Geophysical Journal International

Geophys. J. Int. (2012) **189,** 877–897 doi: 10.1111/j.1365-246X.2012.5404.x

A comparison of the gravity field over Central Europe from superconducting gravimeters, GRACE and global hydrological models, using EOF analysis

David Crossley,[1] Caroline de Linage,[2] Jacques Hinderer,[3] Jean-Paul Boy[3,4] and James Famiglietti[2,5]

[1] *Department of Earth and Atmospheric Sciences, St. Louis University,* MO 63108*, USA. E-mail: crossley@eas.slu.edu*
[2] *Department of Earth System Science, University of California Irvine,* CA 92697*, USA*
[3] *Ecole et Observatoire des Sciences de la Terre, University of Strasbourg, CNRS, France*
[4] *NASA Goddard Space Flight Center, Greenbelt, MD, USA*
[5] *UC Center for Hydrological Modeling, University of California, Irvine, CA, USA*

Accepted 2012 January 30. Received 2012 January 30; in original form 2011 February 4

SUMMARY

We analyse data from seven superconducting gravimeter (SG) stations in Europe from 2002 to 2007 from the Global Geodynamics Project (GGP) and compare seasonal variations with data from GRACE and several global hydrological models—GLDAS, WGHM and ERA-Interim. Our technique is empirical orthogonal function (EOF) decomposition of the fields that allows for the inherent incompatibility of length scales between ground and satellite observations. GGP stations below the ground surface pose a problem because part of the attraction from soil moisture comes from above the gravimeter, and this gives rise to a complex (mixed) gravity response. The first principle component (PC) of the EOF decomposition is the main indicator for comparing the fields, although for some of the series it accounts for only about 50 per cent of the variance reduction. PCs for GRACE solutions RL04 from CSR and GFZ are filtered with a cosine taper (degrees 20–40) and a Gaussian window (350 km). Significant differences are evident between GRACE solutions from different groups and filters, though they all agree reasonably well with the global hydrological models for the predominantly seasonal signal. We estimate the first PC at 10-d sampling to be accurate to 1 μGal for GGP data, 1.5 μGal for GRACE data and 1 μGal between the three global hydrological models. Within these limits the CNES/GRGS solution and ground GGP data agree at the 79 per cent level, and better when the GGP solution is restricted to the three above-ground stations. The major limitation on the GGP side comes from the water mass distribution surrounding the underground instruments that leads to a complex gravity effect. To solve this we propose a method for correcting the SG residual gravity series for the effects of soil moisture above the station.

Key words: Time series analysis; Satellite gravity; Time variable gravity.

1 INTRODUCTION

Satellite gravity from the Gravity Recovery and Climate Experiment (GRACE) solutions relies heavily on the use of models to remove many of the known effects on gravity variations, and much effort has gone into fine tuning the data processing and updating the models to improve the GRACE products (e.g. Wahr *et al.* 2004, 2006). The task of validation has largely rested on making comparisons with large-scale hydrological models and ground-level data linked to water balance (e.g. Swenson *et al.* 2003). Such studies are in a sense the ultimate ground truth because the GRACE mission was designed for hydrology interpretation, but they are not equivalent to a direct comparison of GRACE solutions with ground gravity measurements.

Several studies have used regional hydrological data over large areas in the United States to test GRACE solutions. Soil moisture data from Illinois confirms that seasonal variations observed by GRACE are similar to those on the ground (Swenson *et al.* 2006), with similar results for soil moisture in Oklahoma (Swenson *et al.* 2008). A study over the High Plains Aquifer (southwestern United States) used observed groundwater levels and a soil moisture model from the Global Land Data Assimilation System (GLDAS) hydrology model (Rodell *et al.* 2004), to confirm GRACE predictions (Longuevergne *et al.* 2010). Wang *et al.* (2011) show comparisons between GRACE and the *in situ* water data in the Three Gorges Reservoir.

Not all GRACE validation has relied on soil moisture or groundwater *in situ* data. The closest analogy to ground gravity has been

the use of ocean bottom pressure (OBP) gauges. Several studies have shown correlations between time-series in linear arrays of two to three instruments, over a distance of 1000 km or so, that agree with GRACE predictions (e.g. Kanzow *et al.* 2005; Rietbroek *et al.* 2009). Park *et al.* (2008) compared a large number of OBP gauges (46) with GRACE over an area 600 km × 600 km off the east coast of Japan. They did not use the EOF technique, and seemed to take a mean over their grid as their spatial average.

The time-variable part of the GRACE gravity field is difficult to validate with ground gravity because the highest accuracy of the GRACE field, a few microGal (1 μGal = 10 nm s^{-2} = 10^{-8} m s^{-2}), is only achieved by averaging over distances of 400 km or more, for example, Swenson *et al.* (2006, 2008). The time sampling is either every 10 d or monthly. These requirements cannot be met with conventional spring gravimeters, because the irregular drift for even the newer instruments (e.g. the Scintrex CG5 or microG gPhone) is combined with an accuracy of 5–10 μGal (Riccardi *et al.* 2011). Absolute gravimeters (AGs) have a limited accuracy (1–2 μGal), but their irregular and sparse time sampling—at least for normal observational campaigns—leads to aliasing of several important effects at weekly and monthly timescales. The only suitable gravity instrument is the superconducting gravimeter (SG) whose data are recorded for the Global Geodynamics Project (GGP) network (Crossley & Hinderer 2009; http://www.eas.slu.edu/GGP/ggphome.html).

We are interested in comparisons at the resolving limit (400 km) of GRACE solutions, and down to shorter time scales (10 d) than usually considered. Central Europe is the only place with seven SG stations close enough together to compare with the GRACE footprint. We here use one version of the empirical orthogonal function (EOF) method to provide a seasonal variation of gravity that is to be compared with GRACE data. This was first done by Crossley *et al.* (2004), and then Neumeyer *et al.* (2008) who used the same computer algorithm.

Our previous studies used techniques similar to those here (Crossley *et al.* 2005, 2006; Hinderer *et al.* 2006). We here extend the common time period to 6 yr (2002–2007) and consider four versions of the CSR/GFZ RL04 monthly solutions and the 10-d CNES/GRGS RL02 solutions. Although we have done an EOF analysis on GGP data for the whole period 1997–2007, the main focus is on GRACE comparisons since 2002. Significant improvements here are the use of station EOFs for GGP data, inclusion of different hydrological models (GLDAS-NOAH, WGHM and ERA-Interim) and an explicit treatment of error estimates in all the EOF variables from GRACE, global hydrological and GGP data.

We here introduce the abbreviations for the institutions providing the online GRACE fields: Center for Space Research, U. Texas (CSR), GeoForschungsZentrum, Potsdam (GFZ) and Groupe de Recherche en Géodésie Spatiale at the Centre National d'Etudes Spatiales, in Toulouse, (GRGS/CNES). We also identify the hydrological models: GLDAS, the Water Gap Hydrology Model (WGHM) and the European Center for Medium Range Weather Forecasting (ECMWF) whose hydrology model is ERA-Interm.

2 METHODOLOGY

2.1 Properties of ground and satellite data

SG measurements have high accuracy (better than 0.1 μGal), high sampling (1 s) and also include vertical ground motion (e.g. Van Camp *et al.* 2005; Hinderer *et al.* 2007; Klügel & Wziontek 2009; Wziontek *et al.* 2009a). They are not time aliased, but spatial aliasing is a problem because only a few SGs are available, and a small number of them at separations less than tens of kilometres (Wuhan, Taipei). GRACE solutions are time aliased due to the undersampling of short-period phenomena (e.g. tides) that must be modelled and removed (e.g. Seo & Wilson 2005). On the other hand GRACE data are unaffected by vertical motion of the ground. Both ground and satellite data need to be consistently processed for the removal of known processes (tides, Earth rotation, oceanic and atmospheric effects) before they can be compared to each other or with ground-based hydrology.

GRACE field products from CSR and GFZ need to be filtered by the user to suppress the higher degree (and noisier) harmonics of the field expansion. By comparison, in the GRGS-constrained solutions, filtering occurs at the same time as estimating the coefficients. Several methods for filtering, or localizing, the field at ground level have been used in GRACE processing (e.g. Han *et al.* 2005), and they confirm that the GRACE fields are stable at length scales between 400 and 1000 km (harmonic degrees 50–20), depending on the type of averaging. It is not possible to provide GRACE solutions at points on the ground (Wahr, personal communication, 2001) because instabilities arise from the higher harmonics when constructing solutions at a point so far below the satellite altitude (see Wahr *et al.* 1998, for discussion of the averaging). This is a fundamental problem in comparing a single ground gravity measurement with GRACE data.

Velicogna & Wahr (2001) raised a concern on the use of ground gravity based on their analysis of the spatial coherence of worldwide SG residual data series, using uncorrected data from the GGP database, indicating it was insufficient for the requirement for ground–satellite comparisons. Certainly the irregular and sparse spacing of SG stations cannot address hydrological variability in the 1–100 km length scale over large continental areas, and this is unlikely to change by future deployment of instruments. Nevertheless, a small area such as Central Europe should be amenable to the EOF method of extracting common modes data from both GGP and GRACE.

Despite the incompatibility of scales, GRACE data are sometimes shown as time-series at specific locations on the ground, and compared directly with ground-based measurements. Examples include Neumeyer *et al.* (2006, 2008) and Weise *et al.* (2009) for the European group of stations, and Lambert *et al.* (2009) for AG data and river flow rate at widely separated sites in Canada and the northern United States. Despite the seemingly good correlation, a justification for connecting data at such different length scales is surely required. This is rarely done, and instead of finding a mean averaged gravity at the GRACE resolution (that would require an accuracy of a μGal or so), it is assumed the gravity is the same everywhere over such an area.

2.2 Choice of ground stations

The European group of GGP stations extends from Medicina (Northern Italy) to Ny-Alesund (Svalbard, Norway), for example, Crossley *et al.* (2006). To investigate continental hydrology we need to select stations that are connected to the region being considered, and Ny-Alesund is too far removed from Central Europe in this regard.

Previously we included Metsahovi (Finland), but found only weak coherency with Central Europe (Crossley *et al.* 2003, 2005). Metsahovi is about 1000 km from the middle of the main array, and there

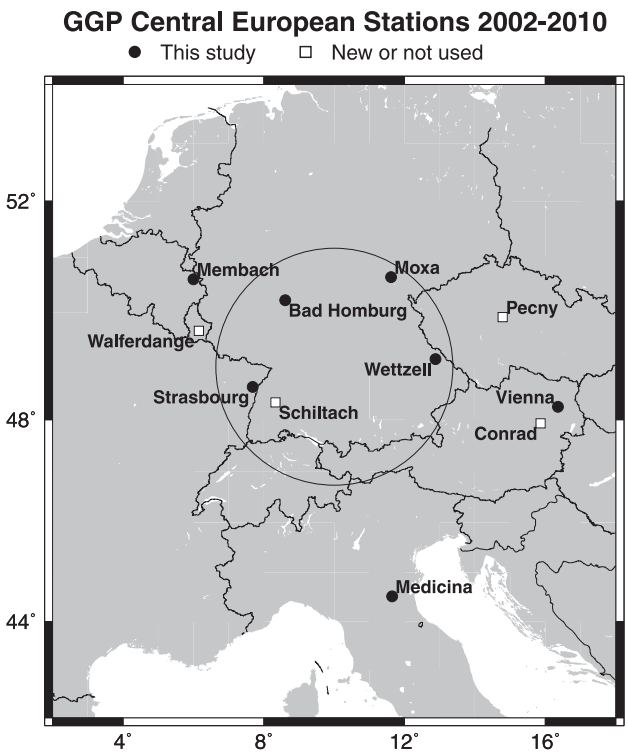

Figure 1. The Central European GGP subarray. We used stations with filled black circles; Pecny, Schiltach and Conrad had insufficient data for our period of study. The circle represents a GRACE footprint of 500 km.

are no supporting SG measurements between the two areas. Further, the Baltic Sea poses special problems because of the added mass loading and the leakage of signal between land and sea (Virtanen & Makinen 2003; Virtanen 2004). We therefore omitted Metsahovi from this study, even though it is a surface station with well-studied hydrology and ocean loading effects (Hokkanen *et al.* 2006,Virtanen *et al.* 2006).

We use the same stations as in Crossley *et al.* (2006)—shown in Fig. 1. These will be referred to by their abbreviations: Bad Homburg (BH), Membach (MB), Medicina (MC), Moxa (MO), Strasbourg (ST), Vienna (VI) and Wettzell (WE). Unlike Neumeyer *et al.* (2008), we keep the underground stations such as ST, MB and VI, arguing instead that their time-series contribute usefully to the EOF solutions. Neumeyer *et al.* (2008) included both Metsahovi and MO, but MO has the least impact on the EOF solutions. Recent additions to the GGP network in Central Europe are Pecny (Pálinkás 2009), Conrad (Meurers 2010) and Schiltach (Widmer-Schnidrig *et al.* 2010) that can be used in future studies.

2.3 Treatment of near-station water mass distribution

The area surrounding a ground gravity station can be separated into three zones, with local (L), regional and global contributions, for example, Virtanen *et al.* (2006). Once the larger signals (such as tides, atmospheric pressure and polar motion) are removed, SG residuals show a seasonal variation due mostly to soil moisture and groundwater, and this is dominated by sources close to the gravimeter, in the local zone, L (meaning within about 1 km of the station). The extent of this zone (also called the near-zone, Wziontek *et al.* 2009b) depends on the topography and soil properties for each station. Although Creutzfeldt *et al.* (2008) showed that 90 per cent of the hydrology signal can be generated within a few hundred kilometres of the station, some studies have found that most of the local effect comes from a radius of 50–150 m (Hokkanen *et al.* 2006; Hasan *et al.* 2008), although investigations around MO considered a larger radius of 1 km (Kroner & Jahr 2006; Kroner *et al.* 2007). All the water transfer processes that take place within this zone are 'local hydrology', but these are not modelled in the present paper. The only hydrological simulation we consider here arises from the global hydrological models that are computed as gravity effects (attraction + loading, in m s^{-2}); this is treated in the next section.

A valid question is whether a correction for near-station water distribution should be done before comparing ground gravity to GRACE data (e.g. Neumeyer *et al.* 2004; Weise *et al.* 2009). When soil moisture or groundwater levels are partially correlated with gravity data they may be considered as a proxy for the local gravity effect. In this case an admittance factor may be used to remove the local effect (if desired), but this will remove all gravity effects from water mass variations (both local and non-local) that are correlated with the gravity data. Neumeyer *et al.* (2006) used this approach to correct residual gravity using groundwater data, finding an admittance between 1 and 10 μGal m^{-1} (water), depending on the hydrological conditions.

In detailed station studies, a better approach is to construct a physical model that could involve soil moisture metres, flow metres, auxiliary gravity measurements and measurements of the physical properties of the soil layers (e.g. conductivity, or porosity). Examples can be found for most of the SG sites here, for example: Naujoks *et al.* (2008, 2010) and Hasan *et al.* (2008) for MO; Van Camp *et al.* (2006) for MB; Creutzfeldt *et al.* (2008, 2010a,b) for WE and Llubes *et al.* (2004) and Longuevergne *et al.* (2009) for ST.

Here we make no corrections for near-station water mass attraction because we are computing a spatial average (i.e. EOF decomposition) that requires the full signal at every surface site. We assume a station is representative of the area around it, and does not require detailed corrections for effects very close to the SG (such as buildings, or vegetation). Although three of our SGs are above ground, and can be used uncorrected, four are buried, sitting between the surface soil moisture (above) and the groundwater (below). In the latter case we refer to the signal as 'mixed', meaning the gravity attractions from above and below the SG oppose each other. In this case one should make a correction to remove the effects of mass variations between the SG and ground level, a topic treated later in the paper.

3 GLOBAL HYDROLOGICAL MODELS

For global hydrological models, the entire continental surface is divided into cells for computing the amount of moisture in the near surface from meteorological forcing—that is, precipitation (rainfall, snowfall) and energy fluxes. These models are calibrated against observed data and are widely used for GRACE comparisons, for example, the Land Dynamics Model (LaD) introduced by Milly & Shmakin (2002a,b), and the GLDAS model from Rodell *et al.* (2004).

Our primary model is GLDAS with the NOAH land-surface (soil moisture and snowfall) model (henceforth just GLDAS). The hydrological storage components not explicitly included are groundwater (water tables) and surface water (lakes, rivers, streams, wetlands and floodplains). The model provides estimates of water storage in four layers within the first 2 m of the ground (soil moisture), at

the surface (snow cover) and in vegetation (canopy water—almost negligible). We used portions of these data, available from 2000 February to the present, available originally at 3 hr intervals. The highest GLDAS resolution for Europe is 0.25°, a grid cell 18 km (E–W) by 25 km (N–S) at mid-latitudes.

The output can be used directly for comparison with GRACE if the latter solutions are expressed in equivalent centimetres of water (Wahr *et al.* 1998). The mean field is subtracted to be consistent with GRACE that cannot retrieve absolute storage. To compute the ground gravity, the water in a surface layer generates a gravity effect due both to loading of the elastic earth model and to Newtonian attraction. The loading is computed by convolving a point mass load on a realistic earth model (e.g. PREM) with the hydrological input to each cell of the model (Farrell 1972). For the Newtonian attraction, the location of the soil moisture with respect to the gravimeter must be correctly assessed. The attraction is separated into two contributions. The first is local-cell component (which is not the near-zone effect referred to in Section 2.3) for the model grid cell in which the station is located (this is effectively a Bouguer plate effect with the equivalent water height in the cell as thickness) and the second is a non-local (regional and global) contribution from the other cells. The latter arises from the Earth's sphericity; a more detailed description may be found in de Linage *et al.* (2009) and Pfeffer *et al.* (2011).

The three contributions—loading, local-cell Newtonian attraction and non-local Newtonian attraction—are shown in Fig. 2 for the seven stations from 2002 May to 2007 December. It is assumed the gravimeter is above all soil moisture and groundwater sources, and that site variability is derived entirely from meteorology and the cell soil properties, ignoring topography. The local-cell attraction (5–10 μGal) dominates, as expected, and the deformation and non-local attraction are similar at about 1–2 μGal. To compare directly with local gravity, the sign of the local-cell attraction should be reversed if the station is below the ground surface (e.g. Boy & Hinderer 2006). This would be applicable for stations MB, ST and VI, but is fully justified only when the soil moisture is completely above the SG. Station MO is underground in a hillside, thus partly below the surface soil moisture. At the remaining three sites, BH, MC and WE, almost all the soil moisture is at the surface, that is, below, the SG. In most cases groundwater is beneath the SG, thus leading to a mixed signal at below-ground stations.

From the data in Fig. 2 we compute the ratio of (local-cell attraction)/(total gravity) for each station in the form of a histogram (not shown) of daily ratios which peaks at values between 0.75 and 0.80, except for MC (0.85). We conclude that 75–85 per cent of the estimated ground gravity arises from the local cell (within about 20 km), and the rest comes from global water storage (e.g. Llubes *et al.* 2004). The regional attraction (1–1000 km) is generally small because it is lateral with respect to the station.

We also compare GLDAS with two other models—WGHM given by Döll *et al.* (2003), and an interim version of a hydrological model based on ECMWF, called ERA-Interim (Dee *et al.* 2011). WGHM solutions are internally based on daily data, but available monthly on a 0.5° grid, as used for example by Wziontek *et al.* (2009b) in their comparison of the same three above-ground stations as we use here. A notable feature of the WGHM model is its extensive treatment of surface water (lakes, rivers and wetlands) and groundwater, and it is calibrated against a large number of global river discharge measurements.

A new product from ECMWF is available (Dee *et al.* 2011); called ERA-Interim, this is a comprehensive atmospheric model based strongly on data reanalysis and assimilation on a 12 hr cycle. The ERA-Interim solutions are produced every 6 hr on a 0.5° grid and include four layers up to 2.55 m, and snow. As with the GLDAS model, we resample this series to 10-d GRGS epochs (see later).

A comparison of these three global hydrological models over the study area is shown in Fig. 3 for the climate extremes in winter and summer 2003. GLDAS has the highest seasonal variability and resolution but does not show coherency over the region. By contrast ERA-Interim has a clear signal over the Alps due to snow accumulation, but otherwise shows almost no water storage. WGHM is between these two extremes, both in amplitude and in variability.

At a global scale, a cell of 0.25° is sufficient for large-scale continental hydrology, but it is too large to adequately characterize effects (within a kilometre or so of the station) due to near-station topography and small-scale variations in water storage. Thus, as

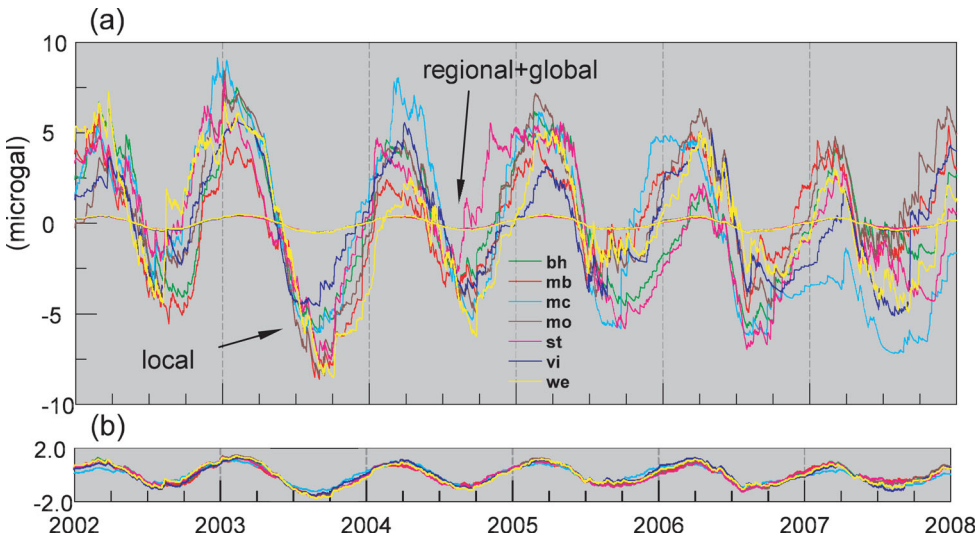

Figure 2. The three components of surface gravity (in μGal) for GLDAS hydrology. Contributions to attraction (a) are a local delta function (from the cell including the station) and a non-local sum that includes both regional and global sources. The loading deformation (b) is similar in amplitude and phase to the non-local attraction.

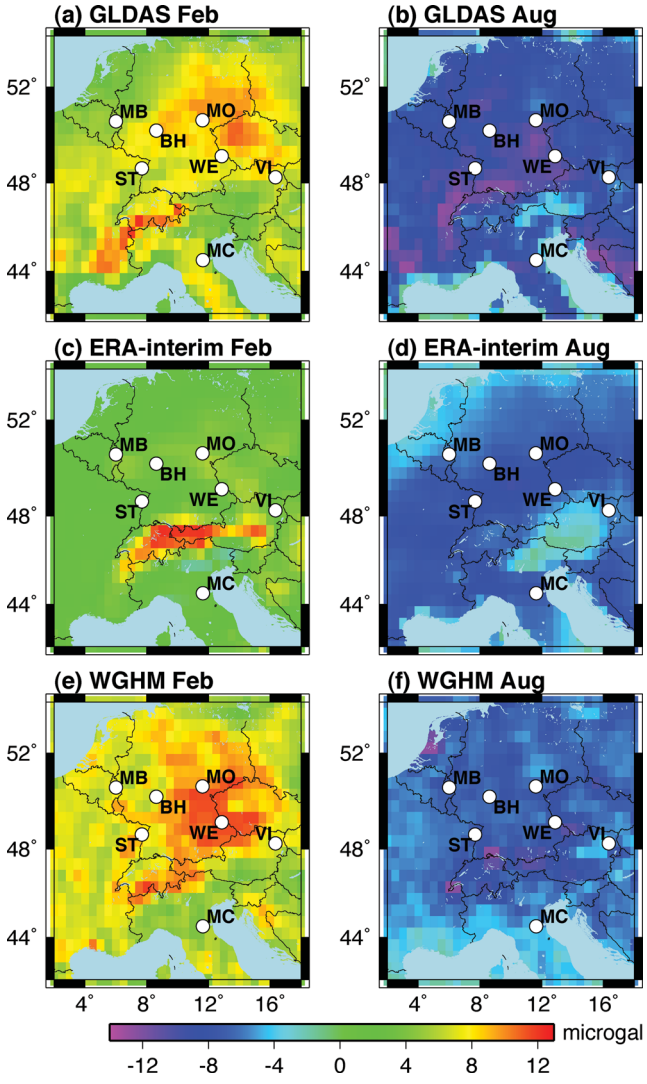

Figure 3. Comparison of total gravity from global hydrological models for 2003 in February (2003.14) and August (2003.68): (a) and (b) for GLDAS, (c) and (d) for ERA-Interim and (e) and (f) for WGHM. The times coincide with the first two vertical dashed lines in Fig. 4.

noted by Longuevergne *et al.* (2009), global hydrological models do not capture detailed near-station water mass changes, but they are important for defining the non-local attraction and loading.

4 GRACE DATA

We use spherical harmonic coefficients provided by GRACE from CSR (Texas) and GFZ (Potsdam) for Level 2 solutions, Release 04 (Bettadpur 2007; Flechtner 2007), from 2002 April to 2007 December. Several months are missing (2002 May–July, 2003 June), due to problems early in the GRACE mission. For the CSR solution, the mean field from 2002 August–2007 August was subtracted, leaving 65 data sets, centred at the 15th day of each month. The GFZ time-series over the same period contains 60 acceptable solutions (a different criteria from CSR for processing and rejection) and the mean field was subtracted.

Two filters are used for the RL04 monthly fields, described in de Linage (2008, fig. 1.13, p. 80). A spherical harmonic of degree ℓ has a half wavelength $\lambda_{1/2} = \pi R/\ell$, which is the resolution on a sphere of radius R. One filter is a cosine taper (CT20-40) applied between degrees 20 and 40, preserving degrees up to 20 but suppressing completely those higher than 40. The second filter is the widely used Gaussian window (Jekeli 1981) that is smooth in both the wavenumber and space domains. We use it with a nominal width of 350 km (GF350) that smoothly damps degrees between 18 and 57 ($\lambda_{1/2} = 1100$, 350 km), thus preserving smaller wavelengths than CT20-40 (de Linage 2008, fig. 1.14). We refer to these RL04 fields as CSRct, CSRgf, GFZct and GFZgf with obvious meanings. Previous experience showed that a filter CT30-50 leaves noisy residuals, and GF500 with 500 km radius is too smooth. We do not use a de-striping algorithm (e.g. Swenson & Wahr 2006) due to the small size of the ground area.

We also use GRACE solutions from the GRGS/CNES in Toulouse (Biancale *et al.* 2010; Bruinsma *et al.* 2010). The GRGS time-series starts in 2002 August and we subtract the mean field. Release 02 is based solely on 10 d of GRACE data, with no smoothing using adjacent 10-d windows (as previously). The solutions are forced by a wavelength-based constraint towards the dynamical reference mean field at each epoch, and further filtering is unnecessary. There are 187 sets of 10-d solutions, omitting months with gaps and poor data. The claimed errors in GRGS limit its resolution to about 400 km, but this is similar to that of the CSR and GFZ monthly solutions.

Processing for the CSR and GFZ solutions are similar (Bettadpur 2007; Flechtner 2007). Solid Earth tides are computed as in IERS (1992) conventions, ocean tides are computed using CSR4.0 for CSR (Eanes 2002), and FES2004 for GFZ (Lyard *et al.* 2006), and polar motion is removed with IERS data. The CSR and GFZ atmospheric and ocean models are ECMWF 0.5°, 6-hr data vertically integrated and combined with the OMCT baroclinic ocean model for Release 04 solutions.

The CNES/GRGS processing includes models for solid and ocean tides (FES-2004) and 3-D atmospheric pressure from the ECMWF 6-hr solution with HUGO-m, which is an update of the MOG2D barotropic ocean model (Carrère & Lyard 2003). Time-variable gravity field variations due to the atmosphere are computed using a 3-D approach (as described in Boy & Chao 2005), but there remains some differences in processing between CSR/GFZ and GRGS for the 3-D atmospheric attraction. GRACE solutions are uncorrected for hydrology and secular effects such as postglacial rebound.

4.1 Correction for vertical displacement

A satellite responds only to changes in the geopotential, which is derived from the effects of variable mass attraction and loading deformation of the Earth. A surface gravimeter, on the other hand, is sensitive to four effects. These are as follows: (1) the direct vertical acceleration of the instrument which dominates surface gravity at seismic normal mode frequencies (<1 hr), but is negligible for the longer period effects considered here, (2) the direct (Newtonian) effect of the mass redistribution, (3) the vertical displacement of the instrument through the invariable part (to first order) of the gravity field, that is, the free-air gravity (or elevation) effect and (4) the perturbation of the geopotential, in this case due to the loading deformation.

Effects (3) and (4) exist only because the Earth elastically deforms when submitted to a load. Effect (4) is the classical Bouguer slab effect in static gravity surveys, provided that the Earth's crust is incompressible with a plane surface, but departs from this value for a fully compressible, spherical Earth (de Linage *et al.* 2007). Satellites

see only (2)+(4), and gravimeters see all effects. Excluding (1), the difference is (3), the free-air effect of loading deformation, which may be expressed as:

$$\Delta g_{\rm GRND} = \Delta g_{\rm SAT} - (2g_0/a)\,\Delta h, \tag{1}$$

where Δh is the vertical displacement of the station. The value of $(-2g_0/a)$ is the well-known free-air gradient $-0.3086\ \mu$Gal mm^{-1}.

Eq. (1) gives a method for correcting either ground or satellite data for vertical displacement. de Linage *et al.* (2007, 2009) discussed the relationship between ground gravity changes Δg and height variations Δh, evaluating this ratio for a variety of situations and surface loads such as the atmosphere and hydrology. A first estimate of $\Delta g_{\rm GRND}$ from $\Delta g_{\rm SAT}$ could come from putting this ratio into eq. (1), that is, $\Delta g_{\rm GRND} = \Delta g_{\rm SAT}/[1 + (2g_0/a)(\Delta g/\Delta h)^{-1}]$. This approach, however, would be only approximate because the computed ratio is not just a simple scalar admittance but depends on the spatial wavelength.

We used instead an alternative approach to estimate the vertical displacement caused by the height-induced loading that would be observed on a standard elastic earth model using the water mass redistribution from GRACE, following Neumeyer *et al.* (2006). The usual GRACE gravity is the vertical derivative of the satellite potential at radius *r:*

$$\Delta g_{\rm SAT}(r,\theta,\lambda) = \frac{GM}{ra}\sum_{n=2}^{n_{\max}}\left(\frac{a}{r}\right)^{n+1}(n+1)\sum_{m=0}^{n}\left[\Delta C_n^m Y_n^{m,c}(\theta,\lambda) + \Delta S_n^m Y_n^{m,s}(\theta,\lambda)\right] \tag{2}$$

where ΔC_n^m and ΔS_n^m are the variations of the Stokes coefficients provided up to degree $n_{\max}$ and the other symbols have their traditional interpretation. The GRACE ground gravity is the above expression including vertical motion of a sensor, evaluated on a spherical surface touching the reference ellipsoid at $a = 6378$ km:

$$\Delta g_{\rm GRND}(\theta,\lambda) = \frac{GM}{a^2}\sum_{n=2}^{n_{\max}}\left(n+1-\frac{2h'_n}{1+k'_n}\right) \times\sum_{m=0}^{n}\left[\Delta C_n^m Y_n^{m,c}(\theta,\lambda) + \Delta S_n^m Y_n^{m,s}(\theta,\lambda)\right] \tag{3}$$

where h'_n, k'_n are the elastic load Love numbers for PREM. Neumeyer *et al.* (2006) did not take into account the term $(1 + k'_n)$ involving the load Love number k'_n. They therefore neglected the large-scale effect of mass redistribution inside the Earth when inverting for the load from the Stokes coefficients; see de Linage (2008, p. 110) for more details. Neumeyer *et al.* (2006) showed that the height effect has a seasonal component at most European stations with an amplitude of about 1 μGal, which is similar to that inferred from the GPS data by Zerbini *et al.* (2001).

We compared the eqs (2) and (3) for each of the two RL04 solutions—CSR (two filters), GFZ (two filters) and for GRGS solution. From the mean field over the array at each epoch we can find the ratio of the GRND to SAT values. For the GRGS solutions the ratio is 1.34 (similar to Neumeyer *et al.* 2006), but for CSRct and CSRgf we find 1.36 and 1.39, respectively, and for GFZct and GFZgf we find 1.11 and 1.08, respectively. Hence, as for the gravity/height effect, this ratio depends on the spectral content of each field; the GFZ solutions have less energy at low degrees (<20) compared to GRGS and CSR (the higher the ratio, the higher the energy at low degrees). All GRACE fields are computed using (3), and the vertical deformation is also included in the gravity prediction from global hydrological models.

4.2 The mean and seasonal fields

We have two CSR solutions with CT20-40 and GF-350 filtering, the same for the GFZ series and the CNES/GRGS solution. The ground gravity is evaluated over a 0.25° grid between latitudes 42–54°N and longitudes 2–18°E and we take the mean field at each epoch as a rough spatial average. The range varies for each data set as seen in Fig. 4.

The two CSR fields are similar, as are the GFZ fields, but there are differences between these and the GRGS mean field. Note that the GLDAS hydrology has generally a higher amplitude than the GRACE ground fields, particularly in winter 2005 and summer 2006.

The mean field, however, is not a sufficient measure of variability, even over such a smaller area as Central Europe. We show in

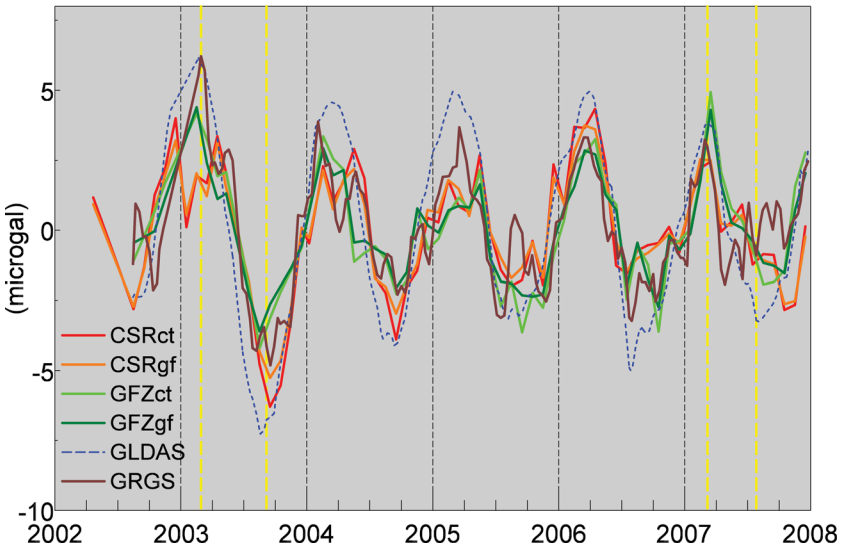

Figure 4. GRACE mean field variations. CSR and GFZ are monthly RL04 solutions; ct, cosine taper; gf, Gaussian filter, and GRGS is the 10-d GRGS solution. GLDAS is given by the dashed curve. The four vertical dashed lines identify the large seasonal extremes from winter to summer in 2003, and the significantly smaller ones in 2007.

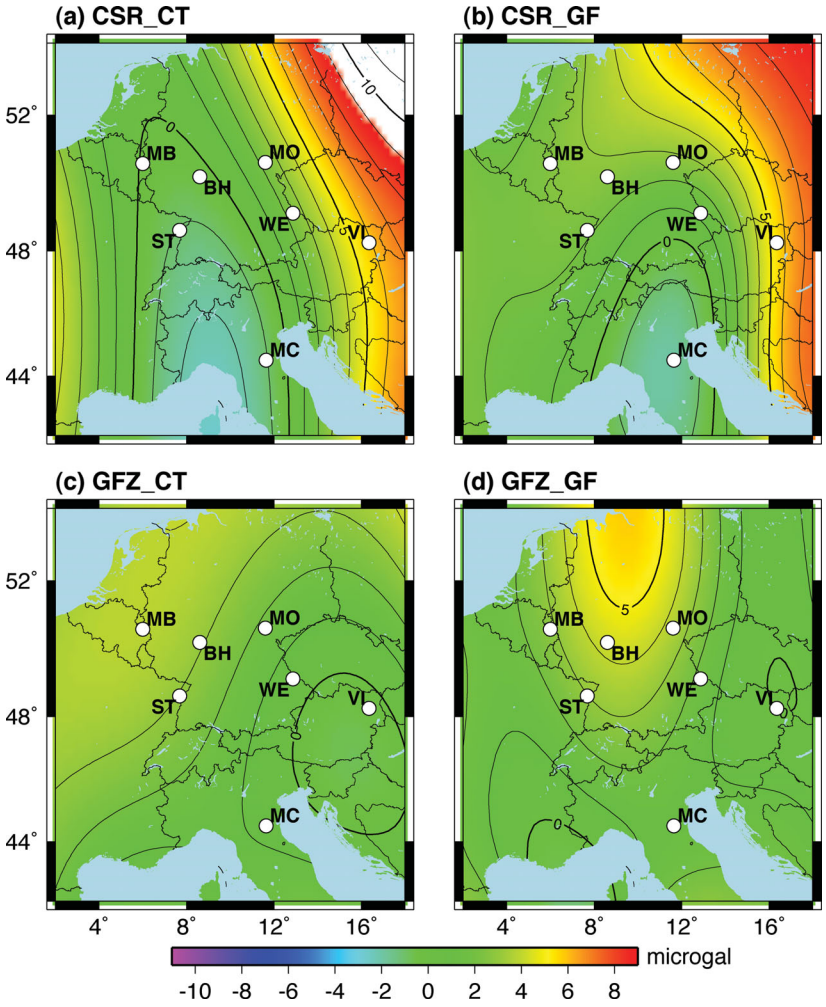

Figure 5. A comparison of GRACE RL04 solutions for February (2007.12). The white area in the first plot just exceeds the limits that are convenient for future comparisons.

Figs 5 and 6 snapshots of the CSR and GFZ fields with both types of filter for the seasonal extremes in 2007 (we avoid the winter of 2003 due to high noise). Differences are most notable between CSR and GFZ for the winter of 2007, otherwise the solutions are quite similar.

The GRGS fields are shown in Fig. 7 for the seasonal extremes both in 2003 and 2007, as noted earlier. The winter high and summer low for 2003 are pronounced over Central Europe, the latter being due to the significant drought noted in many previous studies. The winter of 2007 shows a different pattern for GRGS compared to the CSR and GFZ fields earlier.

5 GGP DATA

We use data from the GGP database at ICET/GFZ Potsdam (http://ggp.gfz-potsdam.de/) between 2002 and 2007, uncorrected and decimated to 1 min, with only a few short gaps. Two methods are used for long-period studies, for example, Hinderer *et al.* (2007). One is a tidal analysis of waves up to monthly periods, combined with pressure and polar motion corrections (e.g. ETERNA 3.4, Wenzel 1996; Baytap-G, Tamura *et al.* 1991). Second is a sequential subtraction of known effects—the approach taken here. Important discussions of the SG processing are given in Klügel & Wziontek, (2009), Van Camp *et al.* (2005) and Wziontek *et al.* (2009a).

5.1 Standard corrections

The pressure signal is first corrected for problems, and previous tidal analyses using ETERNA 3.4 are used to find the local gravimetric tidal factors (δ, κ) at each station. We use a tidal potential with 1200 waves (Tamura 1987) and compute the total synthetic tide from the pre-determined gravimetric factors (automatically including ocean loading) for all waves up to and including monthly periods; the Wahr–Dehant elastic parameters are assumed (Dehant 1987). Tidal waves at semi-annual periods and longer are set to the nominal elastic values (1.16, 0.0) to preserve seasonal variations. Atmospheric pressure is then removed with a nominal admittance of -0.3 μGal hPa^{-1} (Warburton & Goodkind 1977), which allows easier identification of instrument offsets. The series are then decimated to 1 hr samples. Polar motion is subtracted using the IERS rotation pole position provided at http://hpiers.obspm.fr/eoppc/eop/eopc04_05/.

5.2 Treatment of offsets and trends

Removal of instrumental offsets is a critical step and constitutes the most labour intensive part of the SG processing (Hinderer *et al.* 2002). We verified our offset detection against those from station

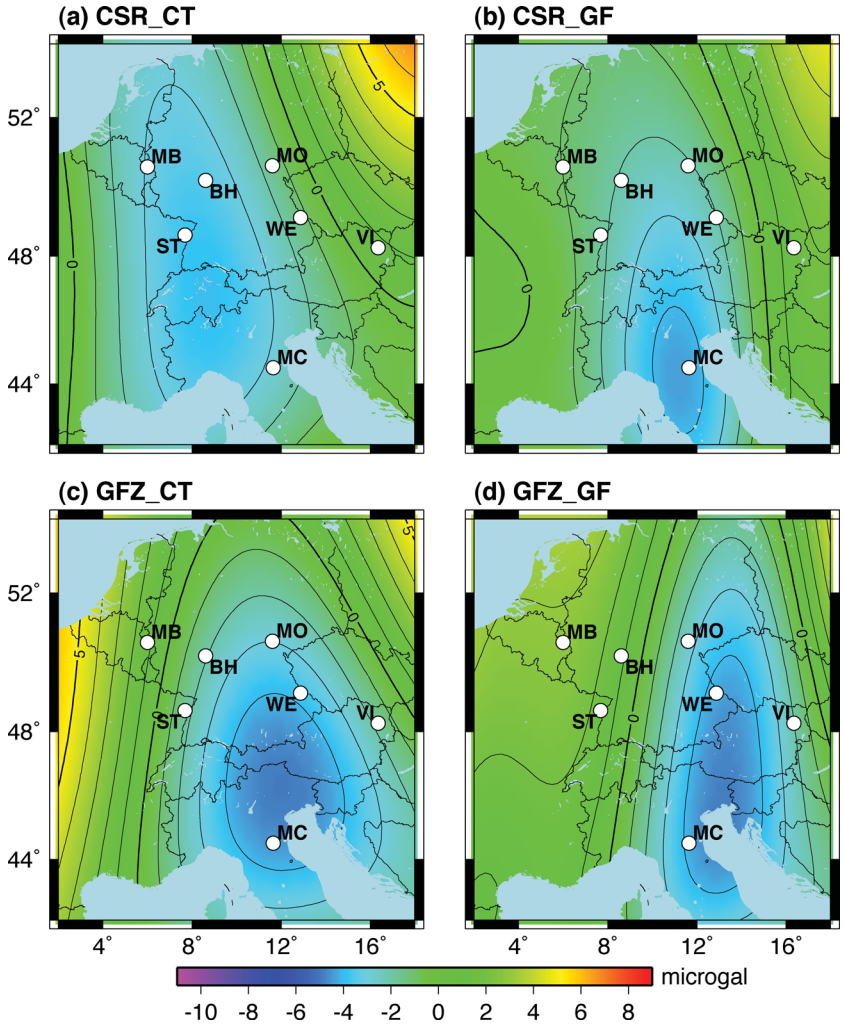

Figure 6. A comparison of GRACE RL04 solutions for July (2007.54).

operators, and use station logs where possible. For these data we removed between 14 and 39 offsets per data set (Table 1).

Offsets are removed along with a linear drift function; this is particularly useful when there are longer data gaps. The linear trend varies between -0.62 and $+3.76$ μGal yr^{-1} (Table 1). For dual-sphere instruments (BH, MO and WE), we average the data from each of the sensors as the difference signal is typically a few tenths of a μGal (Kroner *et al.* 2005). Simultaneous treatment of offsets and instrument drift was also emphasized by Hinderer *et al.* (2002).

5.3 Global pressure corrections

We use three model assumptions: (M1) ECMWF 3-hr global pressure using the inverted barometer (IB) assumption (Boy *et al.* 2002), (M2) ECMWF 3-hr global pressure plus 6-hr HUGO-m non-tidal ocean loading (low resolution, 0.5°, Boy & Lyard 2008) and (M3) ECMWF 3-hr global pressure plus 6-hr HUGO-m (high resolution, 0.25°, Boy *et al.* 2009). These three models differ in the modelling of the ocean's response to surface pressure variations: M1 is static and only valid for periods exceeding several days while M2 and M3 are dynamic (HUGO-m is a barotropic model). We show in Table 2 the standard deviations between these three models at 1 hr sampling for the seven stations, with M1–M3 interpreted as earlier.

There are only small differences M3−M2 between the two HUGO-m computations, but a significant improvement between the HUGO-m model compared to the usual IB assumption (M1). This is clearly seen in the spectrum of gravity residuals at periods of 1–10 d (Fig. 2, Boy *et al.* 2009), indicating the advantage of the HUGO-m model at 10-d sampling (e.g. with the GRACE CNES/GRGS solution). We find that the seasonal variation between the pressure loadings is small compared to the short-term variability, and global pressure is only a small correction to the local admittance already accounted for.

5.4 Corrections for 3-D atmospheric attraction

There is a seasonal warming effect of air masses that is not seen in surface pressure measurements or global corrections. Simon (2002, 2006) found annual variations of about ± 0.6 μGal at the surface with a maximum upward attraction effect in July and a minimum in winter. This is the same sense as the usual annual gravity signal due to soil moisture or groundwater (negative moisture and gravity in summer, the opposite in winter). Zerbini *et al.* (2001) incorporated this correction into the processing of SG data from MC using a fixed annual cosine wave of amplitude ± 0.8 μGal.

Neumeyer *et al.* (2004) recomputed the effect using a more extensive atmospheric model with 50 layers up to an altitude of 60 km and showed that the seasonal effect was indeed smaller than Simon (2002) had first predicted. However, short-period fluctuations can

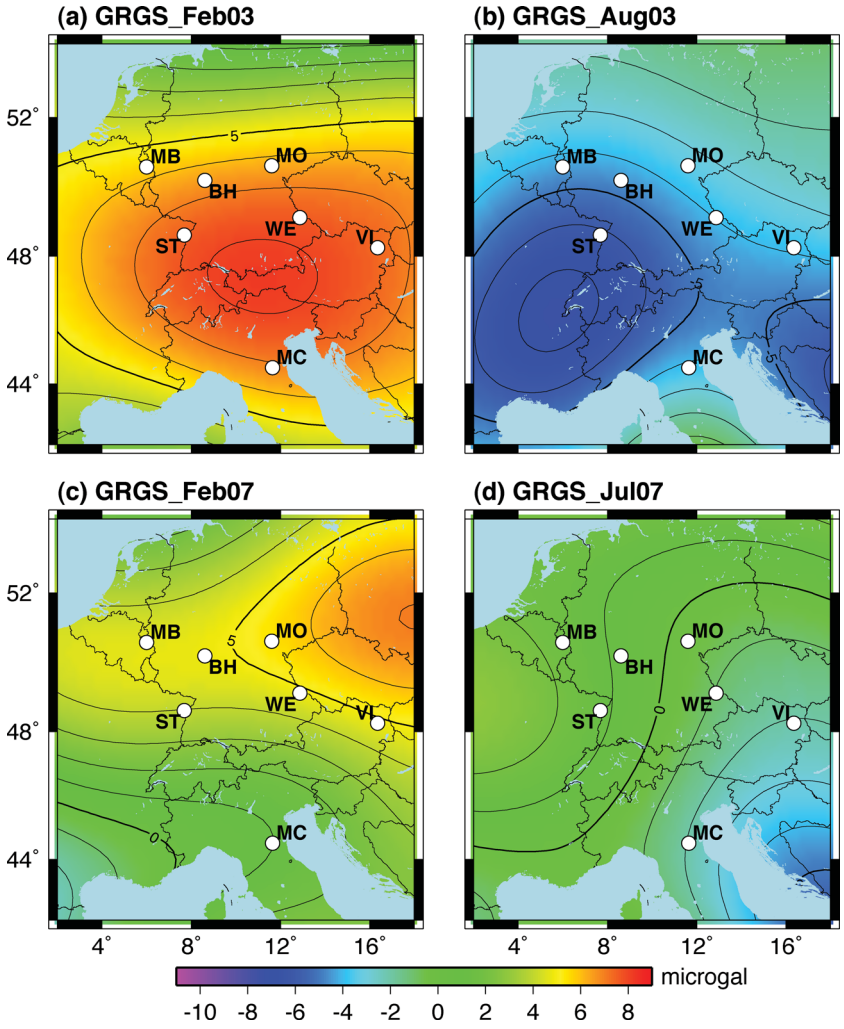

Figure 7. The GRACE–GRGS fields at the times indicated in Fig. 4.

Table 1. GGP station trends and offsets.

Station -sensor	Annual trend (μGal yr^{-1})	Number offsets	Total[a] offsets (μGal)	Accumulated[b] offset (μGal)
BH-h1	0.79	37	1551.3	−75.2
BH-h2	0.82	36	670.9	−16.2
MB	3.37	34	298.6	−35.2
MC	0.78	16	121.8	−21.8
MO-m1	3.76	17	41.4	18.8
MO-m2	1.91	19	108.5	51.4
ST	2.62	14	339.2	59.6
VI	1.55	30	129.1	5.1
WE-w1	0.33	39	953.2	−642.9
WE-w2	−0.62	39	1215.8	−242.0

[a]Total = sum of all offsets disregarding sign.
[b]Accumulated = net effect all offsets on series.

Table 2. Global pressure corrections.

Station	Standard deviation σ (μGal) M3 − local	M3 − M1	M3 − M2
BH	0.541	0.209	0.066
MB	0.417	0.257	0.075
MC	0.342	0.186	0.054
MO	0.373	0.195	0.074
ST	0.531	0.183	0.054
VI	0.308	0.154	0.052
WE	0.341	0.169	0.059

reach several μGal and need to be accounted for in shorter period studies.

We performed a test using four of the seven stations and with the full 3-D computation provided by J. Neumeyer (personal communication, 2005) for years 2002–2004 (the only data we had at the time). The seasonal signal has a maxima/minima of up to 0.4 μGal in summer/winter, respectively, with occasional excursions of 1–2 μGal over shorter periods. Because only the seasonal variation is important, we fitted an annual cosine wave to the full 3-D computation with an overall amplitude of 0.3 μGal and a maximum in August. This is similar to Zerbini *et al.* (2001) but with lower amplitude, and closely fits the 3-D test computations. Such a reduced amplitude is consistent with more recent work by Klügel & Wziontek (2009) and Abe *et al.* (2010) that shows the larger values found in earlier studies were due to an insufficient height ceiling for the atmospheric model.

5.5 Effects not included

For non-tidal seasonal ocean circulation we refer to Boy & Hinderer (2006) and Kroner *et al.* (2009). The latter paper showed

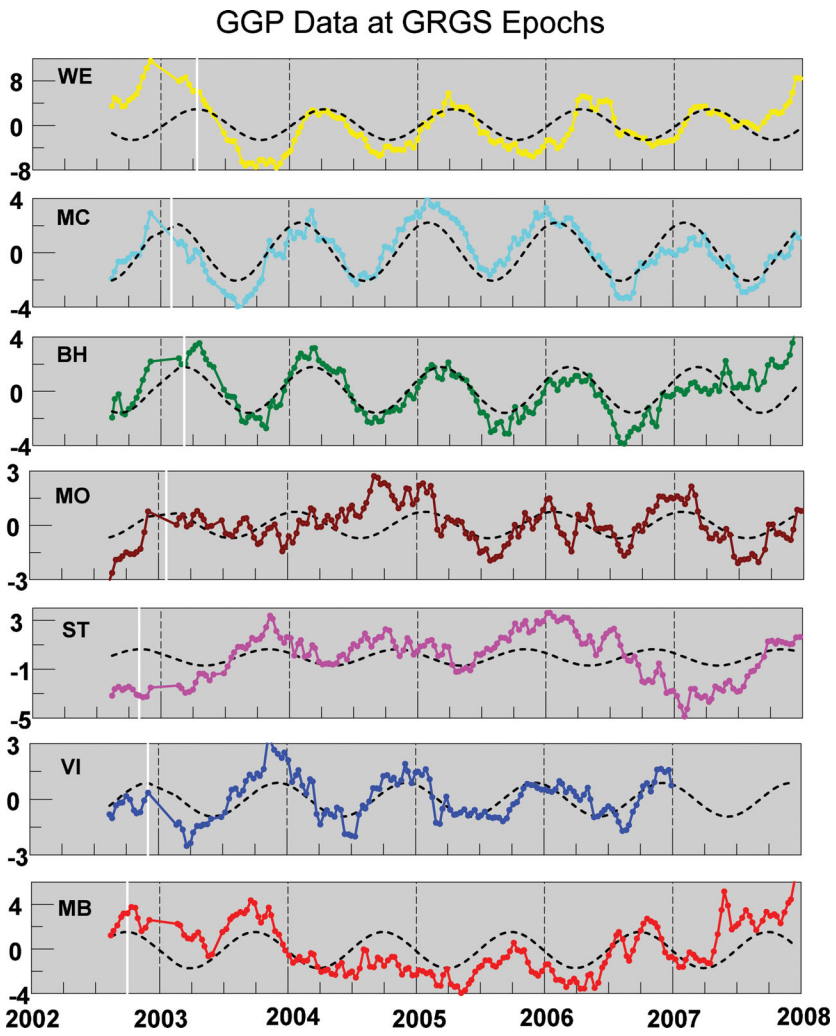

Figure 8. GGP data during the period of the GRACE comparison sampled at GRGS 10-d times. The stations are ordered by increasing depth of burial (in m) from the ground surface above the SG: WE, MC = 0, BH = 2, MO = 6, ST = 6.5, VI = 8 and MB = 48. Also shown are the best fit annual waves (dashed) and their phases (vertical white lines).

that predictions of the OMCT model for station MO gave variations of 0.5–1.0 μGal, but these do not correlate well enough with the observed SG series to significantly reduce the residual gravity. The ocean pole tide and LOD variations are at the level of about 0.15 μGal (Chen *et al.* 2008, 2009), and small compared to the seasonal effects here.

The secular change in gravity over Europe is ignored here as all mean fields are removed. To properly determine the station trends requires AG measurements at each site. Rosat *et al.* (2009), for example, determined in Strasbourg the AG (geodetic) change has been $+0.07 \pm 0.28$ μGal yr^{-1} from 2002 to 2007, and the SG drift (linear) was $+1.20 \pm 0.28$ μGal yr^{-1}. Similarly Zerbini *et al.* (2007) reported for MO a net gravity increase of $+0.72 \pm 0.02$ μGal yr^{-1} for the time period 1998–2005.

5.6 Residual gravity series

We subtract a linear trend from the period 2002 to 2008 (Table 1), and this yields the individual series shown in the Fig. 8. The daily series are resampled with a 10-d Chebyshev filter to the times of the GRGS GRACE solutions and a similar 30-d decimation is done separately to the times of the CSR and GFZ GRACE products.

Fig. 8 shows the GRGS series, ordered top down by the depth of burial of the stations. Note that the three top stations are at, or only slightly below, ground surface, and the rest are significantly below. Also shown are fitted annual cosine waves and their phases; the maxima for the surface stations are somewhat later than the buried stations with a lag reaching 6 months between the deepest (MB) and the shallowest (WE) stations. Further discussion of the annual waves is given in the Supporting Information.

5.7 Single station comparisons

We now show single station comparisons between GLDAS (evaluated at the station), GRACE and GGP data, but this is not our preferred method of comparison, so we use only two examples—stations WE and ST (Fig. 9). For WE, all three fields have similar amplitudes within their errors (see later) and phases, though sometimes the disagreement is striking (especially in 2007). The WE (GGP) series is a good fit to GRACE and GLDAS data which confirms our view that surface stations do not need a significant correction for near-station water mass attraction. We cannot say whether WE is representative of the 400 km area around the station, but this might be inferred from the apparent agreement with

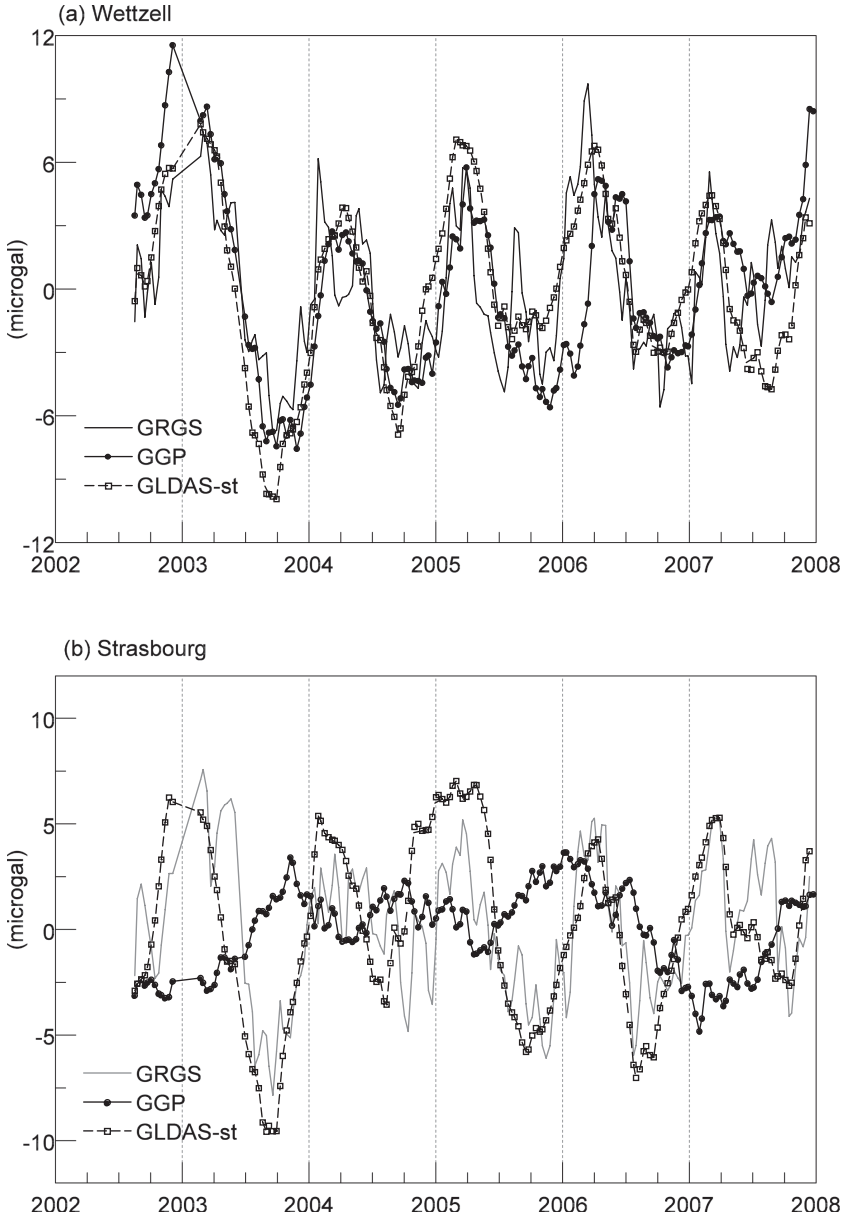

Figure 9. Single station comparisons of GGP, GLDAS-st and GRACE series for (a) WE, showing generally good agreement in amplitude and phase and (b) ST, showing the SG gravity is often anticorrelated with the effect of global hydrology—as expected for an underground station.

GRACE. Comparison between GRACE and a single station rests on a significant assumption of typicality (notwithstanding the errors in the respective data sets).

For ST, the GRACE–GLDAS agreement is again evident (as noted before by Boy & Hinderer 2006; Hinderer *et al.* 2006), but the GGP series is quite different, and well studied (Llubes *et al.* 2004; Longuevergne 2008). It is not simply a question of inverted phase, but the amplitude is smaller as well due to the mixed nature of the attraction (i.e. a thin soil layer above the instrument and groundwater below). This was recognized by Neumeyer *et al.* (2008) as the problem of underground stations.

6 EOF ANALYSIS

EOF is a recognized technique for finding the dominant modes of variability in spatial data sets, and was first used for the GGP-GRACE data sets by Crossley *et al.* (2004). We show in the Supporting Information some properties of the EOFs that are useful in our analysis: (1) the principle components (PCs) can easily be scaled to be in units of μGal (thus giving them a geophysical interpretation), and (2) the phase of the signals from different stations can be opposite without affecting the amplitudes of the PCs.

For the GGP data we use the seven individual station series to form the decomposition, similar to the approach taken by Neumeyer *et al.* (2008), as this introduces no artificial smoothing of the data. This gives PCs that are time-series, but the eigenvectors are only single values and may be gridded to display a 2-D spatial field for each eigenvalue. EOF decomposition gives a diagonal eigenvalue matrix that can be used to find the variance reduction using a certain number of eigenmodes. In all our cases we did not go beyond seven eigenvalues, this being the limit for station EOFs.

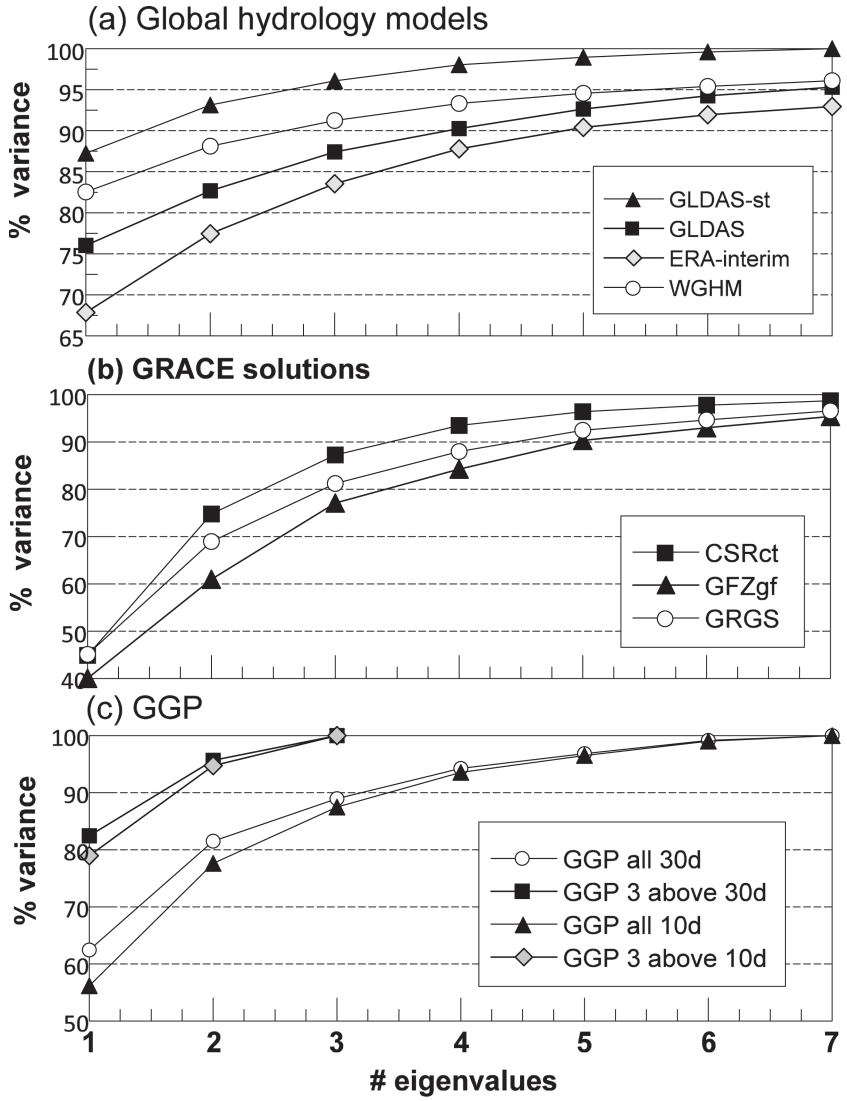

Figure 10. Cumulative variances for various EOF solutions: (a) Global hydrological models, (b) GRACE solutions, and (c) GGP station combinations.

6.1 Results for the EOF comparisons

We show the cumulative variance reduction as a function of number of eigenmodes for all data sets in Fig. 10. At this stage we distinguish between the EOF of the GLDAS data only for the seven stations—called GLDAS-st, as opposed to the EOF of the entire GLDAS array over the study area (GLDAS). The upper panel (a) shows that the GLDAS station EOF has higher variance reduction, and converges faster than the array EOF, indicating a simpler spatial and temporal structure. The WGHM solution has higher variance reduction (83 per cent for PC1) compared to the other two global total water storage simulations hydrology solutions because it is sampled monthly rather than at 10-d intervals. Somewhat surprisingly, the ERA-Interim solution is more complex (72 per cent) compared to GLDAS (76 per cent), perhaps because it includes more smaller wavelength precipitation, and therefore soil moisture.

In panel (b) we se that the GRACE solutions vary between a simpler structure for the cosine taper (showing CSRct, but GFZ is similar) than for the Gaussian filter (showing GFgf, but CSR is similar), due to the different wavenumber content. The GRGS solution is intermediate. Note that PC1 for the GRACE fields explains only 40–50 per cent of the variance reduction, which contradicts the large-scale GRACE spatial smoothing, and must therefore be due to noise in the data.

The GGP solutions are shown in panel (c) and differ markedly between using all seven stations, and using only the three surface stations (the latter as in Wziontek *et al.* 2009b). Differences depend on whether we decimate to 10 or 30 d. The three-station solutions converge quickly (only three eigenvalues), and the variance reduction for mode 1 (>80) per cent is higher than even the GLDAS-st solution.

To characterize the seasonal variability we concentrate on the first two principal components, PC1 and PC2. In Fig. 11 we see that GLDAS-st has higher amplitude than any other PC1 in this study, and GGP generally has the lowest amplitude when using all seven stations. The impact of the 2003 and 2006 summer heat waves (e.g. Teuling *et al.* 2010) is clearly seen in the PC1 of GLDAS, displaying the lowest values of the series at these dates. The 2003 summer heat wave is present in all data sets, and its impact on gravity measurements was first highlighted by Andersen *et al.* (2005). No abnormal low for the summer of 2006, however, is found in the

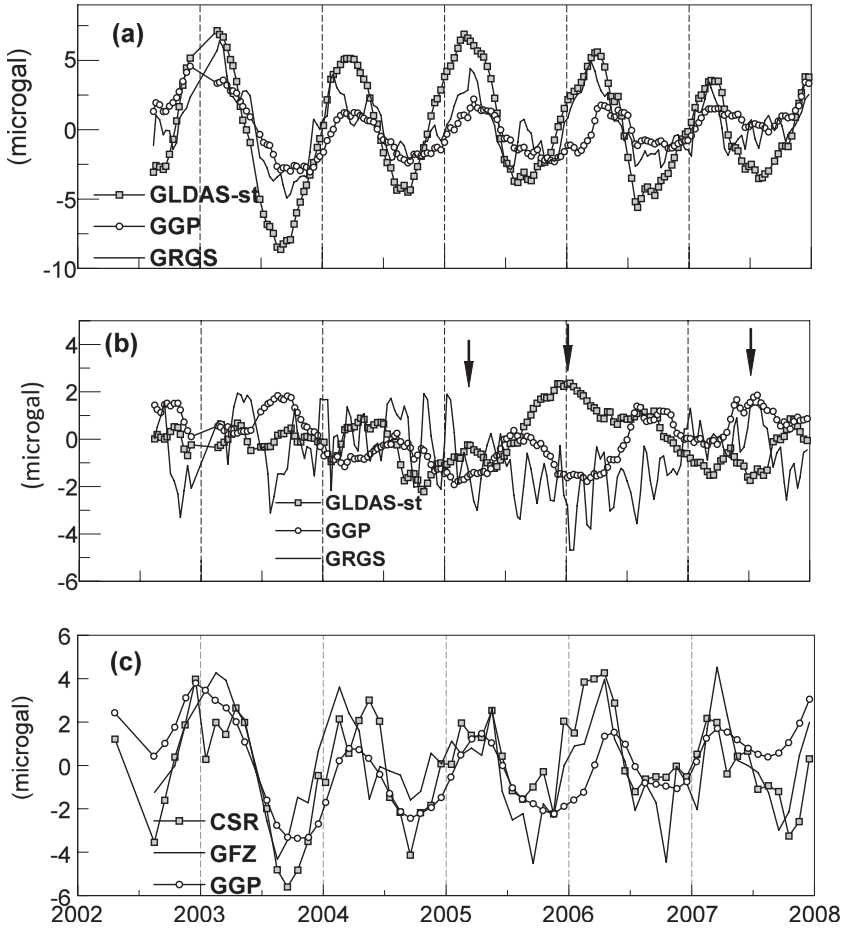

Figure 11. Comparison of principal components (a) 10-d PC1 for GRGS, GLDAS-st and GGP, (b) 10-d PC2 for GRGS, GLDAS-st and GGP (arrows indicate times when GGP and GLDAS-st are opposed) and (c) PC1 for CSR (mean), GFZ (mean) and GGP 30-d solutions.

PCs of GGP and GRACE (as also found by Seitz *et al.* 2008), suggesting that GLDAS simulations may be wrong for the 2006 event.

The next mode PC2 (Fig. 11b) is smaller for all fields than PC1, but the occasional phase opposition of GLDAS-st and GGP shows that PC2 may reflect the inverted water mass attraction for underground stations. The PC2 of GRGS is noisy, but is better correlated with GLDAS than GGP. We infer that PC1 being in phase with GRACE and GLDAS and PC2 being out-of-phase confirms our choice of PC1 as the best measure for the comparison of the data sets.

The rms difference between the CT20-40 and GF350 filters is only 0.5 μGal compared to 1.7 μGal between the PC1s of the CSR and GFZ fields. Henceforth we show (Fig. 11c) only the mean PC1 for the two CSR fields and GFZ fields compared to the GRGS solution. Differences sometimes reach 5 μGal or so, and are neither consistent between any two pairs of solutions, nor are they related to the variance reduction for PC1. Towards the end of 2007, GGP and GRGS agree well, CSR and GFZ agree somewhat, but the two pairs differ notably. The GRACE PC1s all have higher amplitude than GGP.

We compare the global PC1s for the hydrological models in Fig. 12. In panel (a) we see the GLDAS-st variation has higher amplitude than for the GLDAS array, because it is a much simpler solution (Fig. 10). The GLDAS array solution is similar to WGHM, sometimes higher and sometimes lower at the extremes. In panel (b) we compare WGHM and ERA-Interim with GRGS and in this case there is close agreement for the global hydrological models though they both have higher amplitude than the GRGS solution. The ERA-Interim field is the most complex in time, and the lowest amplitude.

The first eigenvector EV1 for the GRACE solutions (Fig. 13) shows less difference between the filters than between the sources (CSR or GFZ). The CSR/GFZ patterns are different, with a large dominant patch centred approximately on the Alps in the GFZ solutions. In the GRGS map this patch is larger and more diffuse (as expected for a 400 km footprint) and is further east, similar to the CSR pattern, suggesting a broader extent of snow cover in winter. In the final panel (f) we plot the GGP eigenvector for the seven stations, showing that the point values clearly differ from the GRACE distribution. Station WE dominates as a high and we see that the underground stations, especially ST and VI, appear as negative spots. Further discussion of the eigenvectors is given in the Supporting Information.

The EV1s for the global hydrological models are shown in Fig. 14. Of all the EV1s, only ERA-Interim shows a large amplitude over the Alps. Recall that ERA-Interim is not a global hydrological model, but the hydrological component of an atmospheric model. The precipitation forcing is, however, clearly better than the one used for GLDAS. WGHM uses GPCC monthly precipitation downscaled using ECMWF outputs of the number of rainy days in the month, so it has good precipitation. The signal in the Alps for ERA-Interim

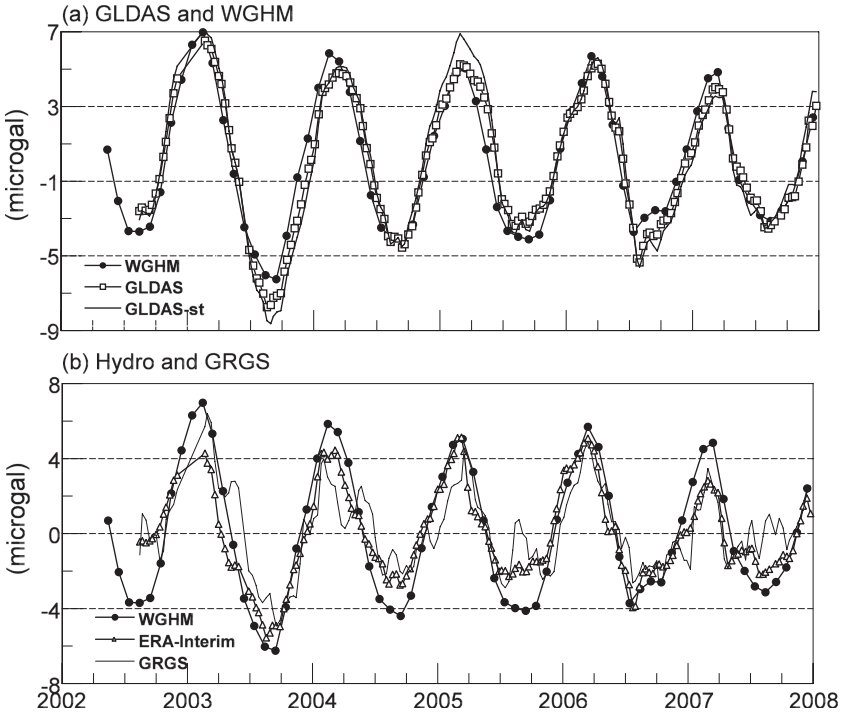

Figure 12. Comparison of principal components (a) global hydrological models (b) WGHM, ERA-Interim and GRGS. Note that the series for WGHM is common to both panels.

maybe due to improper snow modelling, maybe because of the lower temperature.

We have seen the PC1 amplitude of the seven-station GGP solution is lower than GRACE or GLDAS-st (Fig. 11), and a three-station solution has much higher variance reduction (Fig. 10). We therefore investigate the PC1 for different combinations of stations: all seven stations, three above-ground stations and four below-ground stations (Fig. 15a). As noted before, the water mass attraction at underground stations is only partly inverted, so we see a reduced amplitude for PC1 for the four-below stations compared to the seven-station solution. The three-above solution is the one that best represents ground gravity, using stations on the Earth's surface.

The benefit of using three stations is illustrated by comparing their PC1s with the GRGS and global hydrological simulations (Fig. 15b). Here we choose to plot the GLDAS hydrology PC1 to compare better with Fig. 12(b) where we compared GRGS with ERA-Interim. The agreement in both magnitude and phase of GGP with GRACE and global hydrological models becomes much closer than when using all seven stations.

7 DISCUSSION

7.1 Error budget

Errors in hydrological loading can be assessed by comparing different model simulations. Here we use GLDAS with NOAH, as compared to WGHM and ERA-Interim. We computed the standard deviation between these three models as $\sigma = 0.75 \pm 0.44\,\mu$Gal, with the largest σ's occurring in late winter and median values in summer; in autumn the models are very close. This indicates that the treatment of precipitation is a more significant difference between models than the drainage. Our values are consistent with Boy & Hinderer (2006) showing differences between LaD and GLDAS in the time domain of 1–2 μGal. We therefore take a representative error in the PC1 of global hydrological models to be 1 μGal.

Formal errors in the GRACE solutions for the various data sources can be derived from the covariance matrices of the spherical harmonic coefficients, but these are probably biased low. de Linage (2008, p.103) reports an error estimate from GRGS over Europe of about 2 μGal. From the data considered here, the rms deviation of the various PC1s between different filters is about 0.5 μGal, and the difference between CSR and GFZ solutions in Figs 4.3 and 4.4, is 1.7 μGal. Thus we consider values between 1.0 and 1.5 μGal as reasonable for the GRACE error in PC1.

Because we have only one realization of the GGP data set (with the seven stations considered here), errors in the EOF have to be propagated from errors in each stage of processing. The tidal models and polar motion are essentially error free, except for small tidal peaks at 12 and 24 hr, probably due to time-variable ocean loading effects, but these are not important at the 10–30 d averaging periods in this study. To be on the safe side we allow 0.1 μGal, just in case.

The global atmospheric pressure loading (Table 2), shows differences between the IB and two HUGO realizations. The latter is most relevant for our purpose, and we assess a standard deviation of 0.11 μGal. For the 3-D mass attraction effect, we have the rms deviation between Neumeyer *et al*.'s (2004 method) calculations for 2002–2004, and our model, yielding an overall rms error of 0.18 μGal.This is almost 2/3 of the computed amplitude for this correction. We attribute a possible error of 0.2 μGal for the long-term non-tidal ocean circulation, based on the estimates of Boy & Hinderer (2006) for ST and Kroner *et al.* (2009).

As shown in Hinderer *et al.* (2002), decisions about the type and size of offsets can vary significantly depending on individual judgment, and also the type of software available. Automatic procedures cannot deal with offsets, they have to be assessed manually. We took a comparison of the treatment of the ST data set 2002–2008 from Severine Rosat (personal communication, 2009), and one of the

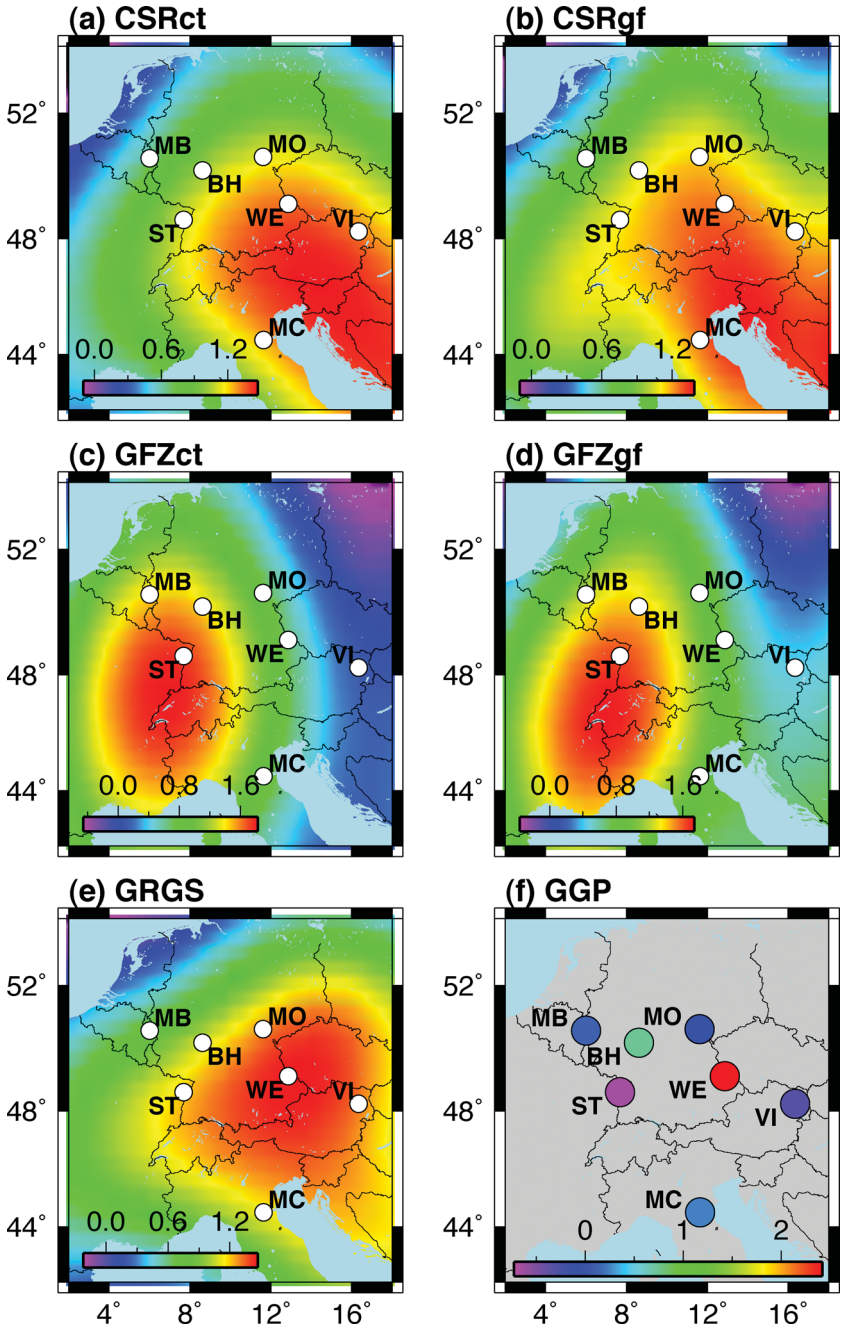

Figure 13. Comparison of EV1s for GRACE gridded solutions and GGP station data. The colour scales vary with the range of each data set to bring out the details; the scale in (f) is enlarged for clarity.

authors (DJC). Station ST has one of the lower number of offsets (Table 1), but the size of accumulated offset is typical. We found the rms deviation between the two processed series, at 10-d sampling, to be 0.82 μGal.

We summarize these errors in Table 3, and determine the overall σ from the sum of the variances. This is dominated by offsets, and comes to 0.9 μGal, which, for the sake of argument, we round up to 1.0 μGal. Note we have taken care to translate all the above estimates to each series with uniform sampling of 10 d, basically that of the GRGS data.

The 1 μGal error is propagated into the EOFs, following standard procedure (e.g. Savage & Svarc 2009), with a Monte Carlo simulation of 500 realizations and Gaussian distribution of error with $\sigma = 1.0$ μGal. The EOF statistics for each parameter (eigenvalues, PCs, and EVs) can then be found. We consider various combinations of stations, listed in Table 4, and show the per cent variance explained by PC1 and PC2.

The observational error becomes 1.4 μGal for PC1 in the three-station solution (Test #3), and 0.4 μGal for the seven stations. Interestingly, the error is less for the underground station Test (#4), probably because the amplitude of PC1 is also reduced, but this is offset by the reduction in variance of PC1. Errors in PC2 are somewhat correlated with those in PC1. Using seven stations, if we compare the PC1 error (0.4 μGal) with the recovered amplitude

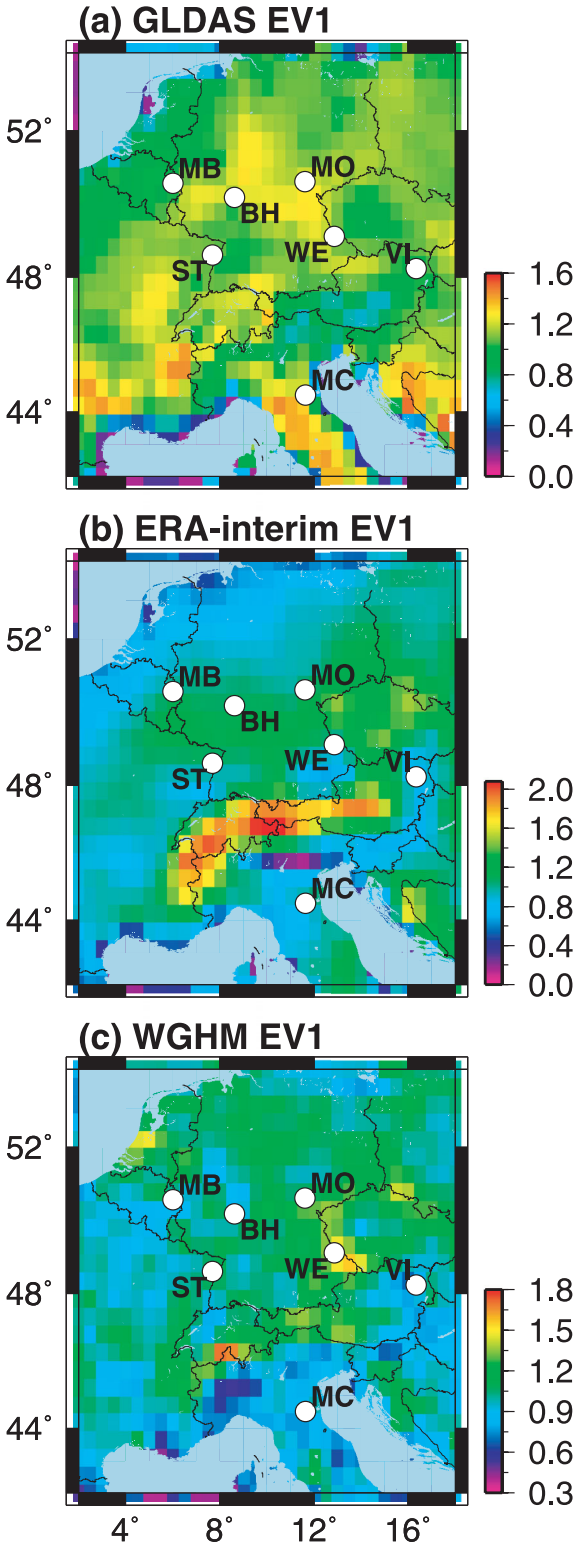

Figure 14. Comparison of EV1s for global hydrological models: (a) GLDAS, (b) ERA-Interim and (c) WGHM. The colour scales differ slightly to enhance the details.

(1.4 μGal) from Test #5 (Supporting Information), and the same for PC2 (0.7 μGal error vs. 0.9 μGal in amplitude), the mode 2 error increases over that of mode 1, while its amplitude decreases.

7.2 Comparing data series

With two series $\{x, y\}$ of length n with errors, there are several measures of similarity. Of the simpler methods we use the rms of centred differences $= \{\sum_i [(x_i - x_m) - (y_i - y_m)]^2/n\}^{(1/2)}$, where x_m and y_m are the series means. This is the rms deviation with means removed, used by Boronina & Ramillien (2008) in their comparison of GRACE and hydrology. We also compute the standard correlation coefficient, and the slope of the scatter plot between two data sets. One way to use the computed errors is to find the percentage of values between data sets that agree within some number of standard deviations.

We considered all possible combinations of GRACE, global hydrological models and GGP station EOFs, always using PC1 as the measure of comparison. As an illustration we take the GRGS with both 1.0 and 1.5 μGal errors and compare with the GGP PC1 for the seven-station and three-station solutions. The plot of PC1s are shown in Fig. 16 including error bars and indications of when the two series overlap within their 1 σ errors. Note that the GGP errors are given by the appropriate column in Table 4 (0.4 and 1.4 μGal) for the seven- and three-station solutions, respectively, and the GRGS error is taken as 1.5 μGal. Overall the agreement is 79 per cent for the seven-station solution.

Using this measure, the agreement with GRACE depends directly on the size of the errors, both for the GRGS solution and for the GGP error which is larger for three-station EOFs compared to seven-station EOFs. Reducing the GRGS error to 1.0 μGal, which is on the low side as discussed earlier, decreases the overlap to 58 and 66 per cent, respectively (Table 5). We see in Fig. 16 that the disagreement occurs mostly in winter months (December–February) of each year 2003–2007, as well as in 2005 summer. A similar three-station comparison (not shown) gives a slight improvement for the winter months of 2003 and 2005, and yields a 90 per cent agreement.

We compute further measures of the comparison in Table 5, noting that we use seven- and three-station GGP solutions, and use the GRGS solution as a reference with a 1.5 μGal error. Ideally the rms deviation should be small, and the correlation coefficient and slope both close to 1.0.

The rms deviation between the GGP series st7, st3 with GRGS lies between 1.6 and 1.9 μGal, less than the sum of the errors in each data set. Fig. 16, however, gives much more detail than this single value. The rms deviations show that GGP is closer to GRGS than the GLDAS or WGHM hydrology models, but the cross correlation is higher between GRGS and global hydrological than for ground gravity. The slopes of the scatter plots indicate the GGP PC1s are always less than GRGS because of the loss in amplitude due to the mixed water attraction. Of the global hydrological models, the ERA-Interim has the closest amplitude (slope = 0.94) to GRGS.

It is clear that a major limitation of the GGP data for GRACE comparisons (as noted by Neumeyer *et al.* 2004) is the location of four of the GGP stations underground. This would not be a factor for the EOF solution if all the soil moisture and groundwater were above the underground stations, but the problem lies in the mixed attraction sources at these stations. This was already evident in our synthetic Test #5 in the Supporting Information).

7.3 Hydrology remediation for underground stations

It is possible to make a gravity correction to account for soil moisture attraction between an underground station and the ground surface. Neumeyer *et al.* (2008) describe local corrections they applied for station MO, but it is not clear it is the same as we propose here. We

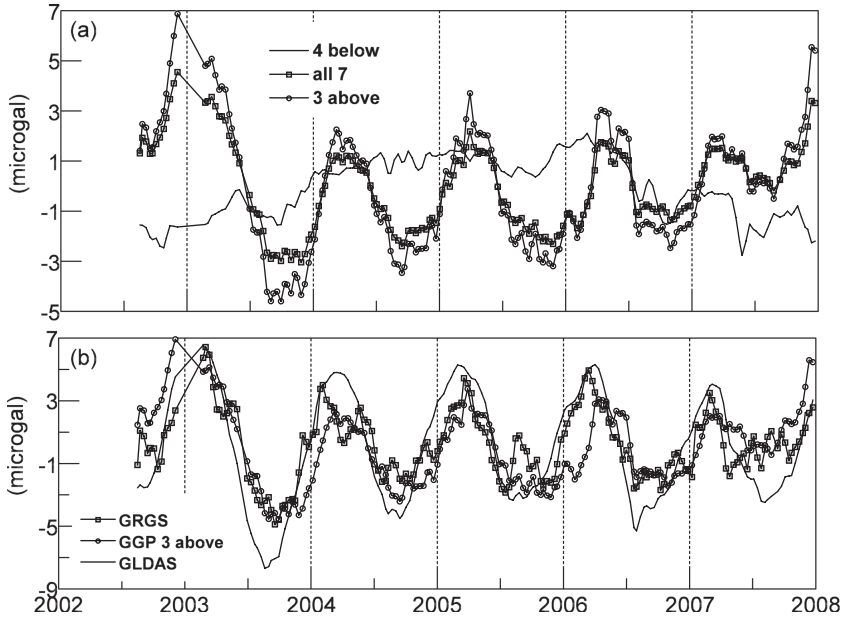

Figure 15. Comparison of PC1s (a) for different combinations of GGP stations; the three stations above ground give higher amplitudes than seven stations, but the four stations below ground show a much weaker seasonal signal, (b) GGP, GRGSG and GLDAS. The agreement between GGP and GRGS in (b) is quantified in Table 5.

Table 3. GGP error budget (1 σ, 10-d sampling).

Signal	μGal	Method
Solid earth, ocean tides	0.10	Nominal
Global pressure	0.11	Difference between IB, HUGO-m models
3-D mass attraction	0.18	Approximation to Neumeyer *et al.* (2004)
Non-tidal ocean circulation	0.20	Kroner *et al.* (2009), Boy & Hinderer (2006)
Uncertainty in offsets	0.82	Intercomparison of processing
Overall	$0.9 \rightarrow 1.0$	

Table 4. Error propagation in PCs, err in μGal.

Test #	Stations	Per cent PC1	PC1 err	PC2 err
1	WE	100	1.0	
2	BH, MC	68	1.43	1.04
3	BH, MC, WE	74	1.39	0.89
4	MB, MO, ST, VI	45	0.52	0.53
5	#4 + BH	36	0.91	0.92
6	#4 + BH, MC	35	0.56	0.54
7	All	49	0.39	0.69

suggest a relatively simple procedure, starting with the location of an SG underground (site A). For the ST J9 site, this is an underground bunker, and in the case of MB, MO and some other stations, the SGs are in tunnels dug into a mountainside. We seek a gravity correction to 'move' or 'transfer' the SG series to a surface site (call it B), so we need to know the gravity difference between site A and B under all environmental conditions.

To do this requires the type of gravity network installed around a gravity station, and realized by Naujoks *et al.* (2008), and also Jacob *et al.* (2010) in the case of an AG base station, but here the network is simple—just two stations. We suggest repeated gravity measurements between A and B, under all precipitation conditions (and monitoring environmental factors such as temperature, wind and relative humidity conditions), to develop a hydrological model of the leaky bucket type (Bower & Courtier 1998; Longuevergne 2008) that would match the observed gravity effect.

The gravity measurements can be made using a portable gravimeter (e.g. a Scintrex, Burris or gPhone), between the sites A and B. This might require measurements fairly often (perhaps daily or at least weekly) for one or two seasonal cycles. Once done, the model can then be extrapolated forward without the need for subsequent (or at least much less frequent) gravity measurements at B. One could even apply such corrections to previous gravity and precipitation data at the same site, provided all the required environmental factors had been adequately recorded.

8 CONCLUSIONS

We summarize our findings as follows.

(1) We have added the free-air effect of loading deformation to several GRACE solutions to make the gravity variations consistent with those observed at the Earth's surface. This gives larger ground amplitudes of about 34 per cent for the GRGS solutions, but this factor depends on the spectral energy content of each GRACE solution and therefore on the highest harmonic degree chosen and type of filter used.

(2) EOF analysis is well suited to the spatial fields in this study, and this method will address the issue of the spatial averaging of point data without correcting surface stations for near-station mass

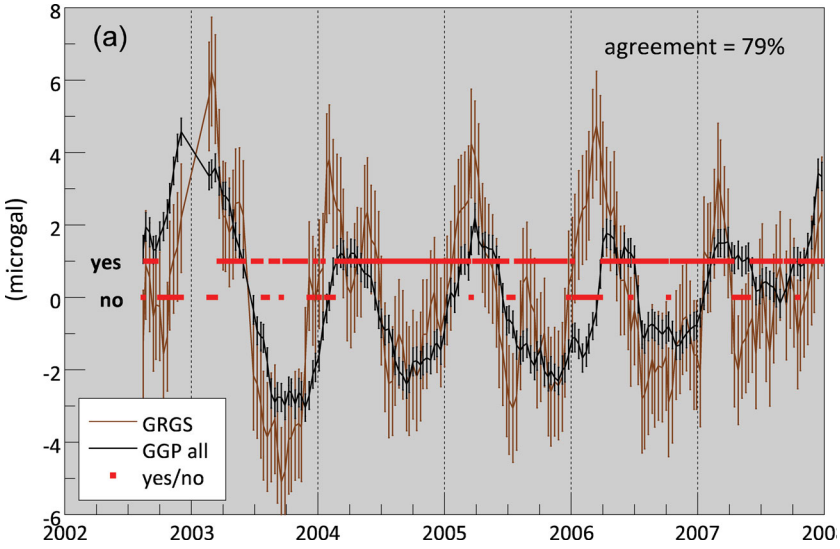

Figure 16. Agreement between ground and satellite solutions within a 1-σ error of 1 μGal for GGP and 1.5 μGal for GRGS using all seven stations.

Table 5. Comparison of PC1s between GRGS, GGP and global hydrological models.

Data	rms Deviation (μGal)	Correlation coefficient	Slope of scatter plot	Agreement[a]
GGP st7	1.67	0.68	0.55 ± 0.03	79 per cent
GGP st3	1.83	0.72	0.86 ± 0.04	90 per cent
GLDAS-st	2.38	0.85	1.51 ± 0.04	–
GLDAS	2.04	0.86	1.38 ± 0.04	–
WGHM	2.13	0.82	1.30 ± 0.04	–
ERA-Interim	1.44	0.82	0.94 ± 0.03	–

[a]Agreement is with respect to errors of 1.5 μGal for GRGS and 1 μGal for GGP.

(soil moisture and groundwater) attraction. Central Europe is the only area with enough SGs on a continental land mass to attempt an EOF analysis.

(3) The first principle components (PC1 and PC2) of the seven GGP stations in Europe are the primary measures for comparing ground data with various GRACE data and GLDAS data sets. The amplitudes of the GRACE fields are similar, but GRGS has more detail (in 10-d sampling) for comparison with GGP. The GLDAS and WGHM hydrological models have a larger seasonal component than GRACE, but the ERA-Interim model is similar in amplitude to GRGS. The three-station GGP solution is significantly better than the full EOF solution in recovering the true amplitude of the ground signal.

(4) The first eigenvectors of the ERA-Interim and GLDAS hydrological models are more similar to GRACE than WGHM, but the GGP spatial pattern is contaminated due to a mixed water attraction from the underground stations.

(5) Of the three global hydrological models considered (WGHM, GLDAS and ERA-Interim), only the latter shows the effect of Alps precipitation on the first eigenvector, but all three models agree internally at the level of 0.7 ± 0.4 μGal.

(6) Using the GRGS solution as an example, we determine that GGP can validate GRACE according to the following: using seven stations and 1.5 μGal error for GRGS we get agreement 79 per cent of the time, with only the three above-ground stations this agreement reaches 90 per cent. We consider the 79 per cent agreement to be well determined from our study.

(7) A further improvement to this type of analysis requires correcting SG data for soil moisture levels for a sensor below ground, but only that part of the signal that comes from attraction between the sensor and ground level. We have indicated how such a hydrological model could be established and calibrated by a high-quality portable spring gravimeter.

ACKNOWLEDGMENTS

We thank the GGP station operators for supplying the data for this paper, especially in some cases directly to us in advance, and also the B. Ducarme (ICET) and B. Ritschel (ISDC) in maintaining the GGP database. This material is based upon work supported by the National Science Foundation under Grant No. 0409381, the CNRS (French Research Ministry) and GRACE Science Team Grant No. NNG04GE99G. During this study J-PB was visiting NASA Goddard Space Flight Center, with a Marie Curie International Outgoing Fellowship (N° PIOF-GA-2008-221753). The GLDAS data used in this effort were acquired as part of the activities of NASA's Science Mission Directorate, and are archived and distributed by the Goddard Earth Sciences (GES) Data and Information Services Center (DISC). We also acknowledge the authors of the generic mapping tools GMT for the availability of this exceptional software.

Finally we sincerely thank the efforts of the two referees and the editor who made many helpful suggestions to improve the paper, in particular to suggest we treat several global hydrological models.

REFERENCES

Abe, M., Kroner, C., Neumeyer, J. & Chen, X.D., 2010. Assessment of atmospheric reductions for terrestrial gravity observations, *Bull. d'Inf. Marées Terrestres,* **146,** 11817–11838.

Andersen, O., Seneviratne, S.I., Hinderer, J. & Viterbo, P., 2005. GRACE-derived terrestrial water storage depletion associated with the 2003 European heat wave, *Geophys. Res. Lett.,* **32,** L18405, doi:10.1029/2005GL023574.

Bettadpur, S., 2007. *CSR Level-2 Processing Standards Document for Product Release 04, GRACE 327-742,* The GRACE Project, Cent. for Space Res., University of Texas at Austin, Austin, TX. Available at http://podaac.jpl.nasa.gov/gravity/grace-documentation#Level2Documentation.

Biancale, R. *et al.*, 2010. 7 *Years of Gravity Variations from GRACE and LAGEOS Data From CNES/GRGS*. Available at: http://grgs.obs-mip.fr/index.php/fre/Donnees-scientifiques/Champ-de gravite/grace (last accessed 2012 March 10).

Boronina, A. & Ramillien, G., 2008. Application of AVHRR imagery and GRACE measurements for calculation of actual evapotranspiration over the Quaternary aquifer (Lake Chad basin) and validation of groundwater models, *J. Hydrol.,* **348,** 98–109.

Bower, D.R. & Courtier, N., 1998. Precipitation effects on gravity measurements at the Canadian Absolute Gravity Site, *Phys. Earth planet. Int.,* **106**(3-4), 353–369.

Boy, J.-P. & Chao, B., 2005. Precise evaluation of atmospheric loading effects on Earths time-variable gravity field, *J. geophys. Res.,* **110,** B08412, doi:10.1029/2002JB002333.

Boy, J.-P. & Hinderer, J., 2006. Study of the seasonal gravity signal in superconducting gravimeter data, *J. Geodyn.,* **41,** 227–233.

Boy, J.-P. & Lyard, F., 2008. High-frequency non-tidal ocean loading effects on surface gravity measurements, *Geophys. J. Int.,* **175,** 35–45.

Boy, J.-P., Gegout, P. & Hinderer, J., 2002. Reduction of surface gravity data from global atmospheric pressure loading, *Geophys. J. Int.,* **149,** 534–545.

Boy, J.-P., Longuevergne, L., Boudine, F., Jacob, T., Lyard, F., Llubes, M., Florsch, N. & Esnoult, M.-F., 2009. Modelling atmospheric and induced non-tidal oceanic loading contributions to surface gravity and tilt measurements, *J. Geodyn.,* **48,** 182–188.

Bruinsma, S., Lemoine, J.-M., Biancale, R. & Vales, N., 2010. CNES/GRGS 10-day gravity field models (release 2) and their evaluation, *Adv. Space Res.,* **45,** 587–601.

Carrère, L. & Lyard, F., 2003. Modeling the barotropic response of the global ocean to atmospheric wind and pressure forcing: comparison with observations, *Geophys. Res. Lett.*, **30**(6), 1275, doi:10.1029/2002GL016473.

Chen, X., Ducarme, B., Sun, H.P. & Xua, J., 2008. Loading effect of a self-consistent equilibrium ocean pole tide on the gravimetric parameters of the gravity pole tides at superconducting gravimeter stations, *J. Geodyn.,* **45,** 201–207.

Chen, X., Kroner, C., Sun, H.-P., Abe, M., Zhou, J., Yana, H. & Wziontek, H., 2009. Determination of gravimetric parameters of the gravity pole tide using observations recorded with superconducting gravimeters, *J. Geodyn.,* **48,** 348–353.

Creutzfeldt, B., Güntner, A., Klügel, T. & Wziontek, H., 2008. Simulating the influence of water storage changes on the superconducting gravimeter of the Geodetic Observatory Wettzell. Germany, *Geophysics,* **73**(6), WA95–WA104.

Creutzfeldt, B., Güntner, A., Thoss, H., Merz, B. & Wziontek, H., 2010a. Measuring the effect of local water storage changes on in-situ gravity observations: case study of the Geodetic Observatory Wettzell, Germany, *Water Resour. Res.,* **46,** W08531, doi:10.1029/2009WR008359.

Creutzfeldt, B., Güntner, A., Wziontek, H. & Merz, B., 2010b. Reducing local hydrology from high-precision gravity measurements: a lysimeter-based approach, *Geophys. J. Int.,* **183**(1), 178–187.

Crossley, D. & Hinderer, J., 2009. A review of the GGP network and scientific challenges, *J. Geodyn.,* **48,** 299–304.

Crossley, D., Hinderer, J., Llubes, M. & Florsch, N, 2003. The potential of ground gravity measurements to validate GRACE data, *Adv. Geosci.,* **1,** 65–71.

Crossley, D., Hinderer, J. & Boy, J.-P., 2004. Regional gravity variations in Europe from superconducting gravimeters, *J. Geodyn.,* **38,** 325–342.

Crossley, D., Hinderer, J. & Boy, J.-P., 2005. Time variation of the European gravity field from superconducting gravimeters, *Geophys. J. Int.,* **161,** 257–264.

Crossley, D., Hinderer, J., Boy, J.-P. & de Linage, C., 2006. Status of the GGP Satellite project, *Bull. d'Inf. Marées Terrestres,* **142,** 11423–11432.

Dee, D.P. *et al.*, 2011. The ERA Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, **137,** 553–597.

Dehant, V., 1987. Tidal Parameters for an inelastic Earth, *Phys. Earth planet. Int.,* **49,** 97–116.

Döll, P., Kaspar, F. & Lehner, B., 2003. A global hydrological model for deriving water availability indicators: model tuning and validation, *J. Hydrol.,* **270,** 105–134.

Eanes, R., 2002. *The CSR4.0 Global Ocean Tide Model*. Available at: ftp://ftp.csr.utexas.edu/pub/tide (last accessed 2012 March 10).

Farrell, W.E., 1972. Deformation of the Earth by surface loads, *Rev. Geophys.,* **10,** 761–797.

Fletchner, F., 2007. AOD1B Product Description Document for Product Releases 01 to 04, GRACE 327-750, CSR publ. GR-GFZ-AOD-0001 Rev. 3.1, University of Texas at Austin, Austin, TX, 43pp. Available at: http://podaac.jpl.nasa.gov/dataset/GRACE_L1B_GRAV_JPL_RL01 (last accessed 2012 March 11).

Han, S.-C., Shum, C.K., Jekeli, C., Kuo, C.-Y., Wilson, C. & Seo, K.-W., 2005. Non-isotropic filtering of GRACE temporal gravity for geophysical signal enhancement, *Geophys. J. Int.,* **163**(1), 18–25.

Hasan, S., Troch, P.A., Bogaart, P.W. & Kroner, C., 2008. Evaluating catchment-scale hydrological modeling by means of terrestrial gravity observations, *Water Resour. Res.,* **44**(8), W08416, doi:10.1029/2007wr006321.

Hinderer, J., Rosat, S., Crossley, D., Amalvict, M., Boy, J.-P. & Gegout, P., 2002. Influence of different processing methods on the retrieval of gravity signals from GGP data, *Bull. d'Inf. Marées Terrestres,* **135,** 10653–10668.

Hinderer, J., Andersen, O., Lemoine, F., Crossley, D. & Boy, J.-P., 2006. Seasonal changes in the European gravity field from GRACE: a comparison with superconducting gravimeters and hydrology model predictions, *J. Geodyn.,* **41,** 59–68.

Hinderer, J., Crossley, D. & Warburton, R.J., 2007. Superconducting gravimetry, in *Treatise on Geophysics,* Vol. 3, pp. 65–122, eds Herring, T. & Schubert, G., Elsevier, New York, NY.

Hokkanen, T., Korhonen, K. & Virtanen, H., 2006. Hydrogeological effects on superconducting gravimeter measurements at Metsähovi in Finland, *J. Environ. Eng. Geophys.,* **11**(4), 261–267, doi:10.2113/jeeg11.4.261.

Jacob, T., Bayer, R., Chery, J., Jourde, H. & Moigne, N.L., 2010. Time-lase microgravity surveys reveal water storage heterogeneity of a karst aquifer, *J. geophys. Res.,* **115,** B06402, doi:10.1029/2009JB006616.

Jekeli, C., 1981. Alternative methods to smooth the Earths gravity field, Report #327, Dept. Geod. and Sci. and Surv., Ohio State University.

Kanzow, T., Flechtner, F., Chave, A., Schmidt, R., Schwintzer, P. & Send, U., 2005. Seasonal variation of ocean bottom pressure derived from Gravity Recovery and Climate Experiment (GRACE): local validation and global patterns, *J. geophys. Res.,* **110,** C09001, doi:10.1029/2004JC002772.

Klügel, T. & Wziontek, H., 2009. Correcting gravimeters and tiltmeters for atmospheric mass attraction using operational weather models, *J. Geodyn.,* **48**(3–5), 204–210.

Kroner, C. & Jahr, T., 2006. Hydrological experiments around the superconducting gravimeter at Moxa Observatory, *J. Geodyn.,* **41,** 268–275.

Kroner, C., Dierks, O., Neumeyer, J. & Wilmes, H., 2005. Analysis of observations with dual sensor superconducting gravimeters, *Phys. Earth planet. Inter.,* **153,** 210–219.

Kroner, C., Jahr, T., Naujoks, M. & Weise, A., 2007. Hydrological signals in gravity foe or friend? In *IAG Symposia Dynamic Planet,* Vol. **130**, pp. 504–510, Springer, Berlin.

Kroner, C., Thomas, M., Dobslaw, H., Abe, M. & Weise, A., 2009. Seasonal effects of non-tidal oceanic mass shifts in observations with superconducting gravimeters, *J. Geodyn.*, **48,** 354–359.

Lambert, A., Henton, J. & Winester, D., 2009. GRACE monthly estimates compared with surface gravity variations at mid-North American Sites, *Paper Presented at NOAA Workshop: Monitoring NA Geoid Change,* Boulder, CO, 2009 October.

de Linage, C., 2008. Mésures gravimétriques au sol et satellitaires: étude du rapport entre variation de pesanteur et déplacement vertical et apport de la mission spatiale GRACE à l'étude des surcharges hydrologiques et des très grand séismes, *PhD thesis,* Univ. Louis Pasteur Strasbourg, Strasbourg. Available at: http://tel.archives-ouvertes.fr/docs/00/34/91/60/PDF/these_octobre2008.pdf (last accessed 2012 March 11).

de Linage, C., Hinderer, J. & Rogister, Y., 2007. A search for the ratio between gravity variation and vertical displacement due to a surface load, *Geophys. J. Int.,* **171,** 986–994.

de Linage, C., Hinderer, J. & Boy, J.-P., 2009. Variability of the gravity-to-height ratio due to surface loads, *Pure appl. Geophys.,* **166,** 1217–1245.

Llubes, M., Florsch, N., Hinderer, J., Longueverne, L. & Amalvict, M., 2004. Local hydrology, Global Geodynamics Project and CHAMP/GRACE/GOCE perspectives: some case studies, *J. Geod.,* **38,** 355–374.

Longuevergne, L., 2008. Contribution à l'hydrogeodésie. *Ph.D. thesis*, UPMC University Paris 6, 300 pp.

Longuevergne, L., Boy, J.-P., Florsch, N., Viville, D., Ferhat, G., Ulrich, P., Luck, B. & Hinderer, J., 2009. Local and global hydrological contributions to gravity variations observed in Strasbourg (France), *J. Geodyn.,* **48,** doi:10.1016/j.jog.2009.09.008.

Longuevergne, L., Scanlon, B. & Wilson, C., 2010. GRACE Hydrological estimates for small basins: evaluating processing approaches on the High Plains Aquifer, USA, *Water Resour., Res.,* **46,** W11517, doi:10.1029/2009WR008564.

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

Meurers, B., 2010. Environmental effects on gravity at Conrad Observatory, Austria, *Geophys. Res. Abs.,* **12,** EGU2010-3656.

Milly, P.C.D. & Shmakin, A.B., 2002a. Global modeling of land water and energy balances, part I: the land dynamics (LaD) model, *J. Hydromet.,* **3,** 283–299.

Milly, P.C.D. & Shmakin, A.B., 2002b. Global modeling of land water and energy balances, part II: land characteristic contributions to spatial variability, *J. Hydromet.,* **3,** 301–310.

Naujoks, M., Weise, A., Kroner, C. & Jahr, T., 2008. Detection of small hydrological variations in gravity by repeated observations with relative gravimeters, *J. Geod.,* **82,** 543–553.

Naujoks, M., Kroner, C., Weise, A., Jahr, T., Krause, P. & Eisner, S., 2010. Evaluating local hydrological modelling by temporal gravity observations and a gravimetric 3D model, *Geophys. J. Int.,* **182,** 233–249.

Neumeyer, J., Hagedoorn, J., Leitloff, J. & Schmidt, R., 2004. Gravity reduction with three-dimensional atmospheric pressure data for precise ground gravity measurements, *J. Geodyn.*, **38,** 437–450.

Neumeyer, J. *et al.*, 2006. Combination of temporal gravity variations resulting from Superconducting Gravimeter (SG) recordings, GRACE satellite observations and global hydrology models, *J. Geod.,* **79,** 573–585,doi: 10.1007/s00190-005-0014-8.

Neumeyer, J., Barthelmes, F., Kroner, C., Petrovic, S., Schmidt, R., Virtanen, H. & Wilmes, H., 2008. Analysis of gravity field variations derived from superconducting gravimeter recordings, the GRACE satellite and hydrological models at selected European sites, *Earth Planets Space,* **60,** 505–518.

Park, J.-H., Watts, D.R., Donohues, K.A. & Jayne, S.R., 2008. A comparison of in situ bottom pressure array measurements with GRACE estimetes in the Kuroshio Extension, *Geophys. Res. Lett.,* **35,** L17601, doi:10.1029/2008GL034778.

Pfeffer, J. *et al.*, 2011. Local and global hydrological contributions to time-variable gravity in Southwest Niger, *Geophys. J. Int.,* **184,** 661–672, doi:10.1111/j.1365-246X.2010.04894.x.

Pierce, D., 2003. *Empirical Orthogonal Function (EOF) Software,* Climate Research Division, Scripps Institute of Oceanography. Available at: http://meteora.ucsd.edu/pierce/eof/eofs.html (last accessed 2012 March 11).

Pálinkás, V., 2009. Modelling hydrological effects on gravity, *Geophys. Res. Abs.*, **11,** EGU2009-11634.

Riccardi, U., Rosat, S. & Hinderer, J., 2011. Comparison of the Micro-g LaCoste gPhone-054 spring gravimeter and the GWR-C026 superconducting gravimeter in Strasbourg (France) using a 300-day time series, *Metrologia,* **48,** 28–39.

Rietbroek, R., Brunnabend, S.-E., Dahle, C., Kusche, J., Flechtner, F., Schröter, J. & Timmermann, R., 2009. Changes in total ocean mass derived from GRACE, GPS, and ocean modeling with weekly resolution, *J. geophys. Res.,* **114,** C11004, doi:10.1029/2009JC005449.

Rodell, M. *et al.*, 2004. The global land data assimilation system, *Bull. Am. Met. Soc.,* **85**(3), 381–394. [Available online at http://disc.gsfc.nasa.gov/services/grads-gds/gldas.shtml.]

Rosat, S., Boy, J.-P., Ferhat, G., Hinderer, J., Amalvict, M., Gegout, P. & Luck, B., 2009. Analysis of a 10-year (19972007) record of time-varying gravity in Strasbourg using absolute and superconducting gravimeters: new results on the calibration and comparison with GPS height changes and hydrology, *J. Geodyn.,* **48,** 360–365.

Savage, J.C. & Svarc, J.L., 2009. Postseismic relaxation following the 1992 M7.3 Landers and 1999 M7.1 Hector Mine earthquakes, southern California, *J. geophys. Res.,* **114,** B01401, doi:10.1029/2008JB005938.

Seitz, F., Schmidt, M. & Shum, C.K., 2008. Signals of extreme weather conditions in Central Europe in GRACE 4-D hydrological mass variations, *Earth planet. Sci. Lett.,* **268,** 165–170.

Seo, K.-W. & Wilson, C.R., 2005. Estimating GRACE aliasing errors, in *IAG Symposia: Gravity, Geoid, and Space Missions,* Vol. 129, 339–345, eds Jekeli, C., Bastos, L.M.C. & Fernandez, J., Springer, Berlin.

Simon, D., 2002. Modeling of the field of gravity variations induced by the seasonal air mass warming during 1998- 2000, *Bull. d'Inf. Marées Terrestres,* **136,** 10821–10836.

Simon, D., 2006. Gravimetric effects induced by vertical air mass shifts at Medicina (1998-2005), Wettzell, Bad Homburg, Moxa, Pecny and Wien (1998-2004), *Bull. d'Inf. Marées Terrestres,* **142,** 11317–11322.

Swenson, S. & Wahr, J., 2006. Postprocessing removal of correlated errors in GRACE data, *Geophys. Res. Lett.,* **33,** L08402, doi:10.1029/2005GL025285.

Swenson, S., Wahr, J. & Milly, P.C.D., 2003. Estimated accuracies of regional water storage variations inferred from the Gravity Recovery and Climate Experiment (GRACE), *Water Resour. Res.,* **39**(8), 1223, doi:10.1029/2002WR001808.

Swenson, S.C., Yeh, P.J.-F., Wahr, J. & Famiglietti, J.S., 2006. A comparison of terrestrial water storage variations from GRACE with in situ measurements from Illinois, *Geophys. Res. Lett.,* **33,** L16401, doi:10.1029/2006GL026962.

Swenson, S., Famiglietti, J., Basara, J. & Wahr, J., 2008. Estimating profile soil moisture and groundwater variations using GRACE and Oklahoma Mesonet soil moisture data, *Water Resour. Res.,* **44,** W01413, 1–12, doi:10.1029/2007WR006057.

Tamura, Y., 1987. A harmonic development of the tide generating potential, *Bull. d'Inf. Marées Terrestres,* **99,** 6813–6855.

Tamura, Y., Sato, T., Ooe, