ASPECTS OF GOCE CALIBRATION

R.Koop[(1)], **P.Visser**[(2)], **C.C.Tscherning**[(3)]

[(1)] *Space Research Organization Netherlands (SRON)*
Sorbonnelaan 2, 3584 CA Utrecht, The Netherlands
Email: r.koop@sron.nl

[(2)] *Delft Institute for Earth Oriented Space Research (DEOS)*
Kluyverweg 1, 2629 HS Delft, The Netherlands
Email: pieter.visser@lr.tudelft.nl

[(3)] *Department of Geophysics, University of Copenhagen (UCPH)*
Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark
Email: cct@gfy.ku.dk

INTRODUCTION

A gradiometer and a space-qualified GPS receiver having the specifications necessary for the high demands of the GOCE mission are under development. Obviously, the technical demands for these instruments are very stringent. However, instruments are never ideal in reality. Deviations which the material realization of the instrument has from the ideal design on paper are referred to as imperfections. In principle, the accelerometers, the gradiometer and the GPS receiver will have deviations from their ideal design leading to non-ideal measurements. When these deviations are out of the specifications, calibration should provide information to correct for the imperfections as good as possible. Apart from such instrumental imperfections, other errors of different origin may occur during the measurement process. Furthermore, somewhere in the measurement process we have to establish that we have obtained the observations in the required (physical) units and that no unknown scale factors or other corrections remain. Also here calibration is required. Finally, not only the observations but also their error-estimates have to be calibrated. Calibration is thus an indispensable activity in the process of obtaining the highest possible accuracy of the observations.

CALIBRATION AND VALIDATION

In general, calibration and validation (cal/val) are important in the context of data quality assessment. Furthermore, in the case of GOCE, cal/val is part of a complicated data processing activity. So it makes sense to start with specifying rather precisely what we mean with the terms calibration and validation in the context of GOCE. It appears to be convenient to link these terms to the GOCE data levels ([4]). So from now on, when we use the word calibration, we refer to level 1b and with validation we refer to level 2. To be more precise: we agree to designate the level 1b data as (fully) calibrated and corrected observations, which means that calibration as an action to be performed has already taken place in the data processing prior to the level 1b data. With validation we refer to any action which is taken on the level 2 data as it is derived from level 1b data. Arguments in favor of this division are:
in the case of GOCE there is a clear division of responsibility with respect to the data levels: ESA is responsible up to level 1b, and the science community is responsible for the data handling from level 1b up to level 3;
the data processing from level 0 to level 1b is more or less well-defined and straightforward, whereas the level 1b $\rightarrow$ 2 processing is a complicated estimation process for which several valid methods exist and which cannot be simply inverted (i.e. in the sense that the actual data is reconstructed from the estimated geopotential coefficients).

Calibration On the one hand calibration is concerned with *methods* to determine certain parameters (e.g. scale factors, biases) to be applied to the measured quantities in order to correct for errors in the measurement process and to obtain from the instrument read-outs the required observations in known physical units. Not only the calibration parameters themselves should result from the methods, preferably also a description of their accuracy and reliability. (For simplicity, we will disregard here the matter of scale factors being dimensionless or not.) On the other hand calibration is sometimes understood as the *procedure* itself (the numerical or computational action) of correcting the measured quantities using the calibration parameters. Of course, both aspects of calibration are important here since a successful level 1 $\rightarrow$ 2 data processing depends a.o. on the reliability and correctness of the methods with which the calibration parameters have been determined and applied. Here we define calibration as: *the procedure to determine parameters and the application of* *these parameters to the instrument read-outs in order to obtain quantities in the required physical units and dimensions and with the required and known accuracy*. In particular the matter of accuracy is important here since the application of calibration parameters is in fact a correction of the data for (systematic) errors of instrumental or other nature, so the given or assumed measurement error budget refers to the calibrated quantities. Error propagation of the uncertainties in the calibration parameters could be included. This may be a complicated issue in case calibration parameters may appear to be frequency dependent. We remark that, for a proper calibration procedure, it has of course to be made perfectly clear exactly where (at which step) in the level $0 \rightarrow 1$b data pre-processing the calibration parameters are to be applied.

