AN INITIAL INVESTIGATION OF THE GOCE ERROR VARIANCE-COVARIANCE MATRICES IN THE CONTEXT OF THE GOCE USER TOOLBOX PROJECT

Rory J. Bingham[1], **Christian Tscherning**[2], **and Per Knudsen**[3]

[1]*Newcastle University, United Kingdom*
[2]*University of Copenhagen, Denmark*
[3]*Danish National Space Center, Denmark*

ABSTRACT

The availability of the full error variance-covariance matrices for the GOCE gravity field models is an important feature of the GOCE mission. Potentially, it will allow users to evaluate the accuracy of a geoid or mean dynamic topography (MDT) derived from the gravity field model at any particular location, design optimal filters to remove errors from the surfaces, and rigorously assimilate a geoid/MDT into ocean models, or otherwise combine the GOCE gravity field with other data. Here we present an initial investigation into the error characteristics of the GOCE gravity field models as they are realised in the calculated geoid anomalies. We examine how the error variances depend on location and spatial scale, and how they differ between models. The inhomogeneous and anisotropic nature of the error covariances are explored, and an example of how this information may be used in ocean data assimilation is provided. Finally, we consider some of the practical issues relating to the handling of the huge files containing the error variance-covariance information.

Key words: GOCE, geoid errors, error covariances.

1. INTRODUCTION

The GOCE satellite was successfully launched in March 2009. The first level 2 datasets, based on two months of observations, were released to the scientific community in June 2010, with a second release, based on 8 months of observations following in March 2011. The primary products of the GOCE mission are the earth gravity models (EGMs), of which three variants are provided. The EGMs allow the geoid to be computed. A distinguishing feature of the GOCE mission is that, for the first time, the full error variance-covariance information associated with the EGMs will be also be provided to the user community, allowing, for instance, the error characteristic of the geoid, or other derived products to be studied and accounted for in subsequent applications. Because this error information has never previously been available to the user community, the tools and methods to fully exploit this information have yet to be developed. Some potential applications include the design of optimal filters and the characterization of errors magnitudes and length scales in mean dynamic topographies (MDT) derived from the gravity field models to enable the rigorous assimilation of an MDT into an ocean model. However, it may well be many years before the user communities have gained sufficient knowledge and experience to fully exploit the potential of the variance covariance information.

The European Space Agency have funded the development of the GOCE User Toolbox (GUT) to facilitate and ease the use of the GOCE products. In mind, has been the wider scientific community who may wish to use these products in their particular scientific area but who may be deterred by the unfamiliar nature of the spherical harmonic expression of many of the gravity field products. Supplied as part of the GUT package, put standing alone from GUT itself, is a set of tools developed by *Balmino* (2009) to enable a range of error calculations using the GOCE error variance covariance information. (So far, it has been considered beyond the scope of the GUT project to fully integrate these tools into GUT, but this may be done at a later date.)

Below we present two examples of the error calculations that may be performed using the Balmino toolbox. This should be considered only a very preliminary investigation of the error characteristics of the GOCE EGMs, and a full interpretation of the results cannot, as yet, be given. We start in the next section with an examination of the error characteristics of the GOCE geoids, considering the differences between the models and how the accumulated errors depend on truncation degree and order (d/o). Then, in Section 3, we examine how the error covariances of the GOCE geoids depend on location and vary between the models. We also show the impact of filtering on the error characteristics and consider how this may feed into a scheme for the assimilation of a GOCE MDT into an ocean model. In Section 4 we provide a brief discussion of some of the practical issues encountered in handling the error information. Finally, some conclusions are given.

