Geophysical Journal International

Geophys. J. Int. (2013) **194,** 751–771
Advance Access publication 2013 May 11

doi: 10.1093/gji/ggt149

Validation of static gravity field models using GRACE K-band ranging and GOCE gradiometry data

H. Hashemi Farahani,[1] P. Ditmar,[1] R. Klees,[1] J. Teixeira da Encarnação,[1] X. Liu,[1,*] Q. Zhao[2] and J. Guo[2]

[1] *Delft University of Technology, Stevinweg* 1, 2628 *CN, Delft, The Netherlands. E-mail: h.hashemi@tudelft.nl*
[2] *GNSS Research Centre, Wuhan University,* 129 *Luoyu Rd., Wuhan* 430079, *China*

Accepted 2013 April 10. Received 2013 March 30; in original form 2012 March 13

SUMMARY

The ability of satellite gravimetry data to validate global static models of the Earth's gravity field is studied. Two types of data are considered: K-band ranging (KBR) data from the Gravity Recovery and Climate Experiment (GRACE) mission and Satellite Gravity Gradiometry (SGG) data from the GOCE (Gravity field and steady-state Ocean Circulation Explorer) mission. The validation is based on analysis of misfits obtained as the differences between the data observed and those computed with a force model that includes, in particular, a static gravity field model to be assessed. To facilitate the model assessment on a region-by-region basis, we convert KBR data into so-called range combinations, which are approximately equivalent to the intersatellite accelerations. We only use the accurately measured components of SGG data, that is, *xx*, *yy*, *zz* and *xz* components with *x*, *y* and *z* being along-track, cross-track and radial axes. We perform the validation in spectral and spatial domain. The latter requires elimination of low-frequency noise in the misfit data with a subsequent averaging over pre-defined blocks. Only 'independent' data are used, that is, those that have not been used in the production of the models under consideration.

The proposed methodology is applied to eight models: EGM2008 (truncated at degree 250), EIGEN-6C (truncated at degree 250), two GRACE-only models (ITG-Grace03 and ITG-Grace2010s) and four (satellite-only) combined GRACE/GOCE models (GOCO01S, EIGEN-6S, GOCO02S and DGM-1S). The latter is a novel model developed at Delft University of Technology in collaboration with GNSS Research Centre of Wuhan University. The GRACE KBR and GOCE SGG data demonstrate a pronounced sensitivity to inaccuracies of EGM2008 in 5–22 mHz (27–120 cycles-per-revolution, cpr) and 10–28 mHz (54–150 cpr) frequency ranges, respectively. The latter data also show a high sensitivity to inaccuracies of ITG-Grace2010s in 25–37 mHz (135–200 cpr) frequency range. From the validation in the spatial domain, it is confirmed that independent data of both types allow a difference in performance of the models to be observed, despite the fact that the duration of these data is much shorter than that of data used to produce those models. It is shown that EGM2008 performs weaker than the combined GRACE/GOCE models (up to the highest spectral sensitivity of the validation data). Considering the root mean square misfits related to the *zz* gravity gradient component, the differences in performance are: 76–83 per cent in the continental areas poorly covered by terrestrial gravimetry measurements (Himalayas, South America and Equatorial Africa); 4–16 per cent in the continental areas well covered by these measurements (Australia, North Eurasia and North America); and 11 per cent in the world's oceans (65°S–65°N). The identified differences in the regions of the latter two categories are related to the added value of the GOCE mission. It is shown that ITG-Grace03 and ITG-Grace2010s are of a much lower accuracy

*Now at: Fugro Intersite B.V., Dillenburgsingel 69, 2263 HW, Leidschendam, The Netherlands.

than EGM2008 in the gravimetrically well-surveyed continental areas: by 62–70 and 19–35 per cent and in the world's oceans: by 54 and 18 per cent, respectively. Nevertheless, the former models show a higher accuracy in the gravimetrically poorly surveyed continental areas: by 62–69 and 69–75 per cent, respectively. This difference is explained mostly by a loss of information content of ITG-Grace03 when it was combined with terrestrial gravimetry data to produce EGM2008. The KBR and SGG test data identify this loss of information content in 4–23 mHz (22–124 cpr) and 9–26 mHz (50–140 cpr) frequency ranges, respectively. It is shown that EIGEN-6C also suffers from a similar problem but in a much less pronounced manner. In South America, for instance, this model is found to perform somewhat poorer than its satellite-only counterpart, that is, EIGEN-6S, by about 12 per cent. The combined GRACE/GOCE models show in the poorly surveyed continental areas a higher accuracy than ITG-Grace2010s: by 23–36 per cent, which is attributed to the added value of the GOCE mission data. GOCO02S outperforms GOCO01S by not more than 2–5 per cent. DGM-1S and GOCO02S show an almost similar performance against SGG test data. However, the former model shows a slightly better agreement with KBR test data. Both models agree with test data better than EIGEN-6S.

Key words: Satellite gravity; Gravity anomalies and Earth structure; Geopotential theory.

1 INTRODUCTION

The accuracy of static, global models of the Earth's gravity field keeps increasing. To a large extent, this is due to the launch of new satellite gravity missions: the Gravity Recovery And Climate Experiment (GRACE) (Tapley *et al.* 2004) and the Gravity field and steady-state Ocean Circulation Explorer (GOCE) (Drinkwater *et al.* 2003; Floberghagen *et al.* 2011). The first of the two was launched in March 2002. It delivers data that have been used in the production of a number of global static gravity field models including: EIGEN-GL04 (Förste *et al.* 2008a), EIGEN-GL05 (Förste *et al.* 2008b), GGM02 (Tapley *et al.* 2005), GGM03 (Tapley *et al.* 2007), ITG-Grace03 (Mayer-Gürr 2006), ITG-Grace2010s (Mayer-Gürr *et al.* 2010a,b) and EGM2008 (Pavlis *et al.* 2008, 2012). The GOCE satellite is in orbit since 2009 March. The data from this mission have been used in the production of both GOCE-only models (Bruinsma *et al.* 2010; Pail *et al.* 2010a; Migliaccio *et al.* 2011) and a number of combined ones, which exploited GRACE data and, sometimes, terrestrial gravimetry and satellite altimetry measurements, for example, GOCO01S (Pail *et al.* 2010b), GOCO02S (Goiginger *et al.* 2011), EIGEN-6S (Förste *et al.* 2011), GOCO03S (Mayer-Gürr *et al.* 2012) and EIGEN-6C (Förste *et al.* 2011).

The GRACE mission consists of two identical satellites co-orbiting with an along-track separation of about 220 km at an altitude of approximately 500 km in a near polar orbit with an inclination of 89.5°. Their primary scientific payload is a K-Band Ranging (KBR) system, which measures biased values of the inter-satellite ranges with a precision of a few microns.

The GOCE satellite followed until 2012 August an orbit with an altitude of 254.9 km and an inclination of 96.7°. Its primary measurements are the second-order derivatives of the gravitational potential, which form a matrix known as the gravity gradient tensor. They are often referred to as Satellite Gravity Gradiometry (SGG) data.

The high accuracy of the new global static gravity field models makes their validation a challenging task. There are a number of ways that have been traditionally used for that purpose based on different sets of control data: (i) applying a model to compute a satellite's orbit parameters, which are then confronted with Satellite Laser Ranging (SLR) and/or Global Positioning System (GPS) measurements acquired on-board the satellite (e.g. Visser *et al.* 2009; Gruber *et al.* 2011); (ii) synthesis of gravity anomalies or deflections of the vertical at the Earth's surface with a subsequent comparison of them with terrestrial gravimetry or astro-geodetic measurements, respectively (e.g. Ihde *et al.* 2010; Hirt *et al.* 2011) and (iii) usage of a model to compute the geoid height differences between various locations with a subsequent comparison of the results with GPS and levelling measurements (e.g. Gruber 2009; Gruber *et al.* 2011). These validation techniques are not free of certain limitations. First, the accuracy of the new combined GRACE/GOCE models is so high that an assessment of them using control SLR- and/or GPS-based satellite orbit data mostly reveals errors in these data, so that different models may demonstrate a very similar performance. Secondly, the control data mentioned earlier, except for GPS-based satellite orbit data, are not currently globally or homogeneously available, which makes validation results obtained on their basis insufficiently representative. Thirdly, due to the spectral distribution of signal versus that of noise in the control data stated earlier, their frequency content is practically limited either to a very low-degree or to a very high-degree range in the spherical harmonic expansion of the gravitational potential. That is, satellite orbit data are only sensitive to low-degree spherical harmonic coefficients, whereas the terrestrial gravity anomalies, deflections of the vertical, and geoid height differences are primarily sensitive to high-degree coefficients.

The aim of this manuscript is to extend the set of tools for the validation of global static gravity field models. To that end, we analyse an alternative validation procedure in which the accuracy of models is assessed in terms of their ability to forecast data delivered by satellite gravity missions themselves. Two types of control data are considered: KBR data from the GRACE mission and SGG data from the GOCE mission. Each of them has its pros and cons. GRACE KBR measurements contain extremely accurate information about the gravity field. These measurements are particularly sensitive to signals at relatively low degrees, so that even temporal gravity field variations due to, for example, natural mass transport can be reliably sensed (Wahr *et al.* 1998). They contain also some information about gravity signal up to high degrees: 180 or even higher (Ditmar *et al.* 2012). In addition, these data cover practically the

entire surface of the Earth due to a nearly polar orbit of the GRACE satellites. Unfortunately, the sensitivity of GRACE KBR data is highly anisotropic. They are much more sensitive to the north–south variations of the gravity field than to the east–west ones. This can be easily understood from the fact that the GRACE configuration can be seen as a realization of a very large, one-component, along-track gradiometer (Keller and Sharifi 2005). Most of the time, the arm of this gradiometer is nearly parallel to the meridional direction and, therefore, almost 'blind' to the east–west gravity gradients. This means that a validation based on GRACE KBR data alone cannot provide comprehensive information about the quality of models. In contrast, GOCE SGG data contain information about spatial variations of the gravity field in all directions. Furthermore, the very low altitude of the GOCE satellite facilitates a high sensitivity of these measurements to small-scale spatial variations of the gravity field, namely signals up to degree 200–250. A drawback of these measurements is an increased noise level at low frequencies, especially at those associated with spherical harmonics of degree 27 and below (Rummel *et al.* 2011). Furthermore, the geographical distribution of GOCE measurements suffers from polar gaps of approximately 1500 km in diameter. Thus, we find it essential to consider control data of both types to benefit from the advantages of each of them in the validation procedure.

Generally, it is always advisable to use only independent data for the validation. In this study, we define independent data as those that have not been incorporated into models under consideration. In this way, we reduce the likelihood of misleadingly obtaining the most favourable results for the models in the computation of which GRACE KBR and/or GOCE SGG data have been over-weighted. Fortunately, acquiring independent data of these types is currently not a problem, as both the GRACE and GOCE missions are still operational.

In this manuscript, we apply the proposed validation methodology to eight global static gravity field models. Seven of them are known from the previous publications: EGM2008 as a state-of-the-art pre-GOCE model based on GRACE, terrestrial gravimetry and satellite altimetry data; EIGEN-6C, which uses GOCE data in addition to the data types considered in the production of EGM2008; two GRACE-only models, namely, ITG-Grace03, which served as GRACE pseudo-data in the production of EGM2008, and ITG-Grace2010s as a state-of-the-art GRACE-only model; and three combined (satellite-only) GRACE/GOCE models, namely, GOCO01S, EIGEN-6S and GOCO02S. The eighth model in the considered list is a new Delft Gravity Model, release 1, Satellite-only (DGM-1S) developed at Delft University of Technology in collaboration with GNSS Research Centre of Wuhan University on the basis of GRACE and GOCE data.

The manuscript is structured as follows. Section 2 is devoted to the description of the validation methodology. The results of its application to six of the previously mentioned models (EGM2008, ITG-Grace2010s, GOCO01S, EIGEN-6S, GOCO02S and DGM-1S) are presented in Section 3. In this section, we (i) compare performance of the combined GRACE/GOCE models; and (ii) study the added value of the GOCE mission to the static gravity field modelling. The latter is performed by comparing the performance of the combined (satellite-only) GRACE/GOCE models with that of either EGM2008 or ITG-Grace2010s depending on which of the two performs better in a given context. In Section 4, as an example of another potential application of the proposed validation procedure, we inspect how successfully terrestrial gravimetry and satellite altimetry data have been combined with ITG-Grace03 in the production of EGM2008 and with GRACE and GOCE data in the production of EIGEN-6C. To that end, we analyse the difference in performance between (a) EGM2008 and ITG-Grace03; and (b) between EIGEN-6C and its satellite-only counterpart, EIGEN-6S. Section 5 is left for discussion and conclusions. Furthermore, Appendix A contains some basic information about the computation of DGM-1S. An in-depth presentation of this model is given in a separate publication (Hashemi Farahani *et al.* 2013).

2 METHODOLOGY

In this section, we describe (i) the functional models exploited to deal with GRACE KBR and GOCE SGG control data in the validation procedure; (ii) the way in which misfit of static gravity field models to these data are computed; and (iii) further steps required prior to the usage of misfit data for an analysis of the quality of the models.

2.1 Validation data

2.1.1 GOCE gravity gradients

The SGG control data from the GOCE mission are processed in the Gradiometer Reference Frame (GRF): a right-handed satellite-fixed frame whose x, y and z axes under normal circumstances are respectively in the along-track (flight), cross-track (orthogonal to the orbital plane) and radial (downwards) directions with an approximation of a few degrees (Rummel *et al.* 2011). It is important to add that these measurements can be considered as '*in situ*'. This means that each of them reflects the gravity field only in the vicinity of the measurement point. In other words, the functional model that relates parameters of the gravitational field with these data is 'local'. Only the accurately measured components in these data, namely, $V_{xx} = \frac{\partial^2 V}{\partial x^2}$, $V_{yy} = \frac{\partial^2 V}{\partial y^2}$, $V_{zz} = \frac{\partial^2 V}{\partial z^2}$ and $V_{xz} = \frac{\partial^2 V}{\partial x \partial z}$ (with V being the gravitational potential) are used in the validation of the models. They are hereafter referred to as *xx*, *yy*, *zz* and *xz* components, respectively.

2.1.2 GRACE range combinations

Local functional models are definitely preferable for a model validation, because they facilitate the assessment in different geographical regions, individually. Therefore, we introduce a local functional model in the context of GRACE KBR control data as well. More specifically, we use these data in the form of so-called 'range combination', proposed by Liu (2008). A range combination $\bar{a}(t)$ is defined as

$$\bar{a}(t) = \frac{\cos\theta_{(-)}\,\rho(t-\Delta t) - 2\rho(t) + \cos\theta_{(+)}\,\rho(t+\Delta t)}{(\Delta t)^2}, \qquad (1)$$

where Δt denotes the sampling interval, $\rho(t-\Delta t)$, $\rho(t)$ and $\rho(t+\Delta t)$ are bias-corrected KBR ranges (see Liu 2008) at three successive epochs, and $\theta_{(\pm)}$ are the angles between the line-of-sight at the middle epoch and the line-of-sight at the previous and subsequent epochs, respectively. This means that $\cos\theta_{(\pm)} = \mathbf{e}(t)\cdot\mathbf{e}(t\pm\Delta t)$ with $\mathbf{e}(t-\Delta t)$, $\mathbf{e}(t)$ and $\mathbf{e}(t+\Delta t)$ being the unit vectors defining the line-of-sight direction at the three successive epochs. One can see that eq. (1) is close to a numerical differentiation with a three-point scheme. Therefore, a range combination is approximately equal to the line-of-sight component of the inter-satellite acceleration vector (Ditmar *et al.* 2012). A more detailed description of the concept of range combinations can be found in (Liu 2008). In particular, it is

shown there that a range combination is equal to the intersatellite acceleration averaged with a certain weight within the differentiation time interval and projected onto the line-of-sight of the central epoch.

2.2 Computing residual data

The validation procedure makes use of the misfits between model-based quantities and observations. These misfits are hereafter referred to as 'residual data' for the sake of brevity. The comparison between sets of residual data associated with various static gravity field models allows conclusions to be drawn regarding the quality of those models. It is worth reminding that observations naturally refer to the instantaneous gravity field, which experiences temporal variations. In addition, GRACE KBR measurements are influenced by non-gravitational forces. Therefore, a computation of residual data requires that either observations are reduced for nuisance signals (associated with temporal gravity field variations and, possibly, non-gravitational forces) or model-based quantities are computed based on a force model that in addition to the static gravity field model takes these nuisance signals into account. The former approach is followed in the context of SGG control data, whereas the latter one in the context of KBR control data.

2.2.1 Computing GOCE residual gravity gradients

The GOCE residual gravity gradients are computed on the basis of the input provided by the European Space Agency (ESA): (1) SGG measurements in the GRF with 1-s sampling; and (2) reduced-dynamic orbits with 10-s sampling (Gruber *et al.* 2010; de Witte 2011). The latter ones are interpolated at the gradiometer measurement epochs using an eleventh-order Legendre interpolation scheme. The model based gravity gradients are evaluated at those positions. The SGG measurements are corrected for the following time-varying signals: (i) direct (astronomical) tides modelled with the Jet Propulsion Laboratory (JPL) DE405 and LE405 lunar and planetary ephemerides (Standish 1998); (ii) solid Earth and pole tides (McCarthy and Petit 2004); (iii) ocean tides given by the FES2004 model (Lyard *et al.* 2006) and (iv) non-tidal mass redistribution in the atmosphere and oceans described by the fourth release of the Atmosphere and Ocean De-aliasing level-1B (AOD1B) product (Flechtner 2007).

2.2.2 Computing GRACE residual range combinations

Computation of GRACE residual range combinations is somewhat more complicated. The force model defined above has to be extended further. First, non-gravitational forces measured by the onboard accelerometers (Case *et al.* 2004) are included into the force model. Secondly, in view of a high sensitivity of GRACE KBR data to temporal variations of the gravity field, the list of time-varying signals, applied in the computation of GOCE-based residuals, is complemented with additional minor signals associated with (i) relativistic effects (McCarthy and Petit 2004) and (ii) ocean pole tide (Desai 2002).

Furthermore, long-term (i.e. with a characteristic time longer than 1 month) gravity field temporal variations are taken into account. At many geographical locations, a prominent annual cycle of mass variations of hydrological origin takes place. Moreover, in some regions (mostly located in the polar areas), a steady accumulation or loss of mass occurs due to postglacial rebound and shrinking of polar ice sheets. Therefore, a discrepancy between a static gravity field model and GRACE KBR control data collected at a certain moment of time may reveal not only the model's inaccuracies but also an evolution of the gravity field in the course of time. To mitigate this effect, we complement the force model with a term described by eq. (2), namely, a model $m^{(\mathrm{LT})}(t)$ of Long Term (LT) gravity field variations. To that end, we use the release four of the Centre for Space Research (CSR)'s GRACE-based monthly gravity field solutions $m^{(\mathrm{CSR})}(t)$ (Bettadpur 2007) processed with an anisotropic filter in accordance with (Kusche *et al.* 2009). Before using these solutions, we correct them for a non-zero mean $\overline{m^{(\mathrm{CSR})}(t)}$:

$$m^{(\mathrm{LT})}(t) = m^{(\mathrm{CSR})}(t) - \overline{m^{(\mathrm{CSR})}(t)}. \tag{2}$$

The latter is computed by averaging the monthly gravity field solutions available in the time interval that coincides with the one covered by GRACE KBR data used in the production of the static gravity field model under assessment. The necessity of this operation can be understood from the fact that long-term variations of the gravity field are not usually removed from GRACE KBR data prior to the computation of a static gravity field model. Since GRACE KBR data are practically the only source of information about temporal gravity field variations, such a model is nothing but the mean gravity field in the time interval covered by GRACE KBR data used in its computation. Thus, by subtracting the non-zero mean $\overline{m^{(\mathrm{CSR})}(t)}$ computed over the same time interval, we ensure that $m^{(\mathrm{LT})}(t)$ approximates nothing but the deviation of the instantaneous gravity field from the reference level defined by the static gravity field model under consideration. This deviation includes, among others, annual variations and linear trends.

An auxiliary input required for computing GRACE residual range combinations are kinematic (or reduced-dynamic) orbits of the GRACE satellites, which were produced in-house. On the basis of those orbits and the adopted force model, we compute dynamic orbits of the GRACE satellites using the Positioning And Navigation Data Analyst (PANDA) software package (Zhao 2004). The duration of orbital arcs is 6 hr. Twelve parameters per orbital arc are estimated: six state vector parameters, three accelerometer bias parameters and three accelerometer scaling factors. The use of the dynamic orbits is twofold. First, they are exploited to compute the intersatellite ranges, which are then subtracted from the observed ones to form the residual ranges. The latter ones are needed to compute residual range combinations. Secondly, these orbits are considered as the source of information about the line-of-sight orientations, which are used to compute the angles $\theta_{(-)}$ and $\theta_{(+)}$ when the residual ranges are converted into residual range combinations, *cf.* eq. (1).

2.3 Suppressing noise

In the first instance, the validation is performed in the spectral domain on the basis of the Power Spectral Density (PSD) of the obtained residual data. However, a spectral analysis of residuals does not allow identifying the geographical regions responsible for misfits of static gravity field models to control data. Therefore, we find it essential to analyse the residual data sets in the spatial domain as well. This requires improving signal-to-noise ratio in the sets of residual data. It is worth noting that 'signal' in this context is defined as inaccuracies of a given static gravity field model propagated into the corresponding sets of residual data.

The residual data of both types show an increased level of noise at low frequencies. It exceeds the noise level at mid-frequencies

approximately 30 and 300 times in the case of the GRACE and GOCE residual data, respectively. In the residuals of the former type, this noise is eliminated using a high-pass filter based on an empirical model composed of seven parameters (Kim 2000):

$$r(t) = x_0 + x_1 t + x_2 \cos\omega t + x_3 \sin\omega t + x_4 t \cos\omega t + x_5 t \sin\omega t + x_6 t^2, \quad (3)$$

where $r(t)$ is a value estimated by the empirical model, $\omega = \frac{2\pi}{T}$ is the orbital angular velocity with T being the orbital revolution time, and $x_0, x_1, \ldots,$ and x_6 are unknown parameters. These parameters are estimated per orbital revolution by means of a least-squares adjustment. Subsequently, the values synthesized with the empirical model are subtracted from GRACE residual range combinations. This operation leads to the elimination of noise up to approximately three cycles-per-revolution (cpr) frequency. Elimination of the low-frequency noise in the case of GOCE residual gravity gradients is achieved in a similar manner, but the empirical model includes a bias, a linear trend and periodic terms up to 27-cpr frequency:

$$r(t) = x_0 + x_1 t + \sum_{k=1}^{27} \{x_{(2k)} \cos k\omega t + x_{(2k+1)} \sin k\omega t\}, \quad (4)$$

where $x_0, x_1, \ldots,$ and x_{55} are unknown parameters. The application of this filter eliminates noise in GOCE residual gravity gradients up to the 27-cpr frequency, which corresponds to the lower bound of the gradiometer measurement band, that is, 5 mHz (Rummel *et al.* 2011).

Furthermore, we find it essential to suppress noise as a whole, including that in high-frequency range where the instrumental noise is particularly strong (above 25 mHz and 100 mHz in the case of GRACE KBR and GOCE SGG data, respectively). To that end, we map the residuals onto the Earth's surface and compute mean residuals per $\Delta^\circ \times \Delta^\circ$ block, where Δ° is the block size in degrees. In the case of GRACE KBR data, each block-mean value is computed on the basis of all measurements for which the mid-point between the GRACE satellites is located inside that block. Assuming that noise in different orbital tracks is uncorrelated, computation of block-mean values results in the suppression of noise at least by the factor equal to the square root of the average number of orbital tracks crossing the blocks. Thus, the longer the duration of a data set considered in the validation procedure is, the better signal-to-noise ratio is expected.

The computation of block-mean values also leads to another positive outcome. It operates as low-pass filtering of residual data. The GRACE and GOCE satellites cross the distance of Δ° in about $t_{\Delta^\circ} \approx \frac{\frac{\Delta^\circ \times \pi}{180^\circ} \times 6370}{v}$ s, where $v \approx 7.4$ km s^{-1} and $v \approx 8$ km s^{-1}, respectively denote the GRACE and GOCE satellites' ground speed and 6370 is the Earth's mean equatorial radius in kilometres. Thus, noise and signal at the frequencies above $f_{\Delta^\circ} = \frac{1}{t_{\Delta^\circ}}$ Hz or equivalently above $f_{\Delta^\circ} \times 5400$ cpr are largely averaged out (5400 is the satellites' orbital revolution in seconds). The frequency $f_{\Delta^\circ} \times 5400$ cpr can be approximately associated with spherical harmonic coefficients of degree $N_{\Delta^\circ} = f_{\Delta^\circ} \times 5400$ and less. This means that the validation of a model becomes largely limited to its coefficients below degree N_{Δ°. For instance, choosing Δ° equal to 6° or 3° associates the outcome of the validation largely with the coefficients of models below degree 60 or 120, respectively. Given a unique sensitivity of GRACE KBR data to low degree spherical harmonics of the gravity field, this is extremely beneficial when assessment of low degree coefficients of a model is the primary focus. In that case, it is sufficient to choose Δ° in accordance with the maximum degree up to which the validation of a model is to be performed.

3 APPLICATION

In this section, we use the methodology described earlier to perform a validation of six static gravity field models already mentioned in Section 1.

(i) ITG-Grace2010s, which is complete up to degree 180 and based on 7 yr of GRACE KBR and kinematic orbit data (August 2002–August 2009).

(ii) GOCO01S, which is complete up to degree 224 and based on a combination of ITG-Grace2010s with 2 months of GOCE SGG data (2009 November–2009 December).

(iii) GOCO02S, which is complete up to degree 250 and based on a combination of ITG-Grace2010s with 8 months of GOCE SGG data, 12 months of GOCE kinematic orbit data, 8 yr of the Challenging Mini-satellite Payload (CHAMP) kinematic orbit data, and 5 yr of SLR data from five satellites.

(iv) EIGEN-6S, which is complete up to degree 240 and based on 7.5 yr of GRACE KBR and kinematic orbit data (January 2003–June 2009), 6.7 months of GOCE SGG data (2009 November–2010 June) and 6.5 yr of SLR data from Laser Geodynamics Satellites (LAGEOS). In this study, we only consider the static part of this model.

(v) DGM-1S, which is complete up to degree 250 and based on 7 yr of GRACE KBR data (2003 February–2009 December), 4 yr of GRACE kinematic orbit data (2006 January–2009 December), 10 months of GOCE SGG data (2009 November–2010 December) and 14 months of GOCE kinematic orbit data (2009 August–2010 December). Appendix A contains a more detailed description of this model.

(vi) EGM2008, which is complete up to degree 2159 (with some coefficients up to degree 2190) and based on a combination of ITG-Grace03 with an extended set of terrestrial gravimetry and satellite altimetry data. In this study, we truncate this model at degree 250 to be consistent with the maximum degree of GOCO02S and DGM-1S and speed up the computations. In any case, the satellite control data used in this manuscript are hardly sensitive to coefficients above degree 250.

Fig. 1 shows the geoid height difference per degree between the first five models and EGM2008. One can see that GOCO02S and DGM-1S in this representation agree with EGM2008 almost equally well, whereas ITG-Grace2010s, GOCO01S and EIGEN-6S show slightly larger

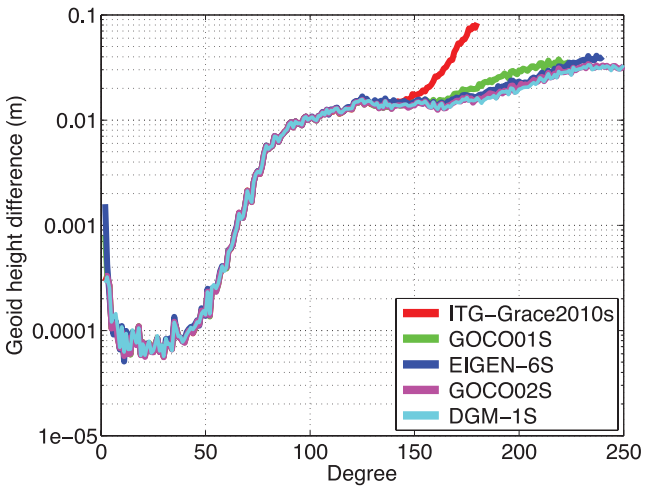

Figure 1. Geoid height difference per degree with respect to EGM2008.

deviations from it at high degrees. Unfortunately, such a comparison hardly allows conclusions to be drawn regarding the actual accuracy of the six models under consideration. The following questions remain open: (1) How to interpret the disagreement between EGM2008 on the one hand and the other five models on the other hand that rapidly increases above degree 50? Is it an evidence of a lower quality of EGM2008 or an indication of an insufficient accuracy of the other five models? (2) Is it fair to interpret a relatively poor agreement of ITG-Grace2010s, GOCO01S and EIGEN-6S with EGM2008 as an indication of a lower accuracy of these three models as compared to GOCO02S and DGM-1S? (3) Is it fair to state that the accuracy of models that match EGM2008 equally well is equal? In this section, we attempt to answer these questions using the proposed validation methodology. The control data utilized for this purpose are:

(i) GRACE range combinations based on KBR data covering all 12 months of 2010 (5-s sampling).

(ii) GOCE gravity gradients based on SGG data from 2011 March to May (1-s sampling).

These data have not been used in the production of any of the considered models. The original data are converted into sets of residuals associated with these models as described in the previous section. In the course of the validation of the models, we also compare the ability of GRACE KBR and GOCE SGG data to reveal model inaccuracies. To make this comparison equitable, we find it important to ensure the same length of the data sets. Therefore, in addition to the yearly set of GRACE range combinations, we also consider a subset of it with the length equal to that of the considered GOCE SGG data (i.e. 3 months) using the data collected in 2010 August–October.

3.1 Validation against GRACE KBR data

3.1.1 Validation in the spectral domain

Fig. 2 shows the square-root of PSD of the sets of GRACE residual range combinations associated with the six aforementioned models (data of 2010 February are used). The shown frequency range in Fig. 2(a) is limited to the one in which a pronounced difference is observed between the performance of EGM2008 and that of the other five models, whereas Fig. 2(b) zooms in the frequency range 3–11 mHz, where some differences are observed between consistencies of the latter models with the KBR control data.

Fig. 2(a) allows one to conclude that EGM2008 is of much lower accuracy than all the other considered models in the frequency range 5–22 mHz (27–120 cpr), which corresponds to signals at spatial scales of 180–800 km (half wavelengths). Practically no difference can be identified in Fig. 2(a) between the performance of the considered models above the upper bound of this frequency range. ITG-Grace2010s and the four combined GRACE/GOCE models in this representation seem to match the GRACE KBR control data almost equally well, so that the corresponding five curves are hardly distinguishable in Fig. 2(a). However, the zoomed-in picture (Fig. 2b) allows some differences in the model performance to be revealed in the frequency range 4–10 mHz (22–54 cpr), which corresponds to signals at spatial scales of 400–990 km. In this frequency range, DGM-1S shows a slightly better agreement with the control data than the other combined GRACE/GOCE models and ITG-Grace2010s.

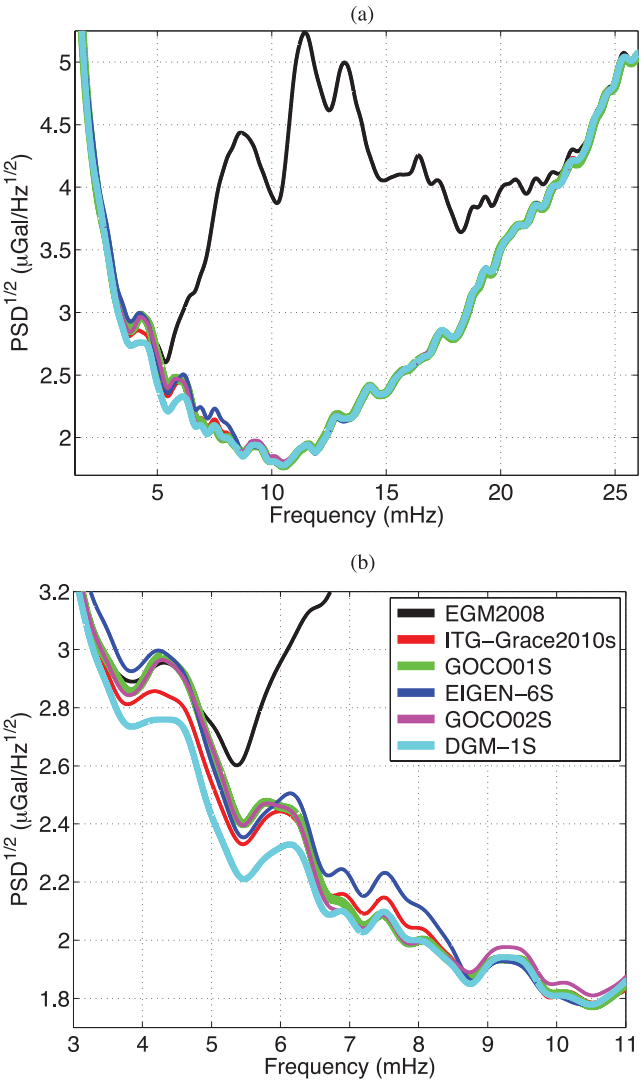

Figure 2. Square-root of PSD of the GRACE residual range combinations for 2010 February.

3.1.2 Validation in the spatial domain

To begin with, we set Δ° equal to 1°, which results in the elimination of high-frequency noise in residual range combinations above 66 mHz. It also leads to a suppression of noise as a whole with a factor of approximately $\sqrt{33} \approx 5.7$, where 33 is the average number of GRACE orbital tracks crossing one block in 2010. For the 3-month data subset, this factor reduces to 3. It is important to note that the chosen size of the blocks allows gravity signal to be largely preserved, since a full suppression of signal would occur only around degree 360, whereas all the considered models are complete, at maximum, up to degree 250. This means that the 'total' accuracy of the models is analysed to the extent possible with GRACE KBR data. Fig. 3 shows the sets of GRACE $1^\circ \times 1^\circ$ block-mean residual range combinations globally. The 112-km wide (in diameter) polar gaps in the GRACE spatial coverage are absent in these maps.

A visual inspection of the maps presented in Fig. 3 suggests splitting the continental areas into three categories: (1) areas that have a poor coverage with terrestrial gravimetry measurements, where EGM2008 shows a relatively poor agreement with the GRACE KBR control data, whereas the other considered models match them

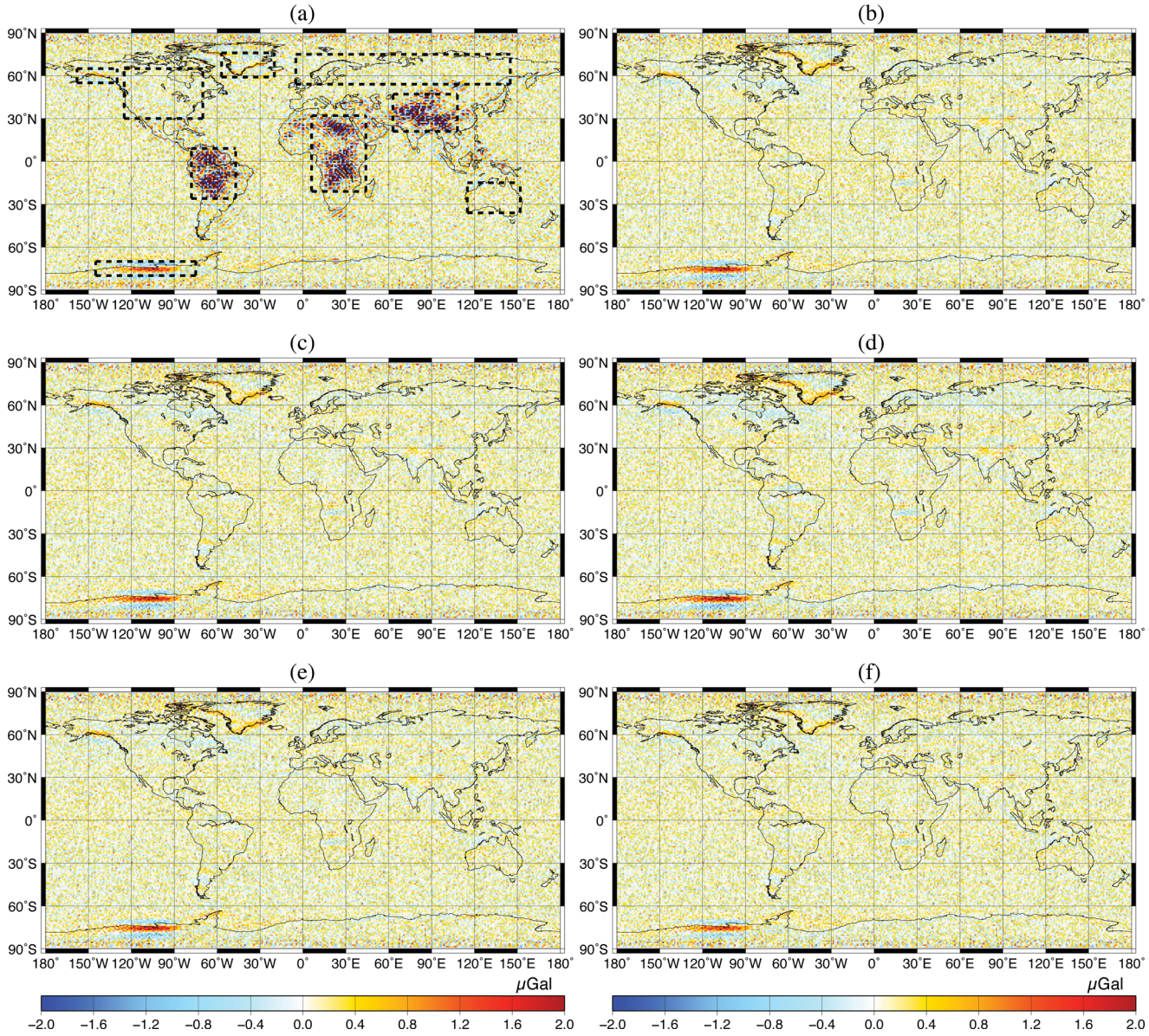

Figure 3. The GRACE 1° × 1° block-mean residual range combinations obtained for (a) EGM2008, (b) ITG-Grace2010s, (c) GOCO01S, (d) EIGEN-6S, (e) GOCO02S and (f) DGM-1S on the basis of the 1-yr data set. The rms misfits are (a) 0.667 μGal, (b) 0.382 μGal, (c) 0.382 μGal, (d) 0.387 μGal, (e) 0.382 μGal and (f) 0.378 μGal. The nine rectangular regions for which regional rms misfits are computed are marked in Fig. 3(a) with block dashed lines.

much better (e.g. the Himalayas, the Northern part of South America and the Equatorial Africa); (2) areas that are well surveyed by terrestrial gravimetry campaigns, where all the six models visually demonstrate a good match to the control data (e.g. Australia, North America and the Northern part of Eurasia) and (3) polar areas with a systematic linear trend in mass variations (mainly caused by a steady ice mass loss there), where all the six models show an increased level of disagreement with the control data (e.g. the South coast of Alaska, the Southern part of Greenland and the coast of Amundsen sea in Antarctica). In addition, we introduce the fourth category: the world's oceans within the latitudes 65°S and 65°N, which is characterized by an excellent coverage with TOPEX/Poseidon satellite altimetry data (see Sandwell & Smith 2009). A good match with the control data is observed there for all the six models. To obtain some quantitative estimates, we define three rectangular regions in the areas of the first, second and third categories (nine regions in total), which are marked in Fig. 3(a). In Table 1, the Root Mean Square (rms) of the sets of GRACE 1° × 1° block-mean residual range combinations related to the considered models are presented for each of these nine rectangular regions as well as for the world's oceans (65°S–65°N) and for the entire globe, both for the 1-yr data set and the 3-month subset.

According to Table 1, the 1-yr data set as compared to the 3-month subset shows a significant reduction of misfit to the considered models: by 20–42 per cent. This is probably a consequence of a more efficient suppression of random noise when a longer data set is considered. The exception is EGM2008 in the regions of the first category, where this reduction is only 5–6 per cent. This is probably because the signal-to-noise ratio in this case is significantly higher than in the other cases. On the other hand, for some combinations of the region and model, this reduction approaches a factor of $\sqrt{\frac{12}{3}}=2$ (with 12 and 3, respectively, referring to the lengths of the 1-yr

Table 1. The rms of the GRACE 1° × 1° block-mean residual range combinations (in μGal) per region, obtained from the 1-yr data set (in the numerator) and from the 3-month subset (in the denominator). The nine considered rectangular regions are defined by longitudinal and latitudinal boundaries marked in Fig. 3(a) with dashed, black lines.

Category	Region	Longitudinal range	Latitudinal range	EGM2008	ITG-Grace2010s	GOCO01S	EIGEN-6S	GOCO02S	DGM-1S
	'Himalayas'	63°E–108°E	21°N–47°N	2.323 2.452	0.363 0.599	0.358 0.596	0.370 0.604	0.361 0.594	0.354 0.592
1	'South America'	78°W–47°W	26°S–9°N	2.211 2.335	0.347 0.557	0.351 0.559	0.356 0.562	0.350 0.559	0.352 0.557
	'Equatorial Africa'	6°E–44°E	21°S–32°N	1.656 1.768	0.354 0.563	0.356 0.563	0.357 0.568	0.354 0.562	0.351 0.562
	'Australia'	115°E–152°E	36°S–15°S	0.386 0.626	0.356 0.614	0.353 0.612	0.357 0.609	0.352 0.607	0.351 0.608
2	'North Eurasia'	5°W–145°E	50°N–75°N	0.385 0.588	0.357 0.569	0.356 0.568	0.366 0.576	0.356 0.568	0.356 0.568
	'North America'	125°W–70°W	30°N–65°N	0.368 0.567	0.350 0.553	0.353 0.559	0.358 0.561	0.350 0.553	0.346 0.553
	'South Alaska'	158°W–130°W	55°N–65°N	0.458 0.657	0.453 0.641	0.449 0.640	0.458 0.644	0.450 0.640	0.418 0.610
3	'Amundsen sea coast'	145°W–75°W	80°S–70°S	0.701 0.872	0.743 0.913	0.745 0.915	0.791 0.959	0.746 0.917	0.639 0.823
	'South Greenland'	57°W–20°W	59°N–76°N	0.444 0.659	0.406 0.638	0.405 0.640	0.431 0.661	0.405 0.640	0.372 0.614
4	'World's oceans'	180°W–180°E	65°S–65°N	0.386 0.595	0.343 0.568	0.343 0.567	0.344 0.568	0.343 0.567	0.343 0.567
	'Globe'	180°W–180°E	89°S–89°N	0.667 0.846	0.382 0.630	0.382 0.629	0.387 0.632	0.382 0.629	0.378 0.626

data set and the 3-month subset), which is the maximum value one can expect under the assumption that noise time-series in different months are not correlated with each other. This allows us to conclude that random noise in the control data plays a substantial, if not dominant, role in the obtained rms misfits. Nevertheless, these rms values still show some differences between the models and between the regions, which means that they also contain valuable information for a model validation. Further analysis is fully based on the misfits to the 1-yr data set, as they are proven to be less contaminated by random noise.

The rms misfits indicate that ITG-Grace2010s and the combined GRACE/GOCE models demonstrate a much higher accuracy than EGM2008 in the regions from the first category: by 79–85 per cent. We find it worth discussing the origin of this difference. EGM2008 utilized GRACE data by including the GRACE-based ITG-Grace03 into the data combination. ITG-Grace03 uses approximately 50 per cent less GRACE data than ITG-Grace2010s. Therefore, the latter model is statistically expected to be more accurate than the former one by a factor of only $\sqrt{1.5} \approx 1.2$ (or 20 per cent). This statistically expected better performance will be confirmed in Section 4.1 by the SGG control data. This means that only a small portion of the performance difference between ITG-Grace2010s and EGM2008 in the first category regions can be attributed to the contribution of the extra GRACE KBR data incorporated into ITG-Grace2010s. The rest of it can only be explained by a lower accuracy of EGM2008 as compared to ITG-Grace03 in these areas. One may find this as an unexpected outcome. Ideally, in an optimal combination of a GRACE-only model with terrestrial gravimetry measurements, the resulting model should perform in the gravimetrically poorly surveyed continental areas at least as well as the GRACE-only model. Nevertheless, as reported by Pavlis *et al.* (2012) themselves, a loss of information content of ITG-Grace03 has indeed occurred in the production of EGM2008 over the areas that are poorly surveyed gravimetrically. In Section 4, we utilize the proposed validation procedure to analyse this deficiency further and present a quantification of its severity.

Furthermore, Table 1 allows us to conclude that ITG-Grace2010s and the combined GRACE/GOCE models consistently demonstrate a slightly higher accuracy than EGM2008 in the regions from the second category (by 5–9 per cent) and in the 'World's oceans' (by about 11 per cent). In addition, Table 1 reveals a comparable performance of ITG-Grace2010s and of the four combined GRACE/GOCE models in the regions belonging to the first and second categories as well as in the 'World's oceans'. In most cases, DGM-1S matches the GRACE KBR control data slightly better than the other models. For instance, in 'Himalayas' DGM-1S seems to outperform ITG-Grace2010s (or GOCO02S), GOCO01S and EIGEN-6S by about 2, 1 and 4 per cent, respectively. Given the fact that ITG-Grace2010s is fully independent from GOCE data and GOCO01S uses only a very limited set of them, it is worth noting that that these two models show practically the same agreement with the GRACE control data as EIGEN-6S, GOCO02S and DGM-1S. They even show a slightly smaller misfit than EIGEN-6S in some regions (e.g. in 'Himalayas' and 'North Eurasia': by 2–3 per cent). We explain this by a low sensitivity of the GRACE KBR control data to the contribution of the GOCE SGG data used in the production of the combined GRACE/GOCE models, which manifests itself at relatively high degrees (see Fig. 1).

In the regions from the third category, DGM-1S fits the GRACE KBR control data by 9–19 per cent better than the other models. Though such a difference seems to be substantial, we recommend interpreting this result with a caution. Time-varying gravity field signals in the regions of the third category are very strong. To demonstrate that, we have re-computed block-mean residual range combinations without including the CSR monthly solutions into the force model. The results obtained this way on the basis of DGM-1S are shown in Fig. 4. A comparison of it with Fig. 3(f) demonstrates that these signals are largely removed in our model validation procedure. It is very likely, however, that the CSR solutions are unable to remove these signals from the GRACE KBR data completely. This may have been caused by a variety of reasons, for example, a limited temporal resolution of the CSR solutions (1 month), a limited spatial resolution of them (degree 60) or a suppression of signal due to anisotropic filtering. Consequently, the rms misfits obtained in the third category regions may be contaminated with relatively large systematic errors.

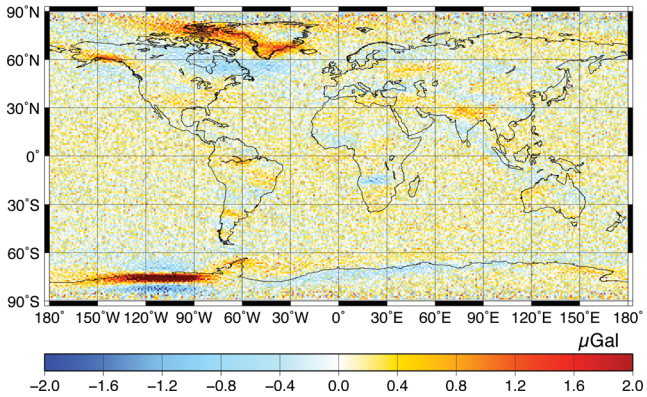

Figure 4. The GRACE 1° × 1° block-mean residual range combinations obtained based on DGM-1S without complementing the force model with eq. (2). The rms number is 0.443 μGal.

Globally, the agreement of EGM2008 with the GRACE KBR control data is worse than of the other five models: by 42–43 per cent. The latter models agree with the KBR data globally almost equally well (DGM-1S matches these data by 1–2 per cent better than ITG-Grace2010s, GOCO01S, EIGEN-6S and GOCO02S).

Due to the fact that the accuracy of satellite-only models reduces with increasing degrees (see Fig. 1), the validation based on 1° × 1° blocks may be insufficiently sensitive to errors in low-degree coefficients. To shed more light on this issue, we perform the validation of the models using 3° × 3° and 6° × 6° block-mean residual range combinations obtained from the 1-yr data set. The average number of GRACE orbital tracks crossing the 3° × 3° and 6° × 6° blocks in year 2010 is 112 and 228, respectively. This allows noise in the sets of residual range combinations to be suppressed by a factor of $\sqrt{112} \approx 11$ and $\sqrt{228} \approx 15$, respectively. Figs 5 and 6 show the obtained sets of 3° × 3° and 6° × 6° block-mean residuals, respectively. The corresponding rms values computed in the previously defined regions are presented in Table 2. The conducted

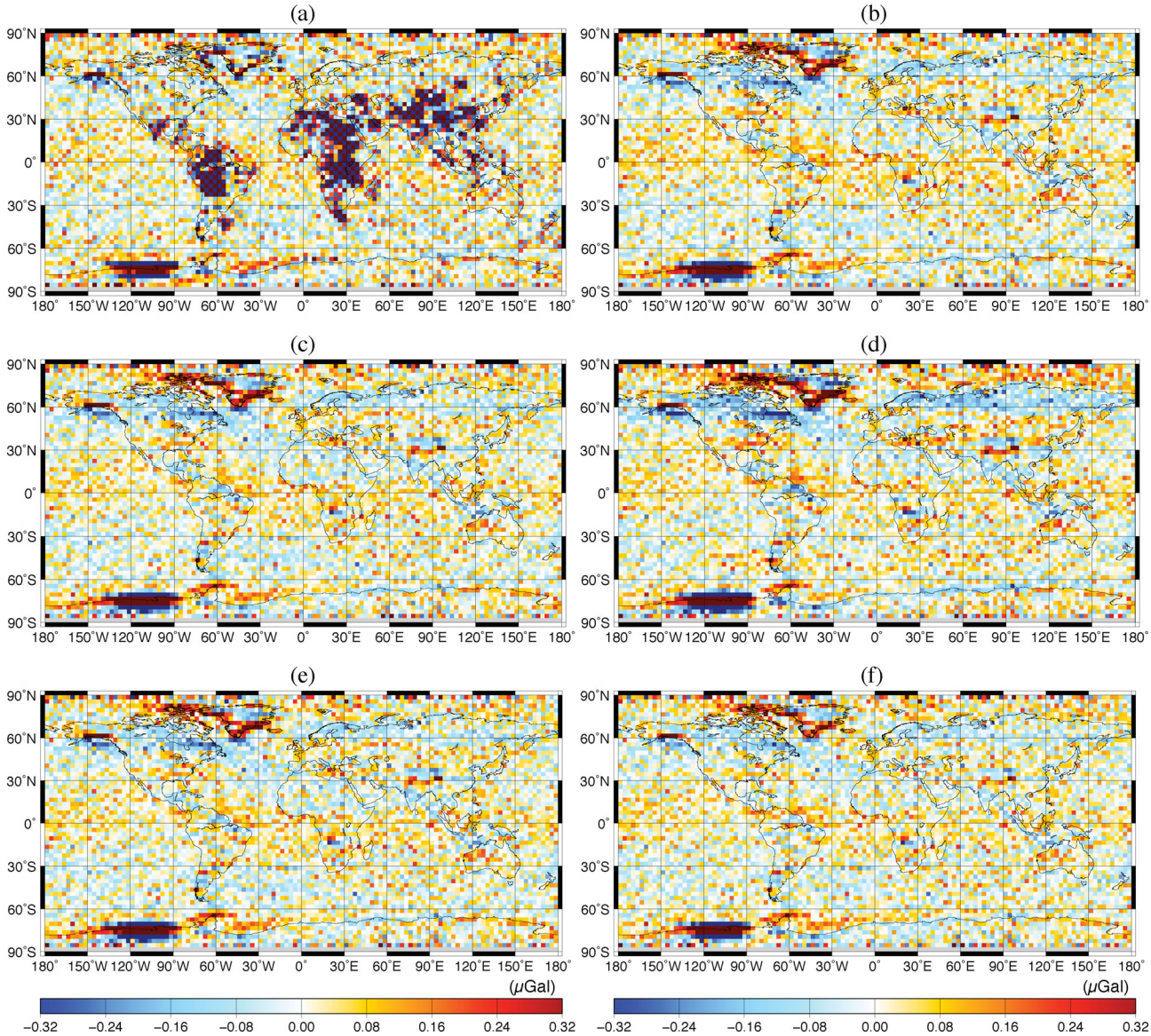

Figure 5. The GRACE 3° × 3° block-mean residual range combinations obtained for (a) EGM2008, (b) ITG-Grace2010s, (c) GOCO01S, (d) EIGEN-6S, (e) GOCO02S and (f) DGM-1S on the basis of the 1-yr data set. The rms misfits are (a) 0.262 μGal, (b) 0.116 μGal, (c) 0.116 μGal, (d) 0.127 μGal, (e) 0.116 μGal and (f) 0.105 μGal.

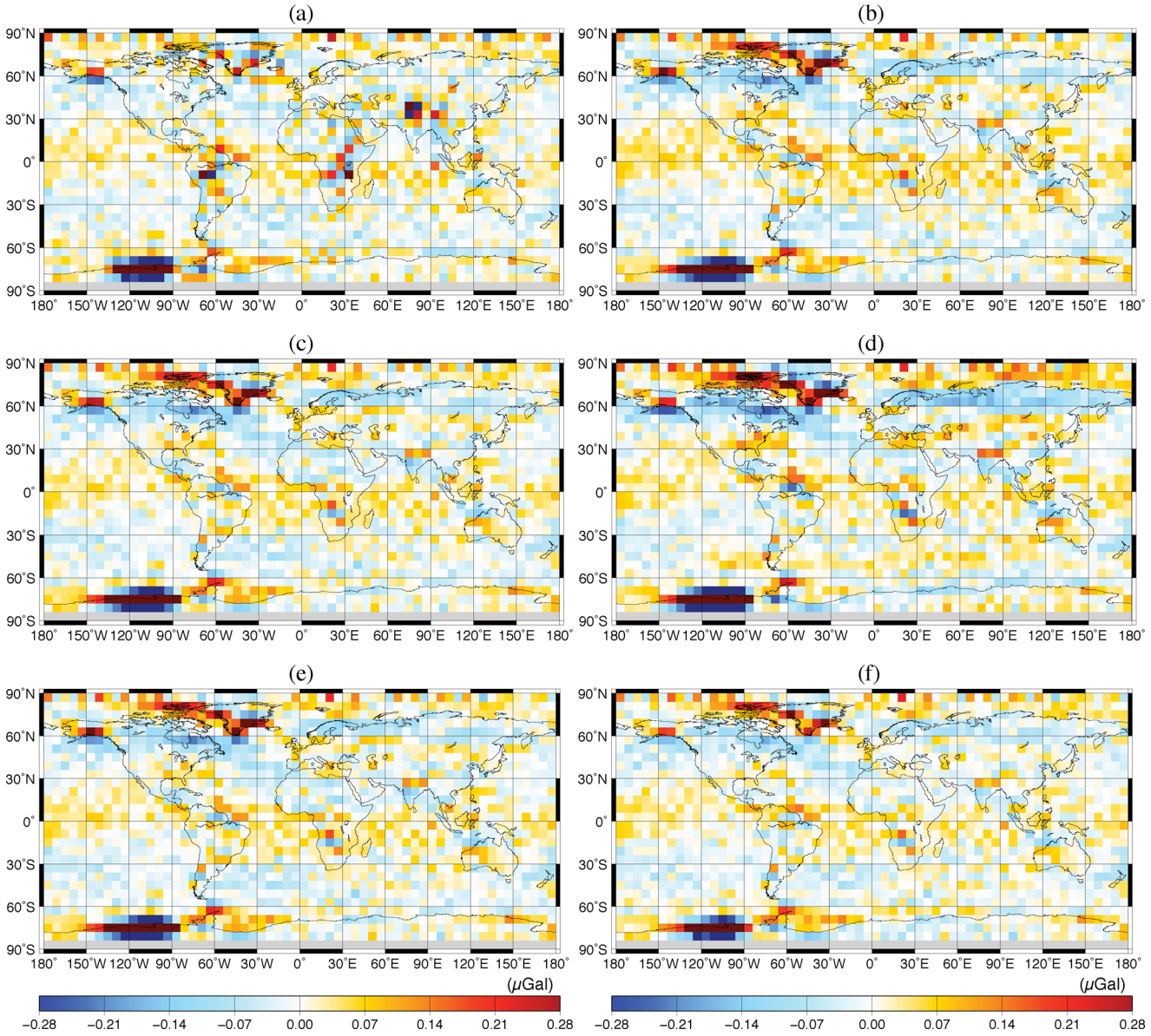

Figure 6. The GRACE 6° × 6° block-mean residual range combinations obtained for (a) EGM2008, (b) ITG-Grace2010s, (c) GOCO01S, (d) EIGEN-6S, (e) GOCO02S and (f) DGM-1S on the basis of the 1-yr data set. The rms misfits are (a) 0.072 μGal, (b) 0.083 μGal, (c) 0.083 μGal, (d) 0.096 μGal, (e) 0.083 μGal and (f) 0.068 μGal.

analysis shows in particular that the low-degree part of DGM-1S is globally by up to 10 per cent more consistent with the KBR control data than the other models. We find it also remarkable that in the case of 6° × 6° blocks, EGM2008 demonstrates globally and outside the poorly studied continental areas a higher level of consistency with the KBR control data than the other models.

3.2 Validation against GOCE SGG data

3.2.1 Validation in the spectral domain

Fig. 7 shows the spectra of GOCE residual gravity gradients associated with all the six models under consideration (the data of 2011 May are used). The shown frequency range in these plots is limited to the one in which differences between the models can be visually identified.

These pictures demonstrate a lower performance of EGM2008 as compared to the other five models in the frequency range 10–28 mHz (54–150 cpr), which corresponds to signals at spatial scales of 140–390 km (half wavelengths). This largely confirms the outcome of the validation against the GRACE KBR data, but the pronounced differences in that case appeared at lower frequencies: 5–22 mHz. Furthermore, Fig. 7 demonstrates a much higher accuracy of EGM2008 and the combined GRACE/GOCE models as opposed to ITG-Grace2010s in the frequency range 25–37 mHz or 135–200 cpr (spatial scales of 105–155 km). This can be explained by the high-frequency information content in GOCE SGG data exploited in the combined GRACE/GOCE models and that in the terrestrial gravimetry and satellite altimetry data exploited in EGM2008. As far as the combined GRACE/GOCE models are concerned, the curves associated with them are practically indistinguishable in the plots. An exception is the frequency range 27–37 mHz or 146–200 cpr (spatial scales of 105–140 km), where

Table 2. The rms of the GRACE 3° × 3° (in the numerator) and 6° × 6° (in the denominator) block-mean residual range combinations (in μGal) per region obtained from the 1-yr data set.

Region	EGM2008	ITG-Grace2010s	GOCO01S	EIGEN-6S	GOCO02S	DGM-1S
'Himalayas'	0.993 0.184	0.116 0.059	0.116 0.057	0.130 0.072	0.116 0.058	0.104 0.052
'South America'	0.963 0.138	0.086 0.040	0.086 0.039	0.101 0.062	0.086 0.039	0.092 0.050
'Equatorial Africa'	0.783 0.124	0.094 0.047	0.095 0.048	0.101 0.052	0.095 0.048	0.089 0.041
'Australia'	0.113 0.049	0.096 0.057	0.094 0.053	0.101 0.059	0.095 0.056	0.088 0.048
'North Eurasia'	0.097 0.036	0.093 0.065	0.090 0.060	0.118 0.101	0.090 0.061	0.084 0.053
'North America'	0.092 0.040	0.111 0.079	0.109 0.078	0.135 0.106	0.109 0.078	0.097 0.063
'South Alaska'	0.266 0.134	0.263 0.216	0.261 0.215	0.260 0.173	0.263 0.219	0.209 0.149
'Amundsen sea coast'	0.475 0.032	0.569 0.186	0.571 0.191	0.623 0.156	0.572 0.192	0.452 0.124
'South Greenland'	0.230 0.143	0.256 0.216	0.258 0.220	0.303 0.264	0.259 0.221	0.200 0.158
'World's oceans'	0.099 0.040	0.080 0.044	0.079 0.042	0.081 0.044	0.079 0.042	0.077 0.039
'Globe'	0.262 0.072	0.116 0.083	0.116 0.083	0.127 0.096	0.116 0.083	0.105 0.068

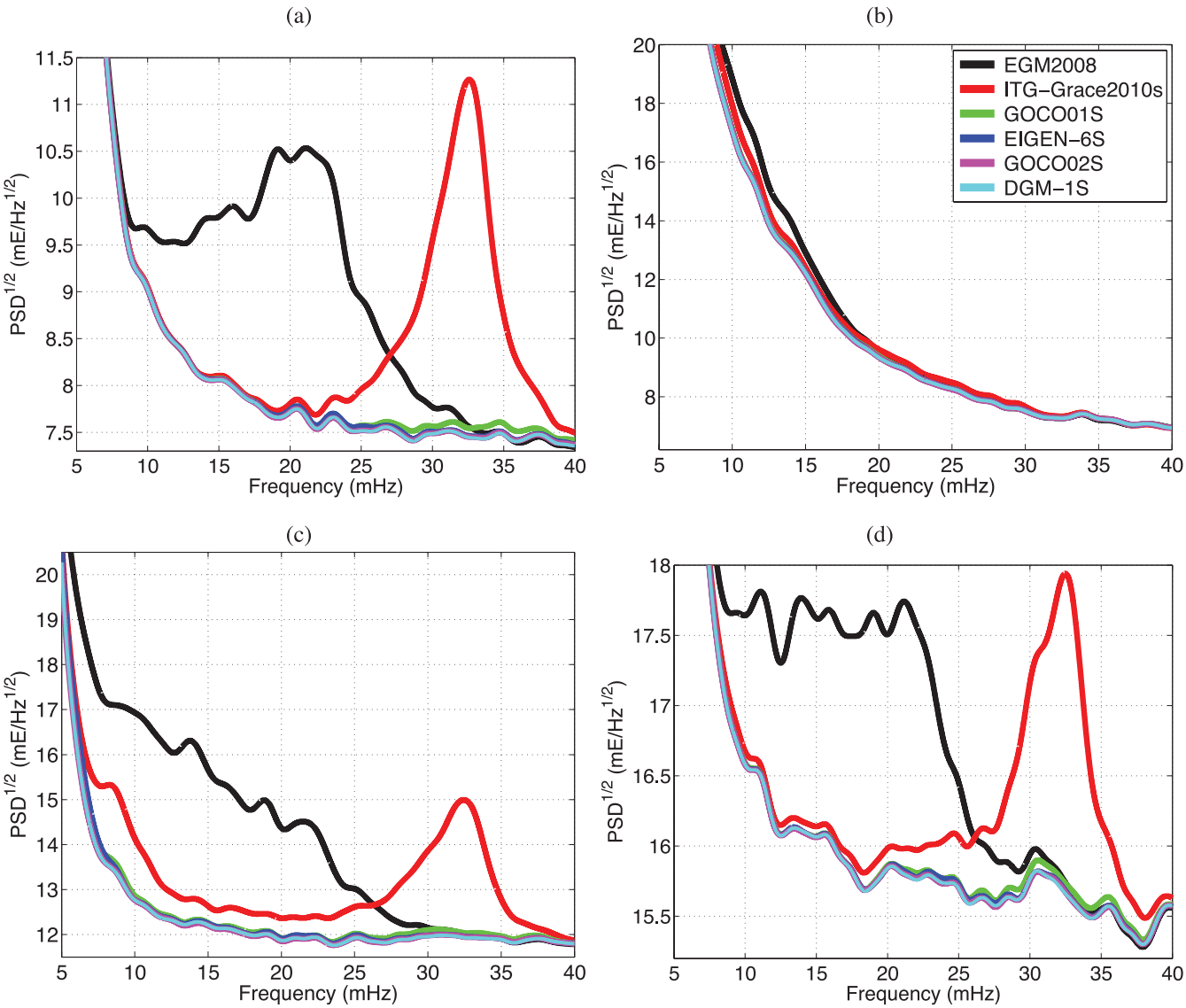

Figure 7. Square-root of PSD of (a) *xx*, (b) *yy*, (c) *zz* and (d) *xz* components of the GOCE residual gravity gradient tensor for 2011 May.

GOCO01S shows slightly larger misfits to the GOCE SGG control data at the *xx* component. The *yy* component in Fig. 7 does not reveal performance differences as pronounced as the other considered components. Besides, the PSD's at the *yy* component shows for the combined GRACE/GOCE models a much higher level than at the *xx* component. This contradicts to the fact that these two components are of an almost similar quality under normal circumstances. A possible reason for these peculiarities is discussed in the next subsection.

3.2.2 Validation in the spatial domain

To perform the validation of the models in the spatial domain, we use the sets of GOCE residual gravity gradients to compute $1^\circ \times 1^\circ$ block-mean residuals. The results for the *xx*, *yy*, *zz* and *xz* components are globally plotted in Figs 8–11, respectively. Only the maps associated with EGM2008, ITG-Grace2010s and DGM-1S are shown, because the maps for the other three models are practically indistinguishable at most locations. The 1500-km wide (in diameter) polar gaps in GOCE data are absent in these maps. The rms values are presented in Table 3 for all the four gravity gradient components and for all the six considered models. They are obtained globally as well as for the nine rectangular regions defined in Section 3.1.2 and for the world's oceans (65°S–65°N).

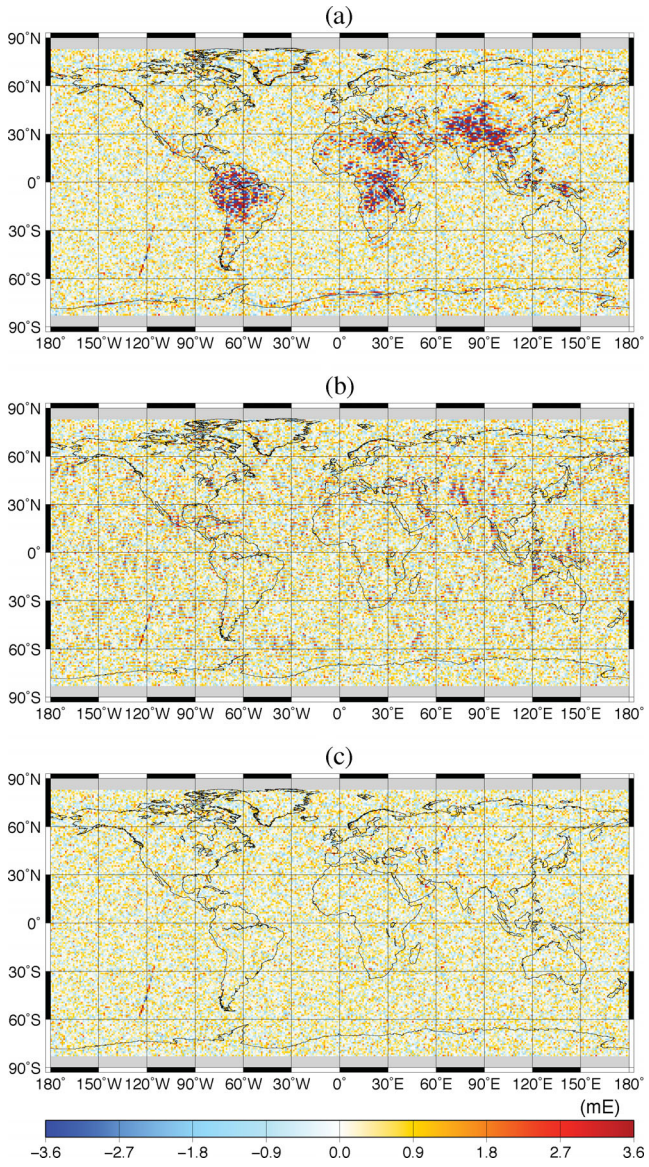

Figure 8. The GOCE $1^\circ \times 1^\circ$ block-mean residual gravity gradients for the *xx* component obtained for (a) EGM2008, (b) ITG-Grace2010s and (c) DGM-1S. The rms misfits are (a) 1.199 mE, (b) 0.944 mE and (c) 0.729 mE.

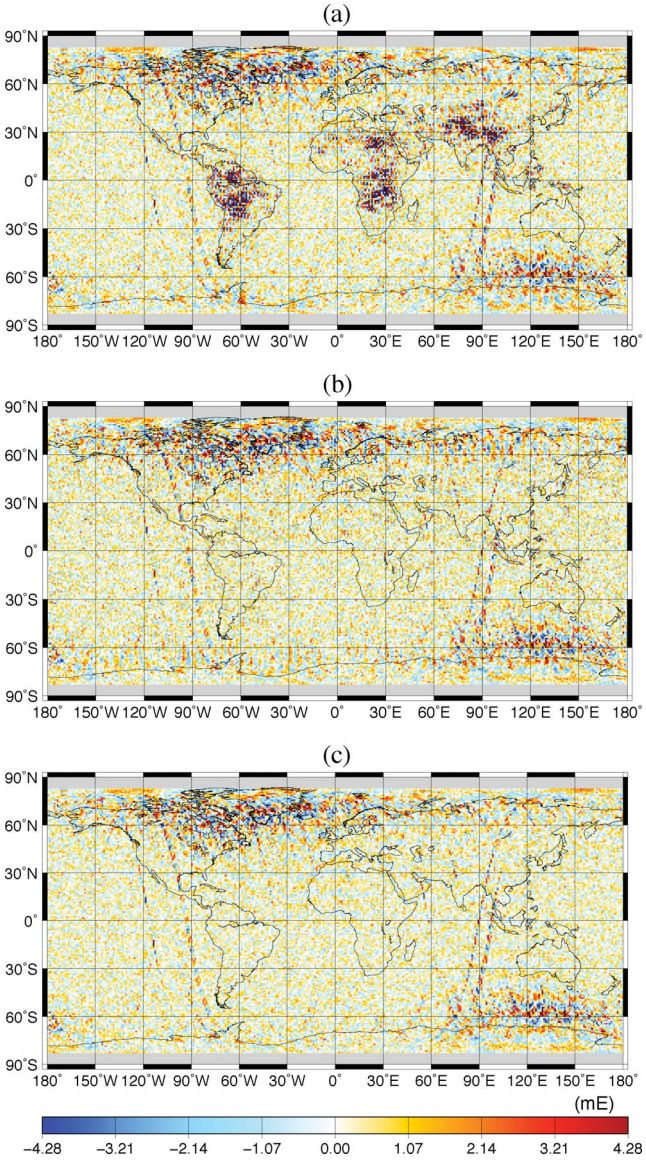

Figure 9. The GOCE $1^\circ \times 1^\circ$ block-mean residual gravity gradients for the *yy* component obtained for (a) EGM2008, (b) ITG-Grace2010s and (c) DGM-1S. The rms misfits are (a) 1.426 mE, (b) 1.226 mE and (c) 1.103 mE.

From Figs 8–10 and Table 3, one can see that the obtained results are, in general, in agreement with those based on the GRACE KBR control data. In particular ITG-Grace2010s and the combined GRACE/GOCE models, as compared to EGM2008, demonstrate in the regions from the first category a better consistency with the control data at all the four components. At the *zz* component, for instance, the rms misfits obtained for ITG-Grace2010s and the combined GRACE/GOCE models are smaller than those for EGM2008 by 69–75 and 76–83 per cent, respectively. In the regions from the second category, the combined GRACE/GOCE models also demonstrate a higher accuracy than EGM2008. The difference in performance is particularly large for the *zz* component over North

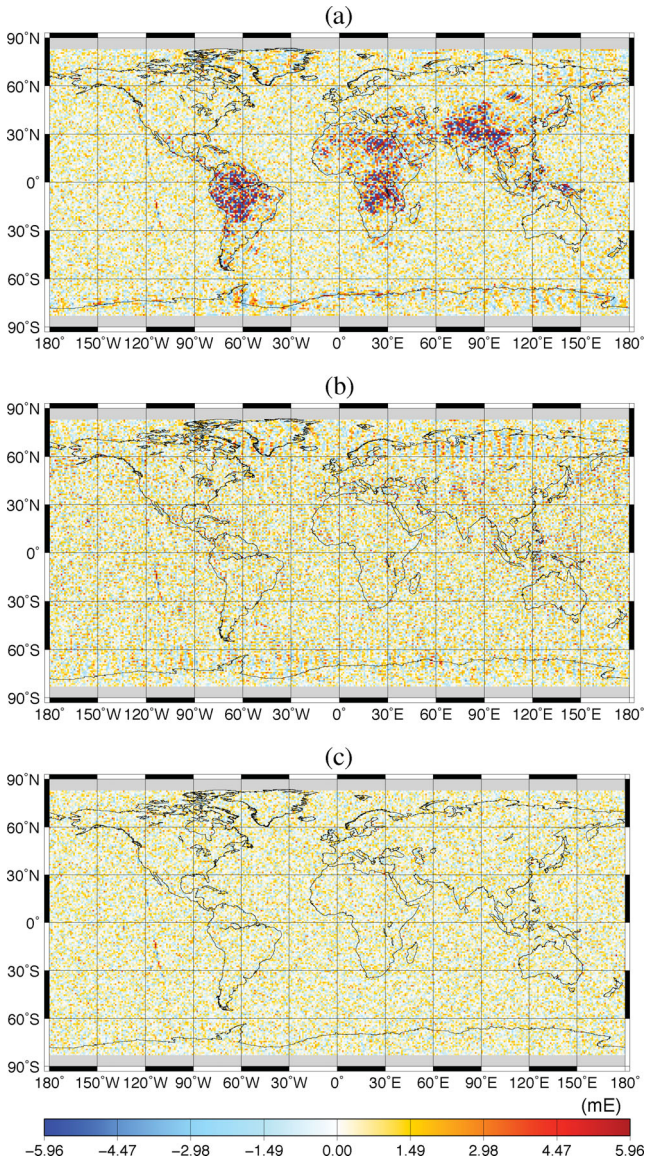

Figure 10. The GOCE 1° × 1° block-mean residual gravity gradients for the *zz* component obtained for (a) EGM2008, (b) ITG-Grace2010s and (c) DGM-1S. The rms misfits are (a) 1.981 mE, (b) 1.426 mE and (c) 1.073 mE.

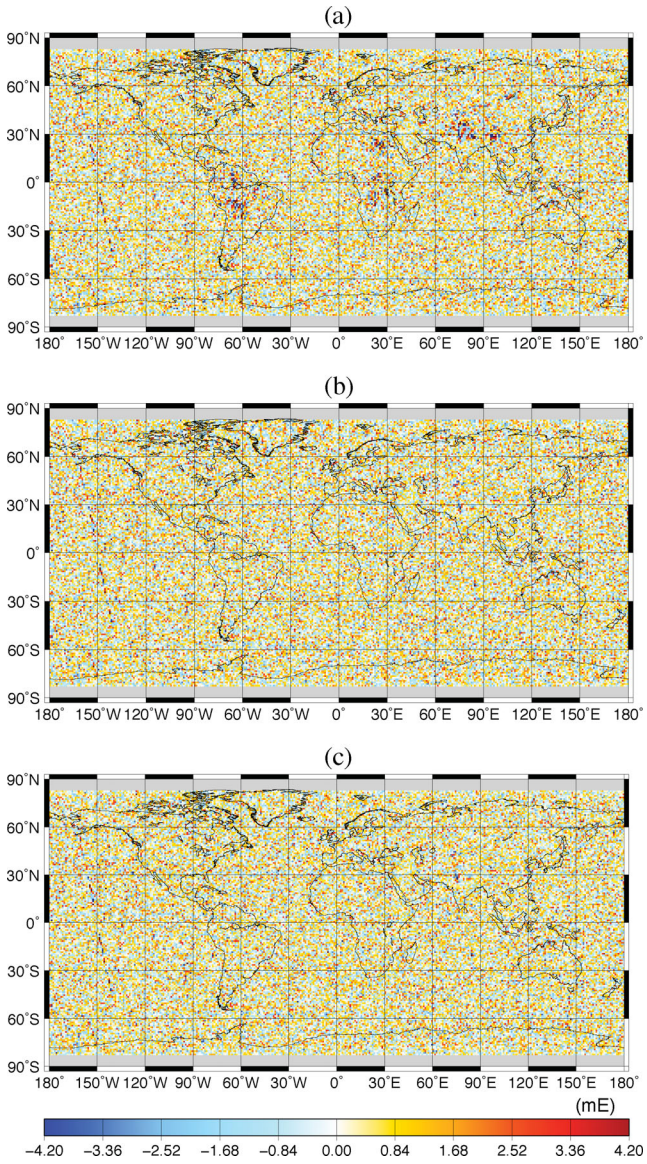

Figure 11. The GOCE 1° × 1° block-mean residual gravity gradients for the *xz* component obtained for (a) EGM2008, (b) ITG-Grace2010s and (c) DGM-1S. The rms misfits are (a) 1.405 mE, (b) 1.366 mE and (c) 1.344 mE.

Eurasia: nearly 15 per cent. For the other two regions of this category, the difference is 3–4 per cent. There are, however, a few exceptions. For instance, EIGEN-6S shows in 'Australia' larger misfits to the *yy* component of SGG data than EGM2008 (by about 4 per cent).

In the 'World's oceans', the combined GRACE/GOCE models also demonstrate a better match to the GOCE SGG control data than EGM2008. At the *xx*, *yy*, *zz* and *xz* components, the observed difference is about 7, 5, 11 and 1 per cent, respectively.

GOCO02S and DGM-1S show, in general, a better match to the GOCE SGG control data than EIGEN-6S. At the *zz* component, for instance, the difference in most regions and globally is 3–4 per cent. In 'Himalayas', the difference reaches about 17 and 15 per cent in the case of *xx* and *zz* components, respectively.

On the other hand, there are some disagreements with the results based on the GRACE KBR control data. First, ITG-Grace2010s shows a much lower accuracy than EGM2008 in the regions from the second category and in the 'World's oceans'. At the *zz* component, the difference reaches 19–35 and 18 per cent, respectively. These differences are even more pronounced at the *xx* component: 18–48 and 23 per cent, respectively. Secondly, the SGG data, unlike the KBR data, clearly show a lower performance of EGM2008 as compared to the combined GRACE/GOCE models in the regions from the third category. At the *zz* component, the difference reaches 12–25 per cent. Thirdly, the GOCE SGG control data reveal an expected better performance of DGM-1S and GOCO02S as compared to GOCO01S: by 4–5 per cent at the *zz* component. Fourthly, practically no difference can be seen in the performance of GOCO02S and DGM-1S. Fifthly, all the combined GRACE/GOCE models under consideration consistently demonstrate a higher accuracy than ITG-Grace2010s in all the considered regions. At the *zz* component, for instance, the difference is 13–36 per cent.

It is also worth noticing that the *xz* component is less sensitive to gravity field signal as compared to the diagonal components. This can be clearly observed by comparing the inconsistencies of EGM2008 associated with the diagonal components (Figs 8a, 9a and 10a) with those associated with the *xz* component (Fig. 11a)

Table 3. The rms of the GOCE 1° × 1° block-mean residual gravity gradients (in mE), per region. The top to bottom numbers in a quarter correspond to the *xx*, *yy*, *zz* and *xz* components, respectively.

Region	EGM2008	ITG-Grace2010s	GOCO01S	EIGEN-6S	GOCO02S	DGM-1S
'Himalayas'	4.018 3.475 6.625 1.963	1.307 1.124 1.783 1.387	0.846 1.027 1.187 1.371	0.975 1.060 1.332 1.365	0.817 0.993 1.133 1.367	0.811 0.998 1.138 1.369
'South America'	3.505 3.525 6.232 1.947	0.914 0.979 1.531 1.378	0.718 0.759 1.112 1.357	0.699 0.747 1.110 1.355	0.692 0.724 1.074 1.354	0.693 0.722 1.075 1.352
'Equatorial Africa'	2.531 2.556 4.551 1.688	0.931 0.913 1.431 1.361	0.741 0.772 1.129 1.351	0.729 0.778 1.130 1.350	0.727 0.749 1.100 1.350	0.722 0.749 1.099 1.350
'Australia'	0.703 0.847 1.104 1.355	1.042 1.006 1.493 1.360	0.696 0.850 1.086 1.348	0.681 0.879 1.094 1.350	0.683 0.823 1.056 1.347	0.682 0.829 1.055 1.347
'North Eurasia'	0.855 1.455 1.273 1.361	0.912 1.626 1.551 1.371	0.780 1.416 1.094 1.345	0.773 1.416 1.099 1.345	0.770 1.401 1.070 1.344	0.769 1.401 1.067 1.344
'North America'	0.766 1.567 1.171 1.392	0.902 1.645 1.397 1.400	0.747 1.570 1.136 1.389	0.753 1.579 1.147 1.389	0.744 1.560 1.106 1.387	0.745 1.563 1.112 1.388
'South Alaska'	0.821 1.330 1.322 1.338	0.876 1.473 1.499 1.362	0.759 1.252 1.197 1.339	0.760 1.233 1.188 1.338	0.754 1.247 1.147 1.336	0.750 1.264 1.140 1.330
'Amundsen sea coast'	0.943 1.022 1.369 1.330	0.839 0.964 1.200 1.285	0.770 0.915 1.051 1.272	0.764 0.901 1.041 1.267	0.765 0.907 1.033 1.266	0.758 0.900 1.045 1.271
'South Greenland'	1.004 2.858 1.452 1.392	0.930 2.976 1.585 1.403	0.779 2.815 1.110 1.346	0.766 2.816 1.098 1.337	0.769 2.808 1.075 1.338	0.763 2.798 1.064 1.337
'World's oceans'	0.778 1.054 1.207 1.370	0.954 1.118 1.419 1.378	0.736 1.017 1.110 1.365	0.726 1.026 1.106 1.365	0.721 1.004 1.076 1.363	0.720 1.005 1.075 1.363
'Globe'	1.199 1.426 1.981 1.405	0.944 1.226 1.426 1.366	0.744 1.114 1.105 1.347	0.738 1.121 1.104 1.346	0.731 1.102 1.073 1.345	0.729 1.103 1.073 1.344

in the regions from the first category, where this model is relatively inaccurate. At the *xx*, *yy* and *zz* components, the performance of this model is worse than GOCO02S by 71–80, 71–79 and 76–83 per cent, respectively, whereas at the *xz* component the difference is only 20–30 per cent. This allows us to conclude that the *xz* component is least informative in gravity field modelling. This justifies the fact that it is usually not exploited in the model production.

Furthermore, one can observe large disagreements between all the considered models and the *yy* component of the GOCE SGG control data (particularly, around the magnetic poles of the Earth, but also along some individual orbital tracks). This leads to a significant spatial variability of the rms misfits associated with this component. For instance, this component shows much larger rms misfits than the other ones in 'North Eurasia', 'South Alaska', 'South Greenland' and 'North America'. Since these discrepancies are observed for all the six models under consideration, we interpret them as an evidence of a low accuracy of the *yy* component in the control data. This finding is consistent with the results of Siemes *et al.* (2012), who explain this by a contamination of the *yy* component with the cross-track non-gravitational acceleration signals due to an imperfect gradiometer calibration.

These inaccuracies may also explain a peculiar behaviour of this component in the validation of the models in the spectral domain. To validate this hypothesis, we re-compute the PSD's related to this component with the same data as in the case of Fig. 7 but only using the latitudinal band 45°S–45°N, where the accuracy of the *yy* component is less degraded. The square-root of the resulting PSD's are exhibited in Fig. 12.

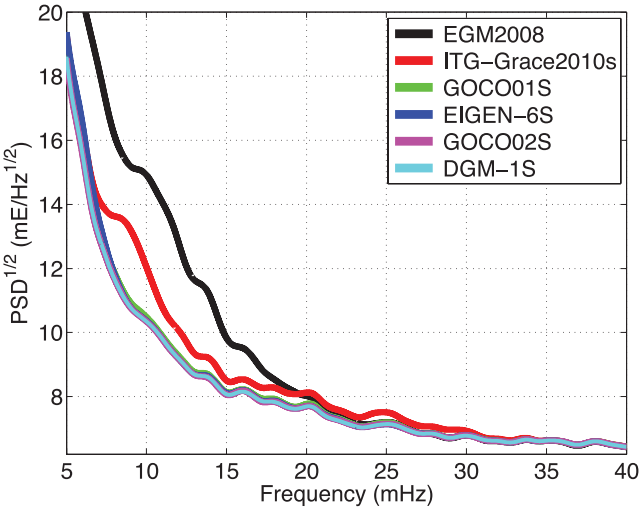

Figure 12. The same as Fig. 7(b), but using the data only in the latitudinal band 45°S–45°N.

This picture does reveal the model differences though only in the low-frequency range: 10–27 mHz. Furthermore, a comparison of the curves in Fig. 12 with those shown earlier for the *xx* component (Fig. 7a) in the context of the combined GRACE/GOCE models allows an expectedly comparable quality of these two components to be observed.

4 INSPECTING DATA COMBINATION OPTIMALITY IN MODELS COMPILED IN THE PRESENCE OF TERRESTRIAL DATA

In this section, we apply the proposed validation procedure to investigate how successfully terrestrial gravimetry and satellite altimetry measurements have been combined (i) with ITG-Grace03 in the production of EGM2008; and (ii) with GRACE/GOCE data in the production of EIGEN-6C. To that end, we analyse the performance difference between EGM2008 and ITG-Grace03 in the first subsection and that between EIGEN-6C and its satellite only counterpart, that is, EIGEN-6S, in the second subsection. We perform these two inspections in the spectral domain as well as in the spatial domain based on the $1^\circ \times 1^\circ$ block-mean values of GRACE residual range combinations related to only the 1-yr data set and on GOCE residual gravity gradients at only the diagonal components.

4.1 EGM2008 versus ITG-Grace03

Fig. 13 shows the square-root of PSD of the GRACE residual range combinations and GOCE residual gravity gradients at the *xx* component with respect to ITG-Grace03 and EGM2008. The shown frequency ranges are limited to those in which pronounced differences are observed. Figs 13(a) and (b) (related to the control KBR and SGG data, respectively) clearly identify a significant loss of information content in ITG-Grace03 in the production of EGM2008 in 4–23 mHz (22–124 cpr) and 9–26 mHz (50–140 cpr) frequency ranges, respectively. The latter picture, on the other hand, confirms a substantial gain in EGM2008 beyond the 26 mHz frequency due to the usage of terrestrial gravimetry and satellite altimetry data. We do not show here the validation spectral results related to the *yy* component due to its relatively low accuracy, as discussed in the previous section. The spectral results related to the *zz* and *xz* components are not shown either, because they lead to similar findings as those based on the *xx* component.

The block-mean values of the GRACE residual range combinations and of the GOCE residual gravity gradients related to ITG-Grace03 are plotted in Figs 14 and 15, respectively. The corresponding rms misfits obtained globally as well as for the previously defined nine rectangular regions and for the 'World's oceans' are presented in Table 4. Those obtained for EGM2008 are also included there to facilitate the comparison. For the sake of completeness, the rms misfits related to the 3-month subset of the GRACE KBR control data are also included.

From a comparison of Figs 14 and 15 with their counterparts associated with EGM2008 (Figs 3a and 8a–10a, respectively), one can clearly see that the data combination performed in the production of EGM2008 has led to a significant deterioration of the model's performance in the regions of the first category. The rms misfits confirm that EGM2008 is of much lower accuracy than ITG-Grace03 in those regions (by 79–85 per cent in the case of the GRACE KBR control data and by 59–69 per cent in the case of the GOCE SGG control data). Furthermore, the GRACE-based rms misfits suggest that EGM2008 also performs slightly poorer than ITG-Grace03 in the other considered regions. For instance, the difference in the regions of the second category is 6–8 per cent and in the 'World's oceans' is about 11 per cent up to maximum spectral ability of KBR control data. The conducted analysis confirms that the data combination in the production of EGM2008 has suffered from a partial loss of the information content of ITG-Grace03 in the areas with a poor coverage with terrestrial gravimetry data. Given

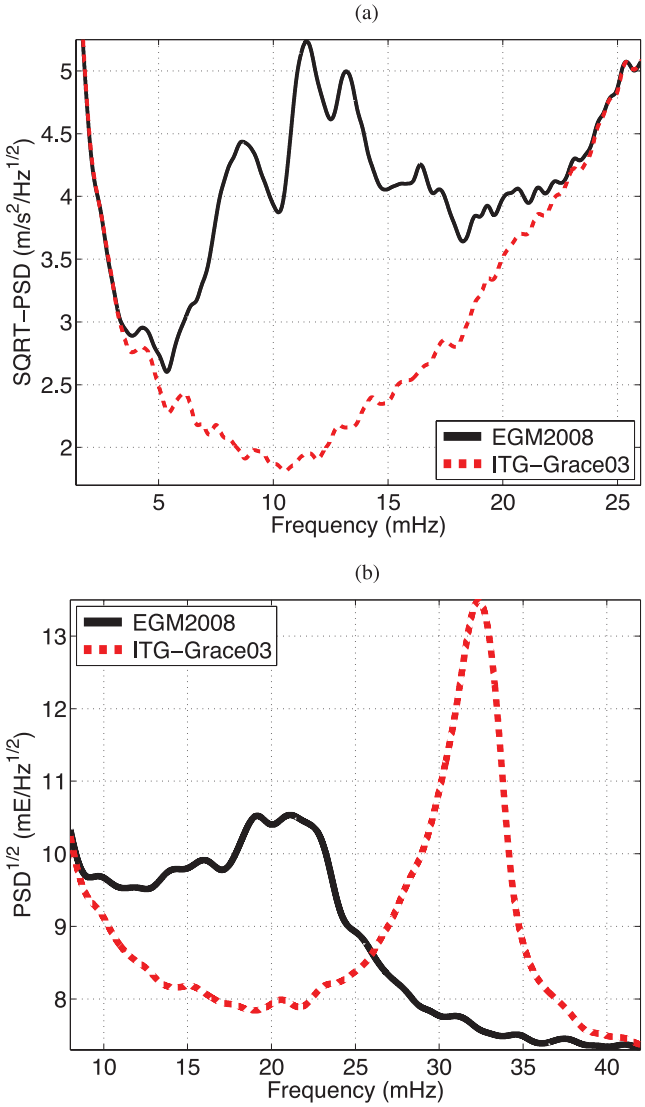

Figure 13. Square-root of PSD of the (a) GRACE residual range combinations and (b) GOCE *xx* residual gravity gradients for 2010 February and 2010 May, respectively.

the results related to KBR control data, a minor loss of information of ITG-Grace03 may have occurred also in the well-studied areas (in the frequency range to which these data are sensitive). On the other hand, the GOCE SGG control data clearly demonstrate that EGM2008 performs much better in the gravimetrically well-studied continental regions as well as in the oceanic areas. At the *zz* component, for instance, the difference in performance is 39–41 per cent and about 35 per cent, respectively.

Furthermore, it is worth noticing that a comparison of Fig. 14 with Figs 15(b) and (c) shows that GRACE-based test fails to reveal the meridional inaccuracies, so-called stripes, in ITG-Grace03, whereas the results based on the *yy* and *zz* gravity gradient components reveal them clearly. We explain the inability of the KBR control data to reveal these errors by the anisotropic sensitivity of these data, which is (as it has been already mentioned in Section 1) a serious limitation of these data.

Finally, it is worth mentioning that the global rms misfit obtained for ITG-Grace03 on the basis of the KBR control data (i.e. 0.380 μGal) is about 0.5 per cent smaller than that for ITG-Grace2010s. The reason for that is not understood.

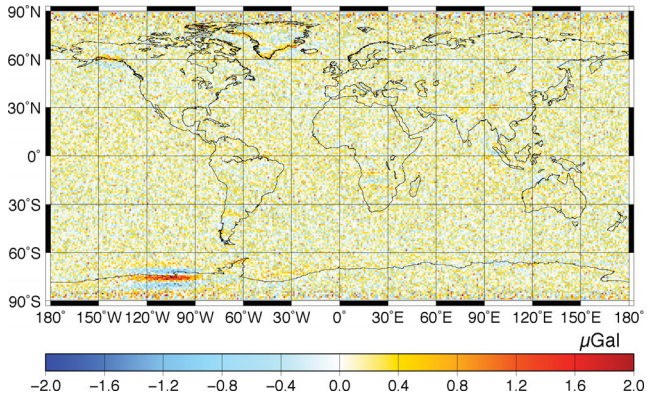

Figure 14. The GRACE 1° × 1° block-mean residual range combinations obtained for ITG-Grace03 on the basis of the yearly data set. The rms misfit is 0.380 μGal.

4.2 EIGEN-6C versus EIGEN-6S

Fig. 16 shows the square-root of PSD of the GRACE residual range combinations with respect to EIGEN-6C and its satellite-only counterpart. The picture in the zoomed frequency range (4–11 mHz or 22–60 cpr) only reveals minor performance differences, but importantly in favour of the satellite-only model. The implication is that a minor loss of information content of satellite data in the production of EIGEN-6C may have also occurred. The PSD's related to the GOCE SGG control data are not shown here, because they fail to reveal these differences due to their weak sensitivity in such low-frequency range.

The global maps of 1° × 1° block-mean values of GRACE residual range combinations and of GOCE residual gravity gradients related to EIGEN-6C are not shown, as they turn out to be visually indistinguishable from those previously shown for EIGEN-6S. This by itself indicates that EIGEN-6C has not suffered from a loss of information content in satellite data as severely as EGM2008. To quantify a potential information loss, we present the rms misfits related to EIGEN-6S and EIGEN-6C in Table 5 globally as well as in the previously defined regions. The results suggest that the combination of GRACE and GOCE data with the terrestrial gravimetry measurements in the production of EIGEN-6C has also led to a minor performance degradation as compared to EIGEN-6S in the first category regions. The maximum loss of information content of satellite data is observed in 'South America': by about 6 per cent at the *yy* component, by about 11 per cent at the *xx* and *zz* components and by about 1 per cent in the case of the GRACE KBR control data. A similar problem is also identified in 'Equatorial Africa' but in a less pronounced manner: by only 1–3 per cent in terms of rms misfits to the GOCE SGG control data. In the well-studied continental and oceanic areas, EIGEN-6C expectedly outperforms its satellite-only counterpart: the difference reaches 2–4 per cent at the *zz* component of gravity gradients. This level of performance difference is rather small. The implication is that our test may have not been sufficiently sensitive to the contribution of surface gravimetry data in a model produced in the presence of GOCE SGG data.

5 DISCUSSION AND CONCLUSIONS

With our study, we showed that both GRACE KBR and GOCE SGG data can be used as control data for a validation and comparison of global static gravity field models. To facilitate the model assessment on a region-by-region basis, we recommend to preprocess

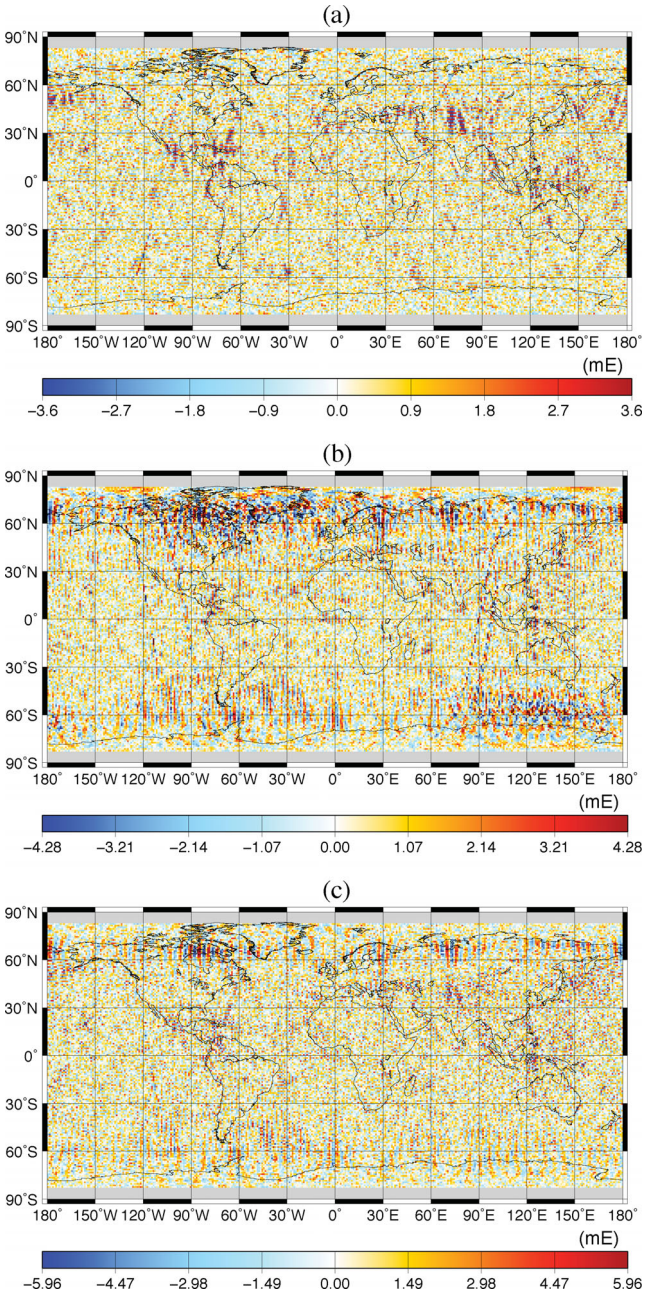

Figure 15. The GOCE 1° × 1° block-mean residual gravity gradients obtained for ITG-Grace03 at the (a) *xx*, (b) *yy* and (c) *zz* component. The rms misfits are (a) 1.076 mE, (b) 1.527 mE and (c) 1.862 mE.

KBR data in such a way that they can be related to gravity field spatial variations locally: for example, to transform them into range combinations with eq. (1).

The validation can be performed in both the frequency and spatial domain. In the latter case, it is advised to represent residual data in the form of block-mean values, so that random noise in the data is suppressed. An advantage of the validation in the spatial domain is that the model performance can be assessed on a regional basis. However, an analysis in the spectral domain can be informative, too. First, it can provide some information about the model accuracy at different spatial scales. Secondly, it can clearly show which parts of the spectrum are overwhelmed by noise and,

Table 4. The rms of the GRACE 1° × 1° block-mean residual range combinations (in μGal) and the GOCE 1° × 1° block-mean residual gravity gradients (in mE) with respect to EGM2008 and ITG-Grace03. The GRACE-based rms misfits are presented for both the 1-yr data set (in the numerator) and for the 3-month subset (in the denominator). The top, middle and bottom rms numbers in a triad correspond to the *xx*, *yy* and *zz* components of GOCE SGG data, respectively.

Region	GRACE-based misfits		GOCE-based misfits	
	ITG-Grace03	EGM2008	ITG-Grace03	EGM2008
'Himalayas'	0.358 0.599	2.323 2.452	1.620 1.535 2.433	4.018 3.475 6.625
'South America'	0.352 0.557	2.211 2.335	1.142 1.290 1.922	3.505 3.525 6.232
'Equatorial Africa'	0.353 0.562	1.656 1.768	1.028 1.135 1.740	2.531 2.556 4.551
'Australia'	0.357 0.614	0.386 0.626	1.091 1.324 1.824	0.703 0.847 1.104
'North Eurasia'	0.354 0.569	0.385 0.588	0.998 1.964 2.067	0.855 1.455 1.273
'North America'	0.345 0.548	0.368 0.567	0.984 2.067 1.988	0.766 1.567 1.171
'South Alaska'	0.440 0.633	0.458 0.657	0.982 1.603 1.881	0.821 1.330 1.322
'Amundsen sea coast'	0.686 0.860	0.701 0.872	0.900 1.085 1.405	0.943 1.022 1.369
'South Greenland'	0.378 0.620	0.444 0.659	0.951 3.157 2.092	1.004 2.858 1.452
'World's oceans'	0.344 0.567	0.386 0.595	1.097 1.449 1.862	0.778 1.054 1.207
'Globe'	0.380 0.627	0.667 0.846	1.076 1.527 1.862	1.199 1.426 1.981

therefore, should be filtered out prior to the validation in the spatial domain.

In this study, we performed a validation and comparison of eight models: two models based on a combination of satellite and surface data: (EGM2008 and EIGEN-6C, both considered up to degree 250 only), two GRACE-only models (ITG-Grace03 and ITG-Grace2010s) and four satellite-only combined GRACE/GOCE models (GOCO01S, EIGEN-6S, GOCO02S and DGM-1S). Table 6 shows a summary of the validation results in the spatial domain. This table presents rms misfits of the 1° × 1° block-mean residuals glob-

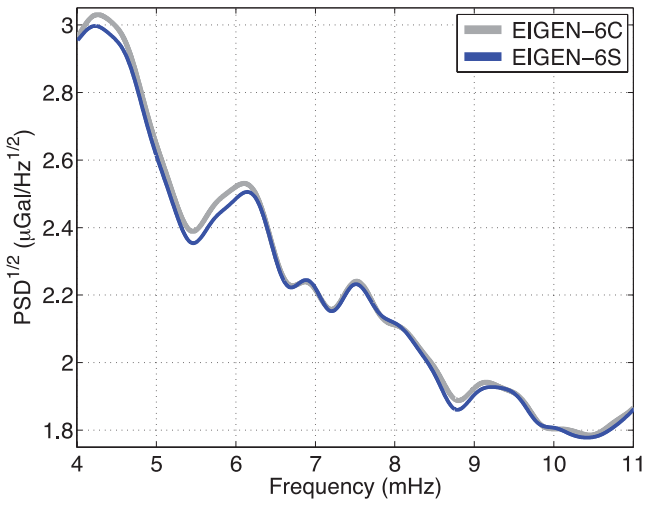

Figure 16. Square-root of PSD of the GRACE residual range combinations for 2010 February.

Table 5. The same as Table 4, but for EIGEN-6C and EIGEN-6S.

Region	GRACE-based misfits		GOCE-based misfits	
	EIGEN-6S	EIGEN-6C	EIGEN-6S	EIGEN-6C
'Himalayas'	0.370 0.604	0.371 0.603	0.975 1.060 1.332	0.967 1.065 1.342
'South America'	0.356 0.562	0.360 0.564	0.699 0.747 1.110	0.744 0.837 1.241
'Equatorial Africa'	0.357 0.568	0.356 0.567	0.729 0.778 1.130	0.736 0.804 1.144
'Australia'	0.357 0.609	0.359 0.612	0.681 0.879 1.094	0.673 0.828 1.048
'North Eurasia'	0.366 0.576	0.366 0.577	0.773 1.416 1.099	0.771 1.406 1.082
'North America'	0.358 0.561	0.359 0.558	0.753 1.579 1.147	0.745 1.568 1.107
'South Alaska'	0.458 0.644	0.460 0.648	0.760 1.233 1.188	0.752 1.248 1.165
'Amundsen sea coast'	0.791 0.959	0.790 0.958	0.764 0.901 1.041	0.768 0.903 1.044
'South Greenland'	0.431 0.661	0.433 0.665	0.766 2.816 1.098	0.764 2.814 1.071
'World's oceans'	0.344 0.568	0.344 0.568	0.726 1.026 1.106	0.718 1.008 1.075
'Globe'	0.387 0.632	0.387 0.633	0.738 1.121 1.104	0.734 1.111 1.085

ally as well as for the pre-defined nine rectangular regions and for the 'World's oceans' in terms of percentages, defining that RMS misfit obtained for a given region for EGM2008 is 100 per cent. Since the RMS misfits for EGM2008 in most cases is the largest among all the models, the chosen presentation manner in Table 6 helps us to highlight the difference between the observed performance of EGM2008 on the one hand and of the other models on the other hand. The results based on both the KBR and SGG control data are included. The GRACE-based numbers correspond to the 3-month subset, so that their comparison with the GOCE-based numbers is more equitable. As far as the GOCE-based values are concerned, the minimum of the numbers obtained for the considered components are shown. In other words, we consider the best-case validation scenario among those based on the GOCE SGG control data. It is worth noting that these smallest numbers are mostly related to the *zz* component.

Table 6 demonstrates, in particular, a significantly poorer performance of EGM2008 in the regions of the first category as compared to the other models, including ITG-Grace2010s. This implies that the added value of the GOCE mission to determining the gravity field in continental areas void of high-quality terrestrial measurements can be quantified more fairly by comparing the performance of the combined (satellite-only) GRACE/GOCE models with ITG-Grace2010s rather than with EGM2008. In view of that, the validation of the models against the KBR control data does not allow this added value to be seen. The validation of the models against the SGG control data, however, quantifies the added value of the GOCE mission in these regions as an improvement of models quality by 21–36 per cent.

At the same time, validation against the SGG control data indicates a higher accuracy of EGM2008 than ITG-Grace2010s in the regions of the second category as well as in the 'World's oceans'. This implies that it is fairer to quantify the added value of GOCE data in those areas by comparing the performance of the combined satellite-only GRACE/GOCE models with that of EGM2008. In the context of the validation against the SGG control data, this turns

Table 6. The rms of the GRACE 1° × 1° block-mean residual range combinations (numerator) and the GOCE 1° × 1° block-mean residual gravity gradients (denominator) obtained for the assessed eight static gravity field models. The rms misfits are shown in terms of percentage with respect to the rms value obtained for EGM2008 in a given region. The GRACE-based values are computed from the 3-month data subset. The GOCE-based values are related to the minimum of those obtained for the considered gravity gradient components. In most cases, these minimum numbers are related to the *zz* component.

Region	EGM2008	ITG-Grace03	ITG-Grace2010s	GOCO01S	EIGEN-6S	EIGEN-6C	GOCO02S	DGM-1S
'Himalayas'	100.0 % 100.0 %	24.4 % 36.7 %	24.4 % 26.9 %	24.3 % 17.9 %	24.6 % 20.1 %	24.6 % 20.3 %	24.2 % 17.1 %	24.1 % 17.2 %
'South America'	100.0 % 100.0 %	23.9 % 30.8 %	23.9 % 24.6 %	23.9 % 17.8 %	24.1 % 17.8 %	24.2 % 19.9 %	23.9 % 17.2 %	23.9 % 17.2 %
'Equatorial Africa'	100.0 % 100.0 %	31.8 % 38.2 %	31.8 % 31.4 %	31.8 % 24.8 %	32.1 % 24.8 %	32.1 % 25.1 %	31.8 % 24.2 %	31.8 % 24.1 %
'Australia'	100.0 % 100.0 %	98.1 % 155.2 %	98.1 % 118.8 %	97.8 % 98.4 %	97.3 % 96.9 %	97.8 % 94.9 %	97.0 % 95.7 %	97.1 % 95.6 %
'North Eurasia'	100.0 % 100.0 %	96.8 % 116.7 %	96.8 % 106.7 %	96.6 % 85.9 %	98.0 % 86.3 %	98.1 % 85.0 %	96.6 % 84.1 %	96.6 % 83.8 %
'North America'	100.0 % 100.0 %	96.6 % 128.5 %	97.5 % 105.0 %	98.6 % 97.0 %	98.9 % 98.0 %	98.4 % 94.5 %	97.5 % 94.4 %	97.5 % 95.0 %
'South Alaska'	100.0 % 100.0 %	96.3 % 119.6 %	97.6 % 106.7 %	97.4 % 90.5 %	98.0 % 89.9 %	98.6 % 88.1 %	97.4 % 86.8 %	92.8 % 86.2 %
'Amundsen sea coast'	100.0 % 100.0 %	98.6 % 95.4 %	104.7 % 87.7 %	104.9 % 76.8 %	110.0 % 76.0 %	109.9 % 76.3 %	105.2 % 75.5 %	94.4 % 76.3 %
'South Greenland'	100.0 % 100.0 %	94.1 % 94.7 %	96.8 % 92.6 %	97.1 % 76.4 %	100.3 % 75.6 %	100.9 % 73.8 %	97.1 % 74.0 %	93.2 % 73.3 %
'World's oceans'	100.0 % 100.0 %	95.3 % 137.5 %	95.5 % 106.1 %	95.3 % 92.0 %	95.5 % 91.6 %	95.5 % 89.1 %	95.3 % 89.1 %	95.3 % 89.1 %
'Globe'	100.0 % 100.0 %	74.1 % 89.7 %	74.5 % 72.0 %	74.3 % 55.8 %	74.7 % 55.7 %	74.8 % 54.8 %	74.3 % 54.2 %	74.0 % 54.2 %

out to be 4–16 per cent in the second category regions and about 11 per cent in the 'World's oceans'.

Furthermore, it was shown that the combination of ITG-Grace03 and terrestrial gravimetry measurements in the production of EGM2008 has significantly worsened the quality of the resulting model as compared to ITG-Grace03 in the gravimetrically poorly studied continental areas. According to the KBR and SGG control data, this problem manifests itself in 4–23 mHz (22–124 cpr) and 9–26 mHz (50–140 cpr) frequency ranges, respectively. We relate this to a consequence of an imperfection in the data weighting performed in the data combination (e.g. due to a lacking stochastic description of noise in some input data). It is worth emphasizing that a loss of information content of ITG-Grace03 in the production of EGM2008 has been reported by Pavlis *et al.* (2012) themselves. The performance differences identified between these two models also explain a large part of those between ITG-Grace2010s and EGM2008 in the poorly studied continental areas.

In addition, we compared the performance of EIGEN-6C with that of its satellite-only counterpart, that is, EIGEN-6S, to inspect if the former model has also experienced a loss of information content of satellite data in the course of data combination. The inspection did reveal this problem in the poorly gravimetrically surveyed areas, but in a significantly less pronounced manner than in the case of EGM2008. In the well-studied continental and oceanic areas, an expected better performance of the EIGEN-6C model was confirmed. We relate the absence of significant information losses of satellite data in the production of EIGEN-6C to the adopted data combination scheme. According to Förste *et al.* (2011), the surface gravity data were permitted to contribute to the production of this model only above degree 160, where the contribution of KBR and SGG data is not substantial.

The results based on the 1-yr KBR control data set and to some extent on the 3-month subset (summarized in Table 6) indicate that DGM-1S shows slightly smaller misfits than the other combined GRACE/GOCE models. One may argue that this is because the same functional model based on range combinations is exploited to handle GRACE KBR data in both the model production and the model validation. This is not the case for the other considered models, all of which use in one way or another KBR data in the form of range-rates. Therefore, we additionally performed the validation of the models using GRACE range-rates. In doing so, we computed the sets of GRACE 1° × 1° block-mean residual range-rates using the 1-yr data set. The corresponding rms misfits obtained globally read as 59 nm s^{-1}, 55 nm s^{-1}, 49.4 nm s^{-1}, 49.8 nm s^{-1}, 49.4 nm s^{-1} and 49.3 nm s^{-1} for EGM2008, ITG-Grace2010s, GOCO01S, EIGEN-6S, GOCO02S and DGM-1S, respectively. Thus, DGM-1S still shows a smaller or similar rms misfit as compared to the other models.

The comparison of the validation results based on the KBR and SGG control data in the spectral domain (respectively presented in Figs 2 and 7) clearly shows a difference in sensitivity of these two data types to model inaccuracies. For instance, the KBR data show inconsistencies with EGM2008 in the frequency range 5–22 mHz with the maximum at about 12 mHz ≈ 65 cpr, which corresponds to degree 65 and less. On the other hand, the SGG data demonstrate inconsistencies with this model at higher degrees. For instance, the *xx* component shows inconsistencies in the frequency range 10–32 mHz with the maximum at 22.5 mHz ≈ 120 cpr. We also note the frequency range 25–37 mHz, in which the SGG data (particularly, at the *xx* component), show large inconsistencies with ITG-Grace2010s, whereas the KBR test data do not allow these inconsistencies to be seen due to an increased noise level at high frequencies. In other words, SGG data expectedly show the maximum sensitivity to model inaccuracies at higher degrees as compared to KBR data.

The results obtained by the validation in the spatial domain, which are summarized in Table 6, offers another way to compare the sensitivity of KBR and SGG data to model inaccuracies. One can see that the SGG data, as compared to the KBR data, reveal a larger discrepancy between the performances of the models. For instance, in the regions from the first and the second categories, the rms misfits associated with the combined GRACE/GOCE models are respectively equal to 17–24 and 80–96 per cent of those related to EGM2008 when the GOCE SGG data are considered, versus

24–32 per cent and 94–98 per cent when the GRACE KBR data are utilized. This indicates to an expectedly higher sensitivity of SGG data. However, this conclusion is based on the mean values over $1^\circ \times 1^\circ$ blocks. The obtained values in this case may capture model inaccuracies up to degree 180 and even somewhat higher. Since model inaccuracies increase with degree, it is not surprising that SGG data act as a more sensitive tool in this situation.

A different frequency sensitivity of KBR and SGG data also explains why the former data practically do not allow discrepancies between GOCO01S and GOCO02S to be seen, whereas the latter data do reveal a difference of up to 4 per cent in terms of rms misfits (see Table 6). As it is explained at the beginning of Section 3, the major additional source of information used in GOCO02S was six extra months of GOCE SGG data. Clearly, the added value of these data manifest itself mostly at high degrees (see Fig. 1). The control data that are sensitive to high-frequency signals offer the best way to sense this improvement.

A better ability of SGG data to reveal model inaccuracies in the range of high degrees may also explain why these data indicate to a much higher accuracy of EGM2008 than ITG-Grace03 or ITG-Grace2010s in the regions of the second category and in the 'World's oceans', whereas the KBR control data show a slightly worse quality of the former model. Furthermore, a different spectral sensitivity of these two data types may also explain a contradiction in the validation outcome in the third category regions, where DGM-1S shows a significantly higher consistency with the KBR control data as compared to, for example, GOCO02S, whereas an almost comparable performance of these two models is observed there in the case of the SGG control data (see Table 6). This contradiction is in line with our hypothesis of an incomplete removal of long-term mass transport signal from the KBR control data in the validation procedure. Since temporal variations of the gravity field occur mostly at low degrees (Wahr *et al.* 1998), they have no effect when a validation of the models is performed against the SGG data.

Furthermore, the analysis performed in the spatial domain demonstrates another limitation of KBR control data: their anisotropic sensitivity. These data fail to reveal the meridional inaccuracies, so-called 'stripes', in ITG-Grace03 (Fig. 14). On the other hand, the test performed based on the SGG control data (Fig. 15) clearly revels those errors.

The presented study proves that GOCE SGG and GRACE KBR data can complement each other in different ways. The GOCE gradiometer is a better tool to assess the integrated quality of a model up to a relatively high degree, that is, 180 or even higher. Importantly, the collection of the GOCE gradiometry data at various components is sensitive to gravity field variations in all directions, so that the validation results obtained on their basis can be considered as sufficiently comprehensive. Of course, the GOCE-based test in the spatial domain requires that the residual gravity gradients are averaged over sufficiently small blocks, for example, of the size $1^\circ \times 1^\circ$, so that a high sensitivity of SGG data at high frequencies can be exploited. At the same time, SGG data are less sensitive to possible model inaccuracies at lower degrees. Thus, a validation based on KBR data must be preferred when the model performance at low degrees is to be investigated, particularly, below degree 27. For this purpose, it is sufficient to increase the size of the block in computing the block-mean values. Of course, one should keep in mind the anisotropic sensitivity of GRACE KBR data, which limits their ability to assess the east–west gravity field variations. Another limitation of KBR data is an increased noise level at very low frequencies (*cf.* Fig. 2), which can make these data poorly sensitive to model inaccuracies at very low degrees, that is, 2–4. Therefore, control data of other types, for example, SLR data, may be needed when the accuracy of a model at very low degrees is the primary concern.

An important outcome of the conducted study is the demonstrated ability of satellite data to sense the difference in the performance of static gravity field models on a region-by-region basis, even if the length of the independent control data is much shorter than the length of data sets used in the production of models. In the considered examples, for instance, all the models were based on several years of GRACE KBR data. Nevertheless, even a 3-month set of independent KBR data allowed a difference in the performance of EGM2008 and the combined GRACE/GOCE models to be seen. The 1-yr set of these data made this difference even more pronounced, as it further improved signal-to-noise ratio in the predefined blocks.

None of the static gravity field models considered in our manuscript uses the *xz* component of GOCE SGG data. Thus, this component is potentially the source of fully independent information for validation of models. This encouraged us to include the *xz* component into the sets of control data. Nevertheless, this component was found to be least informative among all the considered gravity gradient components.

It is worth noting that the model performance shown in the case of EGM2008 is in a good agreement (in terms of geographical distribution of its errors) with the propagated error estimates of this model (Pavlis *et al.* 2012). It is a matter of future studies to inspect if this is also the case for the other tested models. To that end, it is sufficient to compare the actual errors of these models, a quantification of which is presented in this manuscript, with their propagated error estimates.

Importantly, both KBR and SGG data may suffer from systematic inaccuracies. Examples are inaccuracies related to calibration of GRACE attitude determination system (Horwath *et al.* 2011) and inaccuracies caused by an imperfect calibration of GOCE SGG data (Siemes *et al.* 2012). Such inaccuracies contaminate both the data used for the production of models and the control data. Therefore, the proposed validation scheme may not reveal the impact of those inaccuracies, even if the control data are independent in line with our data independency definition. In the context of our study, this caveat does not apply to the validation of three of the tested models, namely, ITG-Grace03, EGM2008 and ITG-Grace2010s, with the SGG control data since these models were produced without GOCE SGG data.

We conclude by saying that a validation based on independent GRACE KBR and GOCE SGG data is a sufficiently general tool in the sense that it can be applied to any global static model of the Earth's gravity field, including the old ones. However, we find it important to emphasize that the proposed validation procedure represents only one way to test the model accuracy. A comprehensive evaluation of a model requires its validation against all types of independent control data, including the terrestrial ones.

ACKNOWLEDGEMENTS

The Stichting Nationale Computerfaciliteiten (National Computing Facilities Foundation, NCF) provided the high-performance computing facilities. The work was financially sponsored by the Nederlandse organisatie voor Wetenschappelijk Onderzoek (Netherlands Organisation for Scientific Research, NWO). The GRACE KBR measurements, and GOCE SGG data and orbits were provided by JPL and ESA, respectively. We thank the editor, Dr. Gary Egbert,

and two anonymous reviewers for their constructive remarks that helped us to improve the quality of the manuscript.

REFERENCES

Bettadpur, S., 2007. CSR Level-2 processing standards document for product release 04, GRACE 327-742, Revision 3.1.

Bruinsma, S.L., Marty, J.C., Balmino, G., Biancale, R., Förste, C., Abrikosov, O. & Neumayer, H., 2010. GOCE gravity field recovery by means of the direct numerical method, in *Proceedings of the ESA Living Planet Symposium,* 28 Jun–2 Jul, 2010, Bergen, Norway.

Case, K., Kruizinga, G.L.H. & Wu, S.-C., 2004. GRACE Level-1B data product user handbook, Jet Propulsion Laboratory, California Institute of Technology, JPL D-22027.

de Witte, S., 2011. GOCE XML Parser, European Space Agency, GO-TN-HPF-GS-0192, Issue: 2, Revision: 7.

Desai, S.D., 2002. Observing the pole tide with satellite altimetry, *J. geophys. Res.,* **107**(C11), 3186.

Ditmar, P. & van Eck van der Sluijs, A.A., 2004. A technique for modeling the Earth's gravity field on the basis of satellite accelerations, *J. Geod.,* **78,** 12–33.

Ditmar, P., Kuznetsov, V., van Eck van der Sluijs, A.A., Schrama, E. & Klees, R., 2006. DEOS_CHAMP-01C-70: a model of the Earth's gravity field computed from accelerations of the CHAMP satellite, *J. Geod.,* **79,** 586–601.

Ditmar, P., Klees, R. & Liu, X., 2007. Frequency-dependent data weighting in global gravity field modeling from satellite data contaminated by non-stationary noise, *J. Geod.,* **81,** 81–96.

Ditmar, P., Teixeira da Encarnação, J. & Hashemi Farahani, H., 2012. Understanding data noise in gravity field recovery on the basis of inter-satellite ranging measurements acquired by the satellite gravimetry mission GRACE, *J. Geod.,* **86,** 441–465.

Drinkwater, M.R., Floberghagen, R., Haagmans, R., Muzi, D. & Popescu, A., 2003. GOCE: ESA's first Earth explorer core mission, *Space Sci. Rev.,* **00,** 1–14.

Flechtner, F., 2007. AOD1B product description document for Product Releases 01 to 04. GRACE 327-750, GR-GFZ-AOD-0001.

Förste, C. *et al.*, 2008a. The GeoForschungsZentrum Potsdam/Groupe de Rcherche de Gèodésie Spatiale satellite only and combined gravity field models: EIGEN-GL04S1 and EIGEN-GL04C, *J. Geod.,* **82,** 331–346.

Förste, C. *et al.*, 2008b. EIGEN-GL05C - A new global combined high-resolution GRACE-based gravity field model of the GFZ-GRGS cooperation, *Geophys. Res. Abstr.,* **10,** EGU2008-A-03426.

Förste, C. *et al.*, 2011. EIGEN-6 - A new combined global gravity field model including GOCE data from the collaboration of GFZ-Potsdam and GRGS-Toulouse, *Geophys. Res. Abstr.,* **13,** EGU2011-3242-2.

Floberghagen, R., Fehringer, M., Lamarre, D., Muzi, D., Frommknecht, B., Steiger, C., Piñeiro, J. & da Costa, A., 2011. Mission design, operation and exploitation of the gravity field and steady-state ocean circulation explorer mission, *J. Geod.,* **85,** 749–758.

Goiginger, H. *et al.*, 2011. The combined satellite-only global gravity field model GOCO02S, *Geophys. Res. Abstr.*, **13,** EGU2011-10571.

Gruber, T., 2009. Evaluation of the EGM2008 gravity field by means of GPS-levelling and sea surface topography solutions, *Newtons Bull. Bur. Gravimétrique Int. (BGI)/Int. Geoid Service (IGeS),* **4,** 3–17.

Gruber, T., Rummel, R., Abrikosov, O. & van Hees, R., 2010. GOCE Level 2 product data handbook. European Space Agency, GO-MA-HPF-GS-0110, Issue: 4, Revision: 3.

Gruber, T., Visser, P., Ackermannm, Ch. & Hosse, M., 2011. Validation of GOCE gravity field models by means of orbit residuals and geoid comparisons, *J. Geod.,* **85,** 845–860.

Hashemi Farahani, H., Ditmar, P., Klees, R., Liu, X., Zhao, Q. & Guo, J., 2013. The static gravity field model DGM-1S based on GRACE and GOCE data: computation, validation, and an analysis of GOCE missions added value, *J. Geod.,* in press.

Hirt, C., Gruber, T. & Featherstone, W.E., 2011. Evaluation of the first GOCE static gravity field models using terrestrial gravity, vertical deflections and EGM2008 quasigeoid heights, *J. Geod.,* **85,** 723–740.

Horwath, M., Lemoine, J.-M., Biancale, R. & Bourgogne, S., 2011. Improved GRACE science results after adjustment of geometric biases in the Level-1B K-band ranging data, *J. Geod.,* **85,** 23–38.

Ihde, J., Wilmes, H., Müller, J., Denker, H., Voigt, C. & Hosse, M., 2010. Validation of satellite gravity field models by regional terrestrial data sets, in *System Earth via Geodetic-Geophysical Space Techniques,* pp. 277–296, eds Flechtner, F., Gruber, T., Güntner, A., Mandea, M., Rothacher, M., Schöne, T. & Wickert, J. Springer-Verlag, Berlin, Heidelberg.

Kim, J., 2000. Simulation study of a low-low satellite-to-satellite tracking mission, *PhD thesis,* Center for Space Research, The University of Texas at Austin, Texas, USA.

Keller, W. & Sharifi, M.A., 2005. Satellite gradiometry using a satellite pair, *J. Geod.,* **78,** 544–557.

Klees, R. & Broersen, P., 2002. *How to Handle Colored Noise in Large Least-Squares Problems - Building the Optimal Filter,* Delft University Press, DUP Science, Delft, The Netherlands.

Klees, R., Ditmar, P. & Broersen, P., 2003. How to handle colored observation noise in large least-squares problems, *J. Geod.,* **76,** 629–640.

Klees, R. & Ditmar, P., 2004. How to handle colored noise in large least-squares problems in the presence of data gaps? in *V Hotine-Marussi Symposium on Mathematical Geodesy, International Association of Geodesy Symposia,* Vol. 127, pp. 39–48, ed. Sansò, F., Springer, Berlin, Heidelberg, New York.

Kusche, J., Schmidt, R., Petrovic, S. & Rietbroek, R., 2009. Decorrelated GRACE time-variable gravity solutions by GFZ, and their validation using a hydrological model, *J. Geod.,* **83,** 903–913.

Liu, X., 2008. Global gravity field recovery from satellite-to-satellite tracking data with the acceleration approach, *PhD thesis,* Delft University of Technology, Delft, The Netherlands.

Lyard, F., Lefevre, F., Letellier, T. & Francis, O., 2006. Modelling the global ocean tides: modern insides from FES2004, *Ocean Dyn.,* **56,** 394–415.

Mayer-Gürr, T., 2006. Gravitationsfeldbestimmung aus der Analyse kurzer Bahnbögen am Beispiel der Satellitenmissionen CHAMP und GRACE, *PhD thesis,* University of Bonn, Bonn, Germany.

Mayer-Gürr, T., Eicker, A., Kurtenbach, E. & Ilk, K.-H., 2010a. ITG-GRACE: Global static and temporal gravity field models from GRACE data, in *System Earth via Geodetic-Geophysical Space Techniques,* pp. 159–168, eds Flechtner, F., Gruber, T., Güntner, A., Mandea, M., Rothacher, M., Schöne, T. & Wickert, J., Springer-Verlag, Berlin, Heidelberg.

Mayer-Gürr, T., Eicker, A. & Kurtenbach, E., 2010b. ITG-Grace2010: The new GRACE gravity field release computed in Bonn, *Geophys. Res. Abstr.,* **12,** EGU2010-2446.

Mayer-Gürr, T. *et al.*, 2012. The new combined satellite only model GOCO03S, in *Proceedings of the International Symposium on Gravity, Geoid and Height Systems (GGHS) 2012,* Venice, Italy.

McCarthy, D.D. & Petit, G., 2004. IERS conventions (2003), *IERS Technical Note 32,* Verlag des Bundesamtes für Kartographie und Geodäsie, Frankfurt am Main, Germany.

Migliaccio, F., Reguzzoni, M., Gatti, A., Sansò, F. & Herceg, M., 2011. A GOCE-only global gravity field model by the space-wise approach, in *Proceedings of the 4th International GOCE User Workshop,* 31 Mar–1 Apr, 2011, Munich, Germany.

Pavlis, N.K., Holmes, S.A., Kenyon, S.C. & Factor, J.K., 2008. An Earth gravitational model to degree 2160: EGM2008, *Geophys. Res. Abstr.,* **10,** EGU2008-A-01891.

Pavlis, N.K., Holmes, S.A., Kenyon, S.C. & Factor, J.K., 2012. The development and evaluation of the earth gravitational Model 2008 (EGM2008), *J. geophys. Res.,* doi:10.1029/2011JB008916.

Pail, R., Goiginger, H., Mayrhofer, R., Schuh, W.-D., Brockmann, J.M., Krasbutter, I., Höck, E. & Fecher, T., 2010a. Global gravity field model derived from orbit and gradiometry data applying the time-wise method, in *Proceedings of the ESA Living Planet Symposium,* 28 Jun–2 Jul, 2010, Bergen, Norway.

Pail, R. *et al.*, 2010b. Combined satellite gravity field model GOCO01S derived from GOCE and GRACE, *Geophys. Res. Lett.,* **37,** L20314, doi:10.1029/2010GL044906.

Rummel, R., Yi, W. & Stummer, C., 2011. GOCE gravitational gradiometry, *J. Geod.,* **85,** 777–790.

Sandwell, D.T. & Smith, W.H.F., 2009. Global marine gravity from retracked Geosat and ERS-1 altimetry: ridge segmentation versus spreading rate, *J. geophys. Res.,* **114,** B01411, doi:10.1029/2008JB006008.

Siemes, C., Haagmans, R., Kern, M., Plank, G. & Floberghagen, R., 2012. Monitoring GOCE gradiometer calibration parameters using accelerometer and star sensor data: methodology and first results, *J. Geod.,* **86,** 629–645.

Standish, E.M., 1998. JPL planetary and lunar ephemerides, DE405/LE405, Jet Propulsion Labratoary, IOM 312.F-98-048.

Tapley, B.D., Bettadpur, S., Watkins, M. & Reigber, C., 2004. The gravity recovery and climate experiment: mission overview and early results, *Geophys. Res. Lett.,* **31,** L09607, doi:10.1029/2004GL019920.

Tapley, B.D. *et al.*, 2005. GGM02 - An improved Earth gravity field model from GRACE, *J. Geod.,* **79,** 467–478.

Tapley, B.D., Ries, J.C., Bettadpur, S., Chambers, D., Cheng, M., Condi, F. & Poole, S., 2007. The GGM03 mean Earth gravity model from GRACE, in *Proceedings of the AGU Fall Meeting, Abstracts,* G42A-03.

Visser, P. *et al.*, 2009. Orbit determination for the GOCE satellite, *Adv. Space Res.,* **43,** 760–768.

Wahr, J., Molenaar, M. & Bryan, F., 1998. Time variability of the Earth's gravity field: hydrological and oceanic effects and their possible detection using GRACE, *J. geophys. Res.,* **103**(B12), 30 205–30 229.

Zhao, Q., 2004. Research on precise orbit determination theory and software for both GPS navigation constellation and LEO satellites, *PhD thesis,* Wuhan University, Wuhan, China.

APPENDIX: DGM-1S MODEL

One of the static gravity field models considered in this study is DGM-1S (Hashemi Farahani *et al.* 2013). Since the manuscript devoted to this model is currently in review, we briefly explain here the basic steps in the production of this model.

The primary data sets exploited in the production of DGM-1S have been already mentioned in the beginning of Section 3. The chosen *a priori* model of the static gravity field was ITG-Grace2010s complete to degree 180. The GRACE KBR and GOCE SGG data sets were respectively converted into residual range combinations and residual gravity gradients with respect to the *a priori* model, as explained in Section 2.2. It is worth mentioning that the force model exploited to process the GRACE KBR data did not include long-term gravity field changes (related to, e.g. hydrology, ice mass loss and postglacial rebound).

The low frequency noise in GRACE residual range combinations was eliminated with a high-pass filter built based on the empirical model of eq. (3). The low-frequency noise in GOCE residual gravity gradients was estimated and eliminated in a similar manner, but the list of unknown parameters only comprised the parameters x_0 till x_4. This led to elimination of noise in these residuals up to 1 cpr frequency.

Furthermore, kinematic orbits of GRACE and GOCE satellites were converted into residual 3-D acceleration vectors with a three-point differentiation scheme. The resulting vectors can be understood as satellite accelerations averaged with a certain weight within the differentiation time interval (Ditmar & van Eck van der Sluijs 2004; Ditmar *et al.* 2006), that is, as a 3-D analogy of residual range combinations.

The four sets of residual data mentioned earlier were jointly inverted into residual spherical harmonic coefficients. Each of the considered sets of residuals is characterized by its own dependency of noise on frequency. To ensure the statistically optimal data processing and combination, the frequency-dependent data weighting was applied to each of the data sets (Klees & Ditmar 2004; Ditmar *et al.* 2007). To that end, each noise PSD was represented in terms of an Auto-Regressive Moving-Average (ARMA) model (Klees & Broersen 2002; Klees *et al.* 2003). To account for a strong temporal variability of noise in GRACE residual range combinations, demonstrated by Ditmar *et al.* (2012), an individual ARMA model of the corresponding noise was built for each particular month.

The spherical harmonic coefficients were computed up to degree 250. A Kaula-type regularization for degrees above 179 was applied. The optimal value of the regularization parameter was obtained empirically by minimizing the global rms geoid height differences with respect to EGM2008 complete to degree 250. Finally, the ITG-Grace2010s coefficients up to degree 180 were added back to the estimated residual coefficients.