Validation Loosely speaking, validation has to do with making sure that the level 2 products, i.e. in the case of GOCE potential coefficients, gridded geoid heights and gravity anomalies, refer to the real Earth's gravity field, according to present-day independent knowledge about this field. This means that we compare the products, which have been derived from the GOCE measurements, with existing, external knowledge about the real gravity field (an agreed upon standard would even be better), in order to make it more certain that they are correct. So we here define validation as: *the application of methods to compare data products derived from the measurements with existing independent data or knowledge in order to assess the quality of the data products and to make sure that the measurement process, error estimation and calibration have been performed well*. The main difference between validation and calibration, as we interpret these terms here, has to do with the fact that calibration concerns corrections and conversions finding their cause in aspects of instrumental nature or measurement setup, while validation concerns no corrections or conversions but merely a test of the results compared to existing independent knowledge. Furthermore, validation tests may reveal shortcomings in retrieval models and methods and may also give some clues pointing at possible outliers or gross errors in the data. In this sense, validation is, like calibration, closely connected to (quick-look) data quality assessment.

Outliers and failures Apart from instrument or satellite induced systematic errors on the data, there may occur more or less randomly outliers of incidental nature. Outliers can occur on all data levels. Usually the best thing to do is to throw such outliers out of the data set or to flag them by assigning a large standard deviation to them, since often there cannot be found a reliable way of correcting such data. In order to do so, thresholds have to be set and the outliers crossing the threshold have to be detected. For this, standard statistical testing procedures are available. Also by means of so-called "continuity checks" outliers could be found. Furthermore, "internal consistency checks" can be performed in case different sensors measuring related quantities are available or making use of redundancy in the data. Also "external consistency checks" can be performed, comparing the observations with independent external knowledge. Such methods are to some extent related to cal/val activities. Furthermore, there is a relation with the detection and resolving of failures, although failures would typically result in offsets in the data or in the error, remaining there until the failure is resolved, while with outliers, individual wrong measurements are meant.

GOCE CALIBRATION

From now on, we will focus specifically on the calibration issue (on level 1b) with GOCE. A full calibration procedure for GOCE is rather complicated, see e.g. [1], [6]. For convenience, we will divide the full calibration procedure into several steps, namely calibration activities that take place before the mission on ground, activities that are performed (repeatedly) in-flight during the mission and post-mission calibration activities. These three steps, which will become more clear in the sequel, will be respectively referred to as: pre-flight calibration, internal calibration and external calibration. The term "full calibration" then refers to the ensemble of activities in the three steps. It should be clear that all three steps are necessary to obtain the best data product and not one step can be seen separately from the others. Table 1 gives a schematic overview of the whole GOCE calibration setup.

Satellite Gravity Gradiometry Observations

Pre-flight calibration This calibration step constitutes (a) a first estimate of the so-called common mode rejection ratio (CMRR) and differential mode rejection ratio (DMRR) to the 10^{-4}-level and (b) makes sure that the accelerometer output is obtained in the required physical units to a certain level of accuracy, say to within 1%. The CMRR and DMRR refer to the coefficients which account for common mode and differential mode couplings due to imperfections of the gradiometer instrument, like mis-alignments, differential scale-factors, etc. Pre-flight calibration is performed before the mission on-ground using a test-bench. Typically, this is a hardware-related job and it is a task of the industry. Performance figures can be obtained from the manufacturers own experience with the instrument (experimental, simulation analysis, etc.) or with other but similar instruments. Actually, the problem with this calibration step is that the 1 g environment on-ground

Table 1: Overview of GOCE calibration

	1) pre-flight	2) internal	3) external
Satellite Gravity Gradiometry (SGG) observations:			
CMRR/DMRR	10^{-4}	10^{-5}	-
units/dimensions	$\sim$ 1%	-	TBD
	on-ground	"shaking"	using external data
Satellite-to-Satellite Tracking (SST) observations:			
	GPS receiver	-	orbit solution
Accelerometer common mode observations:			
	see SGG	see SGG	how?

prevents a proper calibration of the accelerometers since they are designed to operate with the highest accuracy in a 0 g environment. So, the question has to be addressed if the calibration parameters obtained on ground are valid in a 0 g environment and whether they are affected by the launch conditions or not. Item (a) mentioned above relates to internal calibration (coming next), while item (b) is actually part of the external calibration process (see later on).

Internal calibration Basically, internal calibration (also called relative calibration) has to do with the determination of the CMRR and DMRR to the 10^{-5} level. At this level, the influence of couplings between the common mode and differential mode output of the gradiometer remains within the required error budget. Internal calibration is done on board of the satellite (*on-board calibration*) during the mission (*in-flight calibration*) making use of the thrusters. A procedure has been designed which makes it possible to first calibrate the thrusters using the (un-calibrated) accelerometer measurements. Then, the procedure continues by activating the calibrated thrusters in a certain mode, putting a known non-gravitational signal on the accelerometers. The signals of all accelerometers are then analyzed. From the differential mode signals along each arm the internal calibration parameters are deduced. Specifically, the internal calibration procedure gives us information on relative scale factors and relative mis-alignments between accelerometer pairs on each arm and on common mis-alignments of accelerometer pairs on each arm. It is expected that, before calibration, these values will be larger than the requirements, so that indeed the data correction is needed. Information on non-orthogonality of the gradiometer arms is not obtained from the internal calibration procedure, neither is information on on-axis mis-placements of the accelerometers. For these effects it is expected that the instrument is within specifications, so that corrections are not needed, or that they are corrected for in the external calibration step (see below). The internal calibration is sometimes referred to as "relative" calibration, since the ratio of the sum and the difference of the accelerometer scale factors is determined. No "absolute" scale factors are determined from the internal calibration procedure. This makes it clear that another step is needed to do the full calibration in the sense of the definition above, i.e. to obtain quantities in known physical units with known accuracy. This step is the so-called external calibration discussed next.

External calibration In this step it is established that the gradient observations at data level 1b really represent the gradients of the gravity field in Eötvös Units, that remaining unresolved errors are corrected for and that the observations are obtained with the required (or in any case with a known) accuracy. Typically, external data are used in this step, e.g. ground data or existing models, hence its name "external calibration". External calibration is sometimes referred to as absolute calibration, but this term is less favorable.

Satellite-to-Satellite Tracking Observations

Here, when we talk about SST, we refer to the orbit determination from GPS measurements. In the SST level 1b $\rightarrow$ 2 processing the gravity field parameters are derived from the combination of the GPS observations and the accelerometer common mode observations (the latter are also to be regarded a level 1b data product, see below). The calibration issue for SST is less complicated than for SGG, because GPS is a mature technique for which the instrumentation can be properly calibrated on-ground, so that complications as with the SGG internal calibration can be avoided, and because the goals for gravity field determination from SST are less challenging than for SGG in terms of resolution. Furthermore, from the data level definitions for SST we see that, as far as orbit determination is concerned, the level $0 \rightarrow 1$ processing consists of precise orbit determination from GPS, which is a more or less off-line task in which the data to be calibrated on level 1 are the satellite positions. In the case of SST, the pre-flight calibration (see Table 1) consists of calibration of the GPS receiver. This is a rather straightforward task, for example by comparing observations made by it with other GPS receivers, collocated SLR (Satellite Laser Ranging) ground stations, accurately determined baselines, etc. It is fair to assume that the GPS performance will not be affected significantly by the launch conditions, so that no in-flight internal calibration step similar to that for SGG is required. An external calibration step may be required, though, which actually would mean the calibration of the GPS-based orbit solution. This poses a more principal problem, because it requires a comparison of the orbit solution with external data, requiring a second, independent orbit determination from other measurements or models (see also next chapter).

Accelerometer Common Mode Observations

The accelerations derived from the gradiometer common mode output constitute a separate level 1b product. In the level $1 \rightarrow 2$ data processing step for the determination of gravity field parameters from SST observations, the common mode observations are required for a proper separation of the non-conservative accelerations from the gravity field accelerations. In general, these common mode observations thus have to be calibrated as well. In practice the need for accurate calibration depends on the required accuracy of the observations. We remark here that, since both the common mode observations and the gravity gradients are derived from the output of the same hardware device (the six accelerometers), there will be a link between the calibration of the common mode and that of the SGG-observations. In fact, the pre-flight and internal calibration steps, as discussed above for SGG, are part of the common mode calibration procedure as well, since they apply to the different combinations of accelerometer outputs. However, in the same way as an external calibration step is necessary for the gravity gradient observations, as they are derived from the differential mode of the gradiometer output, we need an external calibration step for the common mode accelerations as well. The problem is that there is not a known unique way (yet) in which the differential mode external calibration parameters (i.e. scale factors and biases on the gravity gradients) can be attributed to individual accelerometer outputs so that the common mode would be externally calibrated as well. Although relative scale factors between accelerometer pairs are determined from the internal calibration procedure so that the common mode is corrected for the differential mode coupling, the separation of the rotational terms from the gravitational terms in the differential mode for all three gradiometer arms together prevents a straightforward attribution of SGG calibration scale factors to individual accelerometer outputs. What we do know is that the internal calibration procedure does not tell us anything about relative biases between accelerometer pairs. Whether all this will be a problem depends a.o. on the accuracy of the common mode observation required for SST level $1 \rightarrow 2$ data processing. Furthermore, since DFC (Drag Free Control) is applied, we are dealing here with residual accelerations remaining for certain specific frequency bands.

METHODS FOR EXTERNAL CALIBRATION

An accurate calibration of the level 1b data product would in the ideal case lead to a properly calibrated level 2 product as well. This product will consist of a set of spherical harmonic coefficients that can be converted, following well-known procedures, into grids of gravity anomalies and geoid undulations in a straightforward manner. Note that the calibrated gravity gradients themselves (a level 1b product) are also an important geophysical data product. The level 2 data products can be compared, for example, with existing data sets of in situ gravimetry observations. In the above, this process is referred to as validation. Such validation can thus provide information about the success of the calibration of the level 1b products. This calibration for GOCE has here been identified as a multi-step procedure, consisting of a pre-flight, internal and external calibration step. In the following the focus will be on methods for the external calibration of the level 1b products only. We distinguish here between several classes of external calibration methods: 1. comparison with ground-based gravity data; 2. comparison with existing global gravity field models; 3. comparison with other satellite data from missions like CHAMP or GRACE; 4. inter-comparison between different types of GOCE data. Probably, in reality, a combination of some of these methods will be used, since each of them has its own characteristics in terms of accuracy, frequency band, computational effort, etc. Here we will give a short description of these methods.

1. Comparison with ground-based gravity data The use of terrestrial (or airborne) gravimetry data for the external calibration of GOCE gravity gradients is an obvious choice. The idea is to use gravity data in well-surveyed areas for upward continuation to gravity gradients at satellite altitude, see e.g. [2], [3]. Alternatively, SGG and SST data could be converted into gravity anomalies or geoid heights and continued downward to the surface to be compared with the independent terrestrial gravity data. Although this idea seems very simple, there are a lot of issues to be addressed: the filtering of the data; the (definition of) reference systems and the datum problem; the required quality (covariances, calibration (!)) of the ground data; the use of normal points or not; the use of a priori gravity field models; the resolution of the ground data; the location of the test area; the possible need for taking topography into account or using terrain effect smoothing (RTM); the size of the geographical area; the minimum data collection period to be considered for the satellite data; the error characteristics of the satellite data; methods for converting gravity data into SGG and SST data (including interpolation, filtering, etc.); appropriate upward and downward continuation methods; numerical (in)stability of downward continuation; need for additional supporting measurement campaigns, either terrestrial or airborne.

2. Comparison with existing global gravity field models Over the years, a very extensive data set of both satellite tracking, satellite radar altimetry and terrestrial gravimetry observations has been accumulated from which so-called satellite-only and combined global gravity field models have been derived. The idea is to compute from these models SGG and SST data at satellite altitude and to compare it with the GOCE data. Among the points to be addressed here are: the resolution of the models and their spectral error characteristics; the calibration of the data on which the models are based, i.e. the calibration of the models themselves; the error characteristics of the satellite data. The J2 ("flattening") terms of the gravity field and the central term, as they are part of existing global gravity field models, may also be exploited for a check of the SGG observations. Questions to be addressed here concern the proper operation of the gradiometer at the same frequencies as where the model (J2+central) terms are visible and the sensitivity of the SGG observations at these frequencies (remember high performance in the gradiometer measurement bandwidth and $1/f$ noise below).

3. Comparison with other satellite data from missions like CHAMP or GRACE This method relates to the previous one in as far as updated satellite-only or combined global gravity field models derived from CHAMP and/or GRACE data would be used as "existing" global gravity field knowledge. At the time of the GOCE mission state-of-the-art global gravity field models from CHAMP and GRACE are expected to be available. Important to understand here is in how far the calibration issue of CHAMP and GRACE will be propagated into the GOCE calibration if these data would be used here too. Furthermore, since the accelerometers on either CHAMP, GRACE and GOCE are manufactured by the same manufacturer and probably will have very similar error characteristics, it remains a question whether such a comparison is really "calibration" or more a consistency check. Alternatively, one could investigate the possibility of using the CHAMP or GRACE data (hi-lo SST and lo-lo SST respectively) directly as external data source, convert them to GOCE type SGG and SST data at the respective altitude, and compare with the GOCE measurements. Items to be addressed here are for example: the calibration of GRACE and CHAMP; spectral error characteristics of GRACE and CHAMP compared to GOCE; methods for converting the external data into GOCE type SGG and SST data (including interpolation, filtering, etc.).

4. Inter-comparison between different types of GOCE data One may think of all kinds of methods to inter-compare data from different GOCE sensors and actuators. The main questions with any of such methods concern the in-dependency of these data types and the risk of circular reasoning. Possible methods could be:

a. *SST – SGG*: use the GOCE hi-lo SST data to check the GOCE SGG data, more or less in the same way as with the GRACE and/or CHAMP data as described above.

b. *common mode – SST*: check the gradiometer common mode output with the SST/GPS observations. Issues here are, for instance, the accuracy of the common mode observations, the possibility of cross-interference and aliasing when GPS data is used for both gravity field recovery and checking common mode accelerometer observations, the DFC performance, the accuracy of non-conservative force models, the choice of the a priori gravity field model, etc.

c. *common mode – DFC/atmosphere*: compare the common mode accelerations with information of DFC (telemetry) plus atmospheric models. Issues here are, for instance, the accuracy and resolution of the atmospheric models, the spectral characteristics of DFC, etc.

d. *SST – SLR*: compare the SST orbit data with SLR observations. Although a passive instrument and not an actuator, the SLR retroreflector enables a check of GPS-based orbit solutions. Of course the density of the SLR network will not be enough to ensure by itself a good orbit coverage. Nevertheless, a fit of the SLR observations to the SST-based GOCE orbit solution will give some indication on the accuracy of this orbit and indirectly on the quality of the SST observations, especially at the beginning of the mission.

e. *SGG – star tracker*: the angular motion Ω can be determined by integration from the anti-symmetric part of the measurement matrix. This could be compared to star tracker observations, either by integrating Ω once again or by (numerically) differentiating the angles from the star trackers. Issues here are: the sampling rates of the gradiometer

compared to the star tracker, accuracy of the off-diagonal components (only Γ_{xz} is measured accurately) and of the star tracker observations, calibration of the star tracker observations, etc.
f. *SGG data redundancy*: Here, redundancy is to be understood in the sense of cross-overs and repeated ground tracks. The former are always present, the latter depend on the mission design, but a repeat period of approximately 2 months has been mentioned for GOCE. First of all, overlapping measurements (i.e. measurements having the same location w.r.t. the Earth) give interesting possibilities for quick look data control and outlier detection. Furthermore, if one subset of the data is well calibrated externally, subsequent subsets may be calibrated w.r.t. the first set, like it can be done for subsequent repeat cycles. This is, however, more a kind of "relative" calibration.

End-to-End simulator All the methods described above should be investigated in much more detail to prove their value for the actual calibration once GOCE flies. Obviously this should be done prior to the mission and the discussion here is a first attempt for a qualitative assessment of the methods. Obviously, for a proper quantitative and qualitative assessment we need to do simulation studies. Here, realistic GOCE data time series including errors are necessary. Such realistic time series should come as output of an End-to-End simulator where the GOCE instruments, satellite and measurement process are simulated in detail, including the possibility to inject any kind of (systematic or not) error for which a sensitivity study has to be performed, e.g. [5]. Such an End-to-end simulator forms a necessary tool for testing calibration procedures and for doing sensitivity studies for all kinds of measurement errors.

CONCLUSIONS

A challenging mission like GOCE puts high demands on (external) calibration and on quality assessment in general. We have proposed that a clear terminology is used, where cal/val terms are linked to the GOCE data levels. Hence we use the term "calibration" on level 1b and "validation" on level 2. Calibration for GOCE will consist of a multi-step procedure: pre-flight calibration on ground before the mission, in-flight internal ("relative") calibration on-board and post-mission external ("absolute") calibration. The internal calibration step applies to the accelerometers and the gradiometer but not to SST. Many methods for GOCE external calibration are suggested, although a lot of investigations are still necessary and an assessment of all methods should be done. This can be done with support of the End-to-End simulator, the output of which can be used for studying the feasibility and sensitivity of calibration methods and testing the implementation of these methods. In the end, the whole cal/val process for GOCE will have to be done iteratively, coming back to level 1b after validation to improve the quality even further. Several methods relating to external calibration are under investigation. Many open questions with all these methods remain, concerning e.g. the selection and quality of the external data, correlations (or dependency) if data of different GOCE sensors is inter-compared, "absoluteness" of methods, differences in error characteristics of the GOCE data compared to the external data, calibration of the error-estimates, etc.

References

[1] A. Albertella, F. Migliaccio, F. Sansó and C.C. Tscherning, "The space-wise approach - Overall scientific data strategy," in *From Eötvös to Milligal*, Final Report, H. Sünkel, Ed. ESA/ESTEC Contract No. 13392/98/NL/GD, pp. 157-176, 2000

[2] D. Arabelos and C.C. Tscherning, "Calibration of satellite gradiometer data aided by ground gravity data," *J. Geodesy*, 72(11), pp. 617-625, 1998

[3] D. Arabelos and C.C. Tscherning, "Gravity Field Recovery from Airborne Gradiometer Data Using Collocation and Taking Into Account Correlated Errors," *Phys. Chem. Earth (A)*, 24(1), pp. 19-25, 1999

[4] ESA (1999), Gravity Field and Steady-State Ocean Circulation Mission, *Reports for Mission Selection, The Four Candidate Earth Explorer Core Missions, SP-1233(1)*, European Space Agency, July 1999.

[5] A. Selig, R. Koop, P. Visser, J. Müller and N. Sneeuw, "End-to-End Closed-Loop Simulation Facility for GOCE," Proposal PR-G-001SR/97, Issue 2, 7 January 1998

[6] P. Visser, R. Koop and R. Klees, "Scientific data production quality assessment," in *From Eötvös to Milligal*, Final Report, H. Sünkel, Ed. ESA/ESTEC Contract No. 13392/98/NL/GD, pp. 157-176, 2000