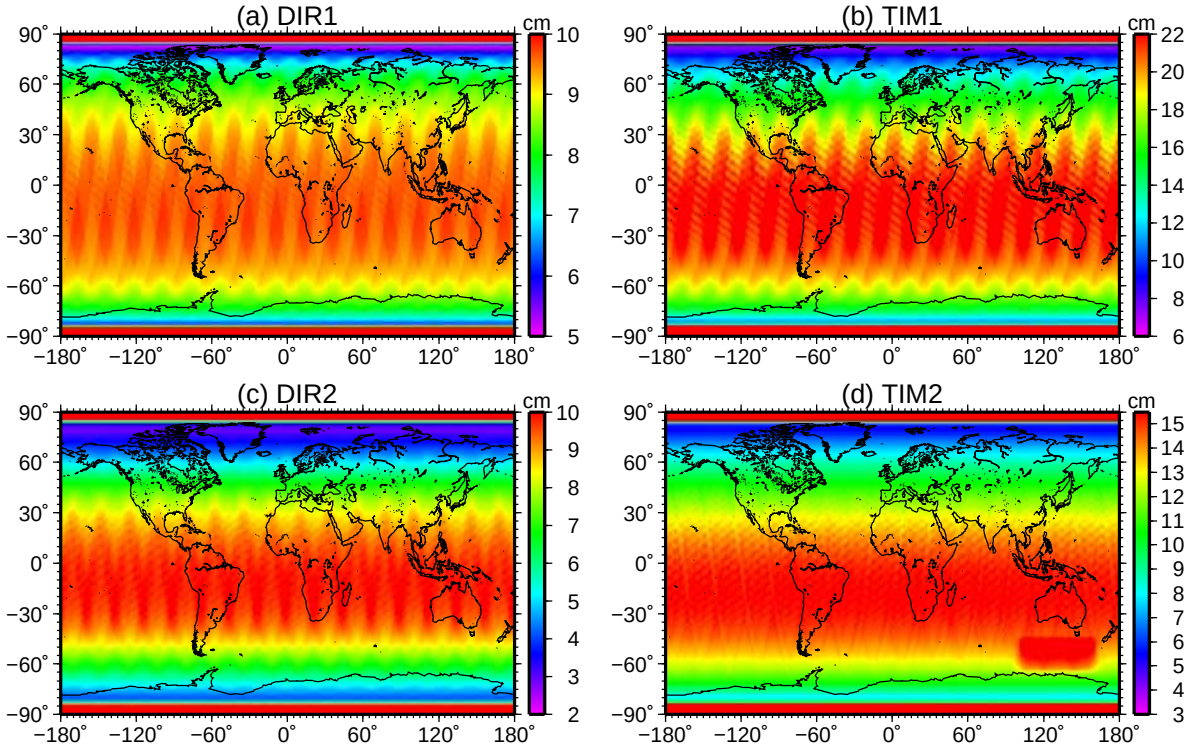

Figure 1. Geoid error maps associated with some of the GOCE gravity models. (a) The first release of the direct solution. (b) The first release of the timewise solution. (c,d) Repeating (a,b) for the second releases. All models are truncated at degree and order 224.

2. GEOID ERRORS

The first question a user who has calculated a GOCE geoid is likely to ask is "what are the magnitude of the errors on the geoid estimate?". Because this will depend on the degree and order to which the geoid has been calculated – errors grow with increasing d/o – it is not possible to provide a single error map. Availability of the error variance covariance information, together with the Balmino covhsmp routine, allows the user to calculate an error map to the required degree and order and on the same grid as the geoid.

Examples of such maps, calculated on a global 1x1 degree grid, for the first and second releases of the direct (DIR1/DIR2) and timewise (TIM1/TIM2) models are provided in Figure 1. Because errors grow with increasing degree and order, in each case the errors correspond to geoids truncated and the highest common degree and order of 224 as set by the TIM1 model. This allows the fairest comparison of the errors between models and generations. For the first generation of models (top row) we see that the errors of the DIR1 geoid are at least a factor of 2 smaller than those of the TIM1 model. This is because the GRACE EIGEN5C gravity model, which also includes surface data, was used as an a priori constraint on the DIR1 solution. The DIR2 model (Fig. 1c) does not use this constraint and for this reason the errors are a little greater, even though the second release is based on four times more GOCE observations. The true impact of the additional data is seen by comparing the error maps for TIM1 with TIM2 (Figs. 1b and 1c). The maximum errors have been reduced by about 6 cm. The greater coverage has also lead to a more zonally homogeneous error map with less trackiness.

In general, for all the models, although some zonal variation results from the satellite orbits, the variation in error magnitude is primarily meridional, with errors at their lowest just below the polar gap (approx. 86 degrees) and at their greatest at tropical latitudes. We can exploit this and characterise the errors by their zonal mean values to allow a more qualitative and succinct examination of how the errors vary between models and grow with truncation d/o. To this end, the zonal mean errors as a function of d/o for the first and second generations of the various GOCE models are shown in Fig. 2. Comparing TIM1 and TIM2 for d/o=224 (the maximum for TIM1), we find that the maximum error in TIM2 has been reduced by 6 cm to just under 16 cm compared with TIM1, while for d/o=250 (the maximum for TIM2) the maximum errors are 23 cm, similar to those for TIM1. For large d/o, a hemispheric asymmetry is clear, with maximum errors at about 15S, and errors a few cm lower at high northerly latitudes than at similar southern hemisphere latitudes. This is related

to the elliptical satellite orbit. As already mentioned, DIR1 uses the EIGEN5C GRACE gravity model as a constraint. For d/o> 150, this suppresses the growth of errors seen in the other models (first and second releases), such that for d/o=240 the maximum errors of DIR1 are nearly half those of TIM2. The DIR2 solution does not use this a priori constraint and consequently the errors are similar to those for DIR1, although still lower than those of TIM2 for d/o> 150. Comparing with the GRACE errors, we see that for d/o< 100, the errors from GRACE are lower than those from the GOCE models, but by d/o=150 the reverse is true. This is what we expect from the two missions.

3. ERROR COVARIANCES

A more abstract notion than error variance (the diagonal elements of a variance-covariance matrix) is error covariance (the off-diagonal elements). In simple terms, the error covariance shows the degree to which errors at two points are related. For any point, it is possible using the GOCE error variance covariance information, together with the Balmino covhs2p routine, to calculate a discrete error covariance function, which shows how the error at that point is related to the errors at surrounding points. With covhs2p it is possible to specify a grid over which to calculate the error covariance functions (the inner-zone) and the boundaries to which the error covariance function for each point will be calculated (the outer zone).

In Fig. 3 we compare the normalised error covariances (scaled by a factor of 10) at three latitudes for the second generations of the TIM (top row) and DIR (middle row) models and the first generation of the SPW model (bottom row). Clearly the solution method has a big impact on the error covariance structure, with the DIR2 model error covariance at lower latitudes being much more anisotropic than the TIM2 or SPW1 solutions. The covariance for the SPW1 model is less structured than for the other models. Towards higher latitudes, the error covariance structures of the models become more similar, with all being more isotropic. To allow direct comparison the three locations chosen for Fig. 3 are the same as those used in Balmino (his Fig 2) where the simulated GOCE error covariance maps are shown. For reasons that are unclear, the actual errors shown here lack the longer wavelength patterns of the simulated covariances.

From the error covariance maps it is clear that the primary anisotropy is with respect to the zonal (x) and meridional (y) directions. In Figure 4, we therefore plot the x and y cross sections of the error covariances for the TIM2 and DIR2 models for the three locations for which maps are given. The top two rows show respectively, the x and y components for the TIM2 model, while the bottom two rows repeat this for the DIR2 model. The left panels show error covariances for the unfiltered geoids, while to the right the error covariance profiles resulting from the application of a 200 km Gaussian filter are shown. In the filtered cases the actual, rather than normalised, values are shown so that the impact of filtering on the error variances can be seen.

Clearly the orbital configuration of the satellite impacts on the error covariance structure. Short wavelength noise tends oscillate between decaying positive and negative correlation in the zonal (across track) direction, but decays more straightforwardly in the meridional (along track) direction. Filtering reduces the error variance, as it should. But it also increases the correlation length scales, since the error covariances now reflect errors in longer wavelength components of the gravity models. Error correlation length scales for the filtered geoid are much greater for the meridional direction because of the satellite orbit. For the DIR2 model, the meridional error covariance profiles decay more slowly, particularly at lower latitudes. This may explain why the filtered errors of the DIR2 geoid are twice those of the TIM2 geoid, even though the unfiltered errors of the former are much lower than those of the latter, and may indicate that anisotropic filtering is more appropriate for the DIR2 model.

3.1. Defining error length scales

In addition to error magnitudes, knowledge of error length scales is required for the assimilation of a geodetic MDT into an ocean model. For now we must assume that MDT errors results entirely from the geoid, since error covariance information for the mean sea surface is not available. Due to the anisotropy discussed above, we can expect long wavelength errors in the meridional (along track) direction to have greater length scales than those for the zonal (across track) direction. For each location the filtered error covariance function is approximately symmetric about the zonal and meridional axes and has an approximately Gaussian form as shown in Figure 4. This permits the zonal and meridional correlation length scales to be defined as the distance at which the covariance drops to e^{-1} of its value at the origin, the so-called e-folding length scale.

The zonal and meridional correlation length scale maps for the geoid/MDT derived from the TIM2 model, filtered with a 200 km Gaussian filter and shown in Figures 5a and 5b. The zonal length scales vary between about 340 to 400 km, and, like the errors themselves, vary mainly in the meridional sense, growing larger towards the poles, and bulging toward the equator. The orbital configuration of the satellite is reflected in the longitudinal variations of the zonal length scale. As noted above, the meridional length scales are much greater than the zonal scales, varying between about 900 and 980 km, because errors in the long wavelength range tend to be related along the satellite orbits. Again, the variation in length scale is primarily meridional, but, compared with the zonal length scales, there is a greater degree of zonal variation, with notable bulges over a range of longitudes where the orbits cross the problematic region south of Australia.

The zonal average scales to be used in the assimilation are shown in Figure 5c. Presently, it is unclear the impact

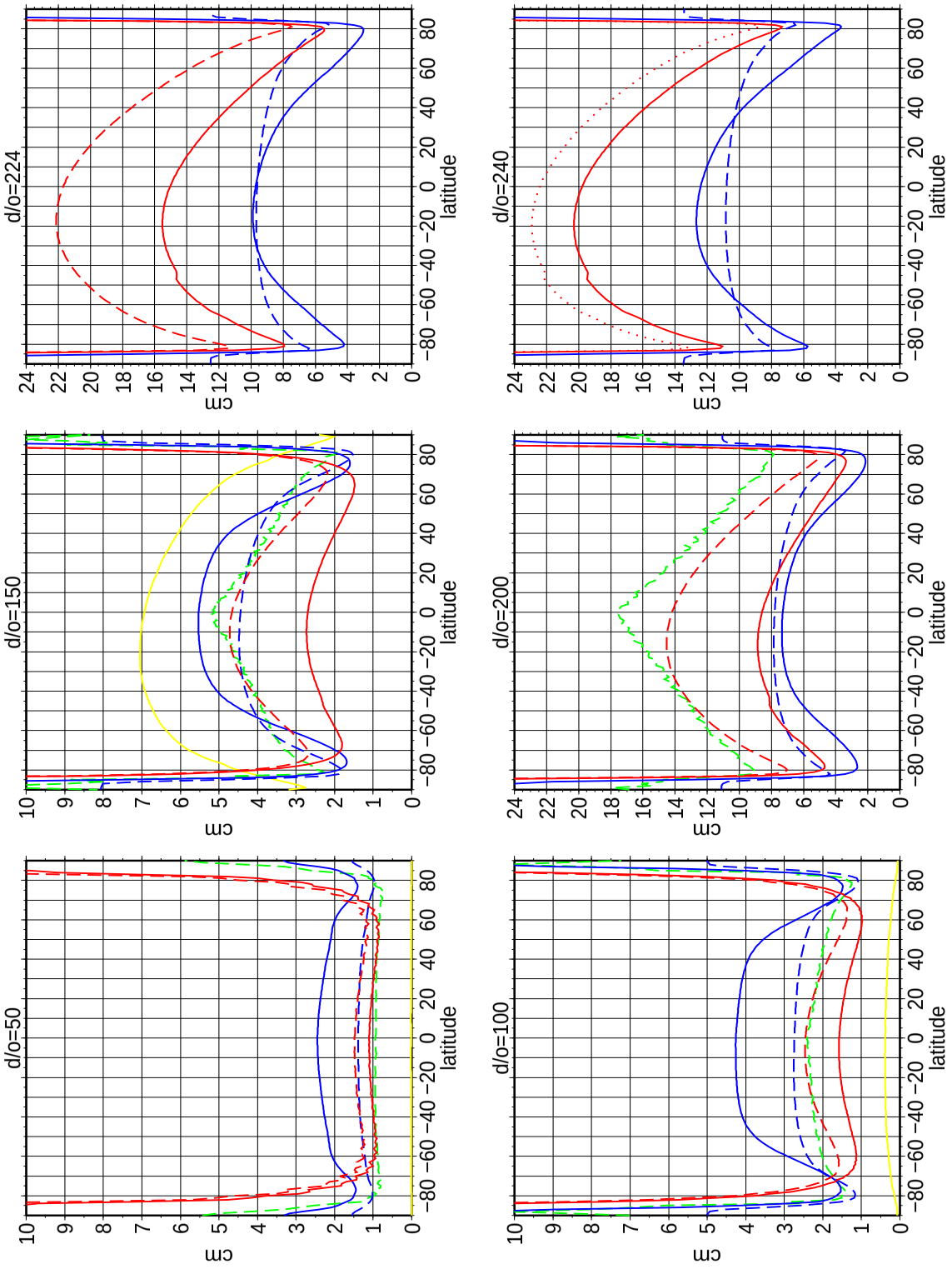

Figure 2. Zonal mean errors for the GOCE geoids: DIR models (blue); TIM models (red); SPW model (green). First generation models are given by the dashed curves. The red dotted curve in the lower right panel is for the TIM2 model at d/o=250. For reference errors for the GRACE EIGEN-GL04S1model are also given to d/o=150 (yellow).

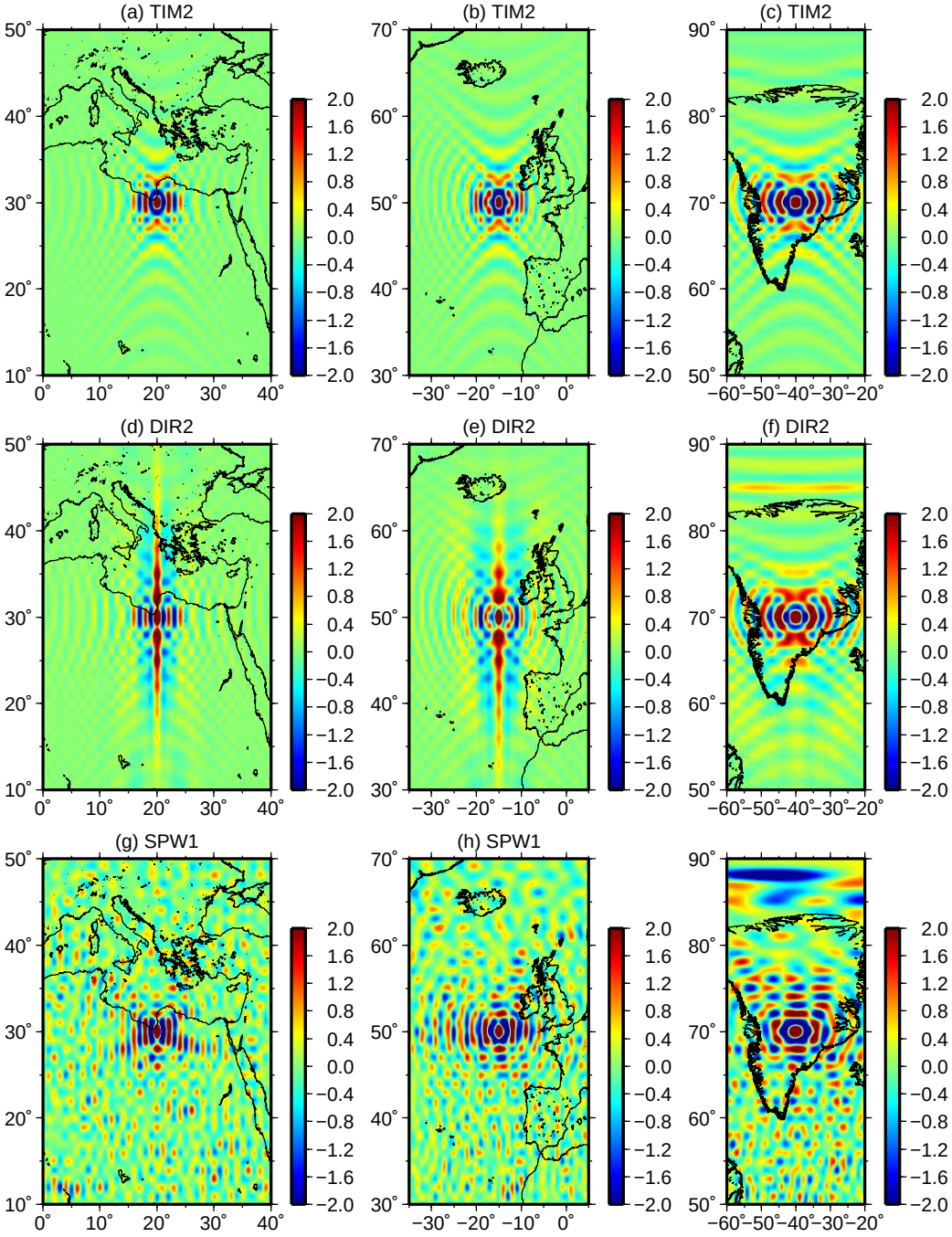

Figure 3. Geoid error covariance maps at three locations for the second generation timewise (top row) and direct (middle row) models and first generation spacewise model (bottom row). (Maps scaled by a factor of 10 for display).

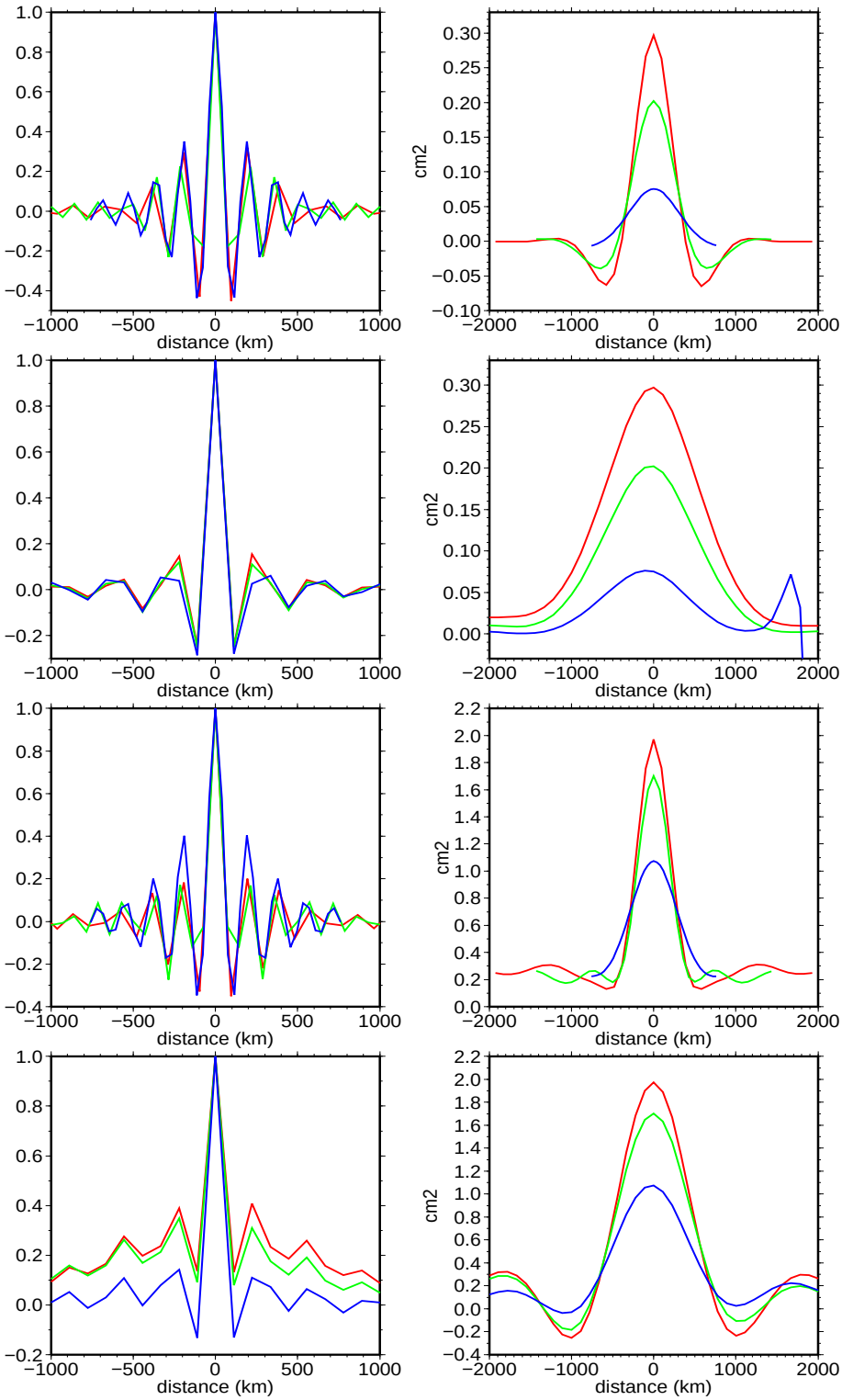

Figure 4. Zonal and meridional cross sections of the geoid error covariances shown in Figure 3. The top two rows show, respectively, the zonal and meridional cross-sections for the TIM2 model, while the bottom two rows repeat this for the DIR2 model. The locations are: 30N,20E (red); 50N,15W (green); 70N, 40W (blue). The right hand panels repeat those on the left but with a 200 km Gaussian filter applied to the geoids. For these, actual, rather than normalised, values are shown.

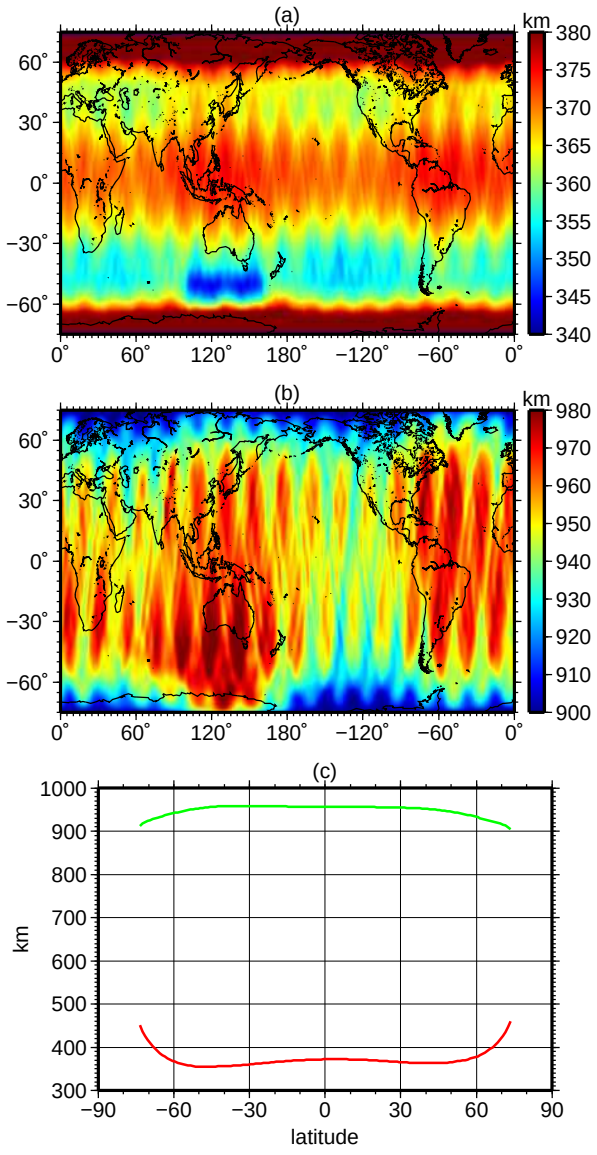

Figure 5. (a) Zonal correlation length scales for an MDT based on the GOCE second generation timewise model filtered with a 200 km Gaussian filter and assuming there are no errors in the MSS. (b) As in (a) but for the meridional correlation length scales. (c) The zonally average zonal (red) and meridional (green) correlation length scales to be used in the assimilation scheme.

these length scales and the large difference between the meridional and zonal lengths will have on the assimilation. (See *Haines et al.* (2011) elsewhere in this volume for further details.)

4. PRACTICAL ISSUES

The GOCE variance-covariance matrices (VCMs) are obtained from the ESA virtual server as compressed tarballs (*.TGZ). The size obviously depends on the maximum degree and order of the model, with the files for the TIM2 being the largest at 20GB. Uncompressing and unpacking one of these files yields a header file (*.HDR) and another tarball (*.DBL in release 1, now *.TAR) which is 52GB in size for the TIM2 model. Unpacking this second tarball, gives a set of ascii files containing the VCM coefficients for each order plus a *.IIH which describes the ordering of the coefficients.

The VCM computations described in this paper were performed using software described by Balmino (2009) which is supplied as part of the GUT package. The first challenge faced by novice users is to convert the ascii VCM data into a format compatible with the Balmino software, which requires a full square matrix given as an unformatted sequential access file with one record corresponding to one row of the VCM. Further, the rows/columns must be ordered by increasing spherical harmonic order, and within this by increase spherical harmonic degree with the even (C) and odd (S) coefficients supplied as pairs. An additional complication is that while the odd and even coefficients for the DIR and SPW models are ordered as pairs as required by the Balmino routines, for the TIM solution they are not, and an extra step is required to reorder the coefficients. For the TIM2 model the required VCM file has a double precision size, as required by the routines, of 32GB. To ease the use of the GOCE VCMs we have developed a small set of Fortran utilities to convert the ascii files into the form required by the Balmino routines. These will soon be made available on the GUT website.

The error maps shown above were computed using the Balmino covhsmp routine on a 1x1 degree global grid. On our average PC the running time for this calculation for the TIM2 model was under 100 minutes. The covariances were computed using the Balmino covhs2p routine with the directing file parameters as in the example supplied in the GUT package. The displayed covariances are therefore taken from a computation where for each point on a 1x1 degree grid between 20-80N and -60-30E, the discreet error covariance function is computed on a 1x1 degree grid in a 40x40 degree window. For this calculation, the computation time on our typical PC for the TIM2 model was about 1 hour. Thus, we see that the despite the large VCMs the error calculations are unlikely to be prohibitively expensive for most users, with the greatest barrier perhaps being disk space, particularly if the maximum degree and order of the GOCE models increases much beyond its present limit of 250.

5. CONCLUSIONS

In this paper we have presented a very preliminary investigation of the error characteristics of the GOCE gravity field models in terms of geoid heights. Calculations were performed using the software developed by *Balmino* (2009), which is supplied as part of the GOCE User Toolbox (GUT). We find the error propagation calculations with this software to be reasonably quick on a standard PC. For instance, calculating a global error map on a 1x1 degree grid for a geoid truncated at d/o=250 took less than 100 minutes. Similarly, the computation of the error covariance map in a 40x40 degree window, took no more than a few minutes per point. This may obviate the need to find fast approximations to the error covariance functions. For the authors, the biggest initial challenge was putting the error variance covariance information as supplied by ESA into the form required by the Balmino routines. Having written the required interfaces these will now be made available to the user community via the GUT webpages. The ascii files containing the error information are large (up to 50GB) which may present a problem for some users.

The interpretation of the error variances is quite straightforward, and show the properties we expect: for a given d/o, errors vary mainly with latitude, are at a minimum just below the polar gaps and at a maximum near the equator; errors grow with increasing d/o; the greater data span for the second release means the errors in the timewise model have been reduced by several cm; the a priori constraint used in the first direct model meant that the errors of the first direct model where much lower than either the timewise or spacewise models, and are similar to the second direct model, for which the a priori constraint was dropped; for d/o up to 100 the errors from GRACE are smaller than those from GRACE, but the situation is reversed at d/o=150. In contrast, the physical interpretation and significance of the error covariance is more challenging, and more work will be required to fully understand and exploit this information.

ACKNOWLEDGMENTS

This work was carried out under the ESA funded GUT project and we acknowledge the many useful discussions with our consortium colleagues.

REFERENCES

Balmino, G. (2009), Efficient propagation of error covariance matrices of gravitational models: application to GRACE and GOCE, *Journal of Geodesy*, *83*, 989–995.

Haines, K., D. Lea, and R. Bingham (2011), Using the goce mdt in ocean data assimilation, in *Proceedings of the 4th International GOCE User Workshop, 31 March - 1 April, Munich, Germany*.