431–1440. https://doi.org/10.1093/gji/ggt485

Featherstone W (2001) Absolute and relative testing of gravimetric geoid models using global positioning system and orthometric height data. Comput Geosci 27(7):807–814. https://doi.org/10.1016/S0098-3004(00)00169-2

Foroughi I, Afrasteh Y, Ramouz S, Safari A (2017) Local evaluation of earth gravitational models, case study: Iran. Geodesy Cartography 43(1):1–13. https://doi.org/10.3846/20296991.2017.1299839

Forsberg R, Olesen AV (2010) Airborne gravity field determination. In: Sciences of geodesy-I, Springer, pp 83–104, https://doi.org/10.1007/978-3-642-11741-1_3

Förste C, Bruinsma S, Abrikosov O, Lemoine JM, Marty JC, Flechtner F, Balmino G, Barthelmes F, Biancale R (2014) EIGEN-6C4-The latest combined global gravity field model including GOCE data up to degree and order 2190 of GFZ Potsdam and GRGS Toulouse. https://doi.org/10.5880/icgem.2015.1

Fotopoulos G, Kotsakis C, Sideris MG (1999) Evaluation of Geoid Models and Their Use in Combined GPS Levelling, Geoid Height Network Adjustments. Tech. rep., Universität Stuttgart, https://www.ifp.uni-stuttgart.de/dokumente/Schriftenreihe/report1999_4.pdf, Accessed 02 Nov 2023

Freeden W, Gervens T, Schreiner M (1998) Constructive approximation on the sphere with applications to geomathematics. Oxford University Press on Demand

Gilardoni M, Reguzzoni M, Sampietro D (2016) GECO: a global gravity model by locally combining GOCE data and EGM2008. Studia Geophysica et Geodaetica 60:228–247. https://doi.org/10.1007/s11200-015-1114-4

Gruber T, Willberg M (2019) Signal and error assessment of GOCE-based high resolution gravity field models. J Geodetic Sci 9(1):71–86. https://doi.org/10.1515/jogs-2019-0008

Heiskanen WA, Moritz H (1967) Physical geodesy. San Francisco W. H, Freeman and Company, USA

Hirt C, Featherstone W, Marti U (2010) Combining EGM2008 and SRTM/DTM2006.0 residual terrain model data to improve quasigeoid computations in mountainous areas devoid of gravity data. Journal of Geodesy 84(9):557–567, https://doi.org/10.1007/s00190-010-0395-1

Hirt C, Kuhn M, Claessens S, Pail R, Seitz K, Gruber T (2014) Study of the Earth's short-scale gravity field using the ERTM2160 gravity model. Comput Geosci 73:71–80. https://doi.org/10.1016/j.cageo.2014.09.001

IGAC (Instituto Geográfico Agustín Codazzi) (2004) Adopción del Marco Geocéntrico Nacional de Referencia MAGNA-SIRGAS como datum oficial de Colombia. https://www.igac.gov.co/sites/igac.gov.co-/files/adopcion.pdf, Accessed 02 Nov 2023

Ihde J, Sánchez L, Barzaghi R, Drewes H, Foerste 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

Klees R, Tenzer R, Prutkin I, Wittwer T (2008) A data-driven approach to local gravity field modelling using spherical radial basis functions. J Geodesy 82(8):457–471. https://doi.org/10.1007/s00190-007-0196-3

Koch KR (1999) Parameter estimation and hypothesis testing in linear models. Springer Science & Business Media

Koch KR, Kusche J (2002) Regularization of geopotential determination from satellite data by variance components. J Geodesy 76(5):259–268. https://doi.org/10.1007/s00190-002-0245-x

Koop R (1993) Global gravity field modelling using satellite gravity gradiometry. Nederlandse Commissie voor Geodesie, Delft

Kotsakis C, Sideris MG (1999) On the adjustment of combined GPS/levelling/geoid networks. J Geodesy 73:412–421. https://doi.org/10.1007/s001900050261

Kvas A, Brockmann JM, Krauss S, Schubert T, Gruber T, Meyer U, Mayer-Gürr T, Schuh WD, Jäggi A, Pail R (2021) GOCO06s-a satellite-only global gravity field model. Earth Syst Sci Data 13(1):99–118. https://doi.org/10.5194/essd-13-99-2021

Li X (2018) Using radial basis functions in airborne gravimetry for local geoid improvement. J Geodesy 92(5):471–485. https://doi.org/10.1007/s00190-017-1074-2

Li X (2021) Leveling airborne and surface gravity surveys. Appl Geomatics 13(4):945–951. https://doi.org/10.1007/s12518-021-00402-2

Liang W, Xu X, Li J, Zhu G (2018) The determination of an ultra-high gravity field model SGG-UGM-1 by combining EGM2008 gravity anomaly and GOCE observation data. Acta Geodaetica et Cartographica Sinica 47(4):425, https://doi.org/10.11947/j.AGCS.2018.20170269

Lieb V (2017) Enhanced regional gravity field modeling from the combination of real data via MRR. PhD thesis, Technische Universität München

Lieb V, Schmidt M, Dettmering D, Börger K (2016) Combination of various observation techniques for regional modeling of the gravity field. J Geophys Res Solid Earth 121(5):3825–3845. https://doi.org/10.1002/2015JB012586

Liu Q (2023) Regional gravity field refinement for geoid height modeling based on the combination of data from various observation techniques. DGK, C 896, Dissertation, Bayerische Akademie der Wissenschaften, München

Liu Q, Schmidt M, Pail R, Willberg M (2020a) Determination of the regularization parameter to combine heterogeneous observations in regional gravity field modeling. Remote Sensing 12(10):1617. https://doi.org/10.3390/rs12101617

Liu Q, Schmidt M, Sánchez L, Willberg M (2020b) Regional gravity field refinement for (quasi-) geoid determination based on spherical radial basis functions in Colorado. J Geodesy 94(99):1–19. https://doi.org/10.1007/s00190-020-01431-2

Liu Q, Schmidt M, Sánchez L (2022) Combination of different observation types through a multi-resolution representation of the regional gravity field using the pyramid algorithm and parameter estimation. J Geodesy 96(80):1–20. https://doi.org/10.1007/s00190-022-01670-5

Matos A, Blitzkow D, Guimarñes G, Silva V (2021) The South American gravimetric quasi-geoid: QGEOID2021. V. 1.0. GFZ Data Services, https://doi.org/10.5880/isg.2021.005

Morelli C, Gantar C, Honkasalo T, McConnell R, Tanner J, Szabo B, Uotila U, Whalen C (1971) The International Gravity Standardization: Net 1971 (IGSN 71). Bureau central de l'Association internationale de Geodésie

Moritz H (2000) Geodetic reference system 1980. J Geodesy 74(1):128–133. https://doi.org/10.1007/s001900050278

Pail R, Fecher T, Barnes D, Factor J, Holmes S, Gruber T, Zingerle P (2018) Short note: the experimental geopotential model XGM2016. J Geodesy 92:443–451. https://doi.org/10.1007/s00190-017-1070-6

Pavlis NK, Holmes SA, Kenyon SC, Factor JK (2012) The development and evaluation of the Earth Gravitational Model 2008 (EGM2008). Journal of geophysical research: solid earth 117. https://doi.org/10.1029/2011JB008916

Pham HT, Claessens S, Kuhn M, Awange J (2023) Performance evaluation of high/ultra-high-degree global geopotential models over Vietnam using GNSS/leveling data. Geodesy and Geodynamics. https://doi.org/10.1016/j.geog.2023.03.002

Reuter R (1982) Über integralformeln der einheitssphäre und harmonische splinefunktionen. PhD thesis, RWTH Aachen University

Rexer M, Hirt C, Claessens S, Tenzer R (2016) Layer-based modelling of the earth's gravitational potential up to 10-km scale in spherical harmonics in spherical and ellipsoidal approximation. Surv Geophys 37(6):1035–1074. https://doi.org/10.1007/s10712-016-9382-2

Rummel R, Balmino G, Johannessen J, Visser P, Woodworth P (2002) Dedicated gravity field missions-principles and aims. J Geodyn 33(1–2):3–20. https://doi.org/10.1016/S0264-3707(01)00050-3

Rummel R, Gruber T, Ihde J, Liebsch G, Rülke A, Schäfer U, Sideris M, Rangelova E, Woodworth P, Hughes C (2014) STSE GOCE+, Height system unifcation with GOCE, Doc. No. GO HSU-PL-002, Issue 1, 24-02-2014

Sánchez L, Sideris MG (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

Sánchez L, Ågren J, Huang J, Wang Y, Mäkinen J, Pail R, Barzaghi R, Vergos GS, Ahlgren K, Liu Q (2021) Strategy for the realisation of the international height reference system (IHRS). J Geodesy 95(33):1–33. https://doi.org/10.1007/s00190-021-01481-0

Schmidt M, Han SC, Kusche J, Sanchez L, Shum C (2006) Regional high-resolution spatiotemporal gravity modeling from GRACE data using spherical wavelets. Geophysical Research Letters 33(8), https://doi.org/10.1029/2005GL025509

Schmidt M, Fengler M, Mayer-Gürr T, Eicker A, Kusche J, Sánchez L, Han SC (2007) Regional gravity modeling in terms of spherical base functions. J Geodesy 81(1):17–38. https://doi.org/10.1007/s00190-006-0101-5

Slobbe C, Klees R, Farahani HH, Huisman L, Alberts B, Voet P, Doncker FD (2019) The impact of noise in a GRACE/GOCE global gravity model on a local quasi-geoid. J Geophys Res Solid Earth 124(3):3219–3237. https://doi.org/10.1029/2018JB016470

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

Soycan M, Soycan A (2003) Surface modeling for GPS-leveling geoid determination. Newton Bull 1:41–52

Tapley BD, Bettadpur S, Watkins M, Reigber C (2004) The gravity recovery and climate experiment: Mission overview and early results. Geophysical Research Letters 31(9), https://doi.org/10.1029/2004GL019920

Varga M, Pitoňák M, Novák P, Bašić T (2021) Contribution of GRAV-D airborne gravity to improvement of regional gravimetric geoid modelling in Colorado, USA. J Geodesy 95(5):53. https://doi.org/10.1007/s00190-021-01494-9

Vergos G, Tziavos I, Andritsanos V (2005) Gravity data base generation and geoid model estimation using heterogeneous data. In: Gravity, Geoid and Space Missions: GGSM 2004 IAG International Symposium Porto, Portugal August 30–September 3, 2004, Springer, pp 155–160, https://doi.org/10.1007/3-540-26932-0_27

Wang Y, Sánchez L, Ågren J, Huang J, Forsberg R, Abd-Elmotaal H, Barzaghi R, Basĭć T, Carrion D, Claessens S, Erol B, Erol S, Filmer M, Grigoriadis V, Isik M, Jiang T, Koç Ö, Li X, Ahlgren K, Krcmaric J, Liu Q, Matsuo K, Natsiopoulos D, Novák P, Pail R, Pitoňák M, Schmidt M, Varga M, Vergos G, Véronneau M, Willberg M, Zingerle P, (2021) Colorado geoid computation experiment - Overview and Summary. Journal of Geodesy 95(127):1–21. https://doi.org/10.1007/s00190-021-01567-9

Willberg M, Zingerle P, Pail R (2020) Integration of airborne gravimetry data filtering into residual least-squares collocation: example from the 1 cm geoid experiment. J Geodesy 94(8):1–17. https://doi.org/10.1007/s00190-020-01396-2

Wittwer T (2009) Regional gravity field modelling with radial basis functions. PhD thesis, Netherlands Geodetic Commission

Wu Y, Zhou H, Zhong B, Luo Z (2017) Regional gravity field recovery using the GOCE gravity gradient tensor and heterogeneous gravimetry and altimetry data. J Geophys Res Solid Earth 122(8):6928–6952. https://doi.org/10.1002/2017JB014196

Wu Y, Wang J, Abulaitijiang A, He X, Luo Z, Shi H, Wang H, Ding Y (2022) Local enhancement of marine gravity field over the spratly islands by combining satellite SAR altimeter-derived gravity data. Remote Sensing 14(3):474. https://doi.org/10.3390/rs14030474

Yilmaz M, Turgut B, Güllü M, Yilmaz I (2017) The evaluation of high-degree geopotential models for regional geoid determination in Turkey. Afyon Kocatepe Üniversitesi Fen Ve Mühendislik Bilimleri Dergisi 17(1):147–153. https://doi.org/10.5578/fmbd.50706

Zingerle P (2022) Advanced methodologies for large-scale gravity field modelling. PhD thesis, Technische Universität München

Zingerle P, Pail R, Scheinert M, Schaller T (2019) Evaluation of terrestrial and airborne gravity data over Antarctica-a generic approach. J Geodetic Sci 9(1):29–40. https://doi.org/10.1515/jogs-2019-0004

Zingerle P, Pail R, Gruber T, Oikonomidou X (2020) The combined global gravity field model XGM2019e. J Geodesy 94(7):1–12. https://doi.org/10.1007/s00190-020-01398-0

Publisher's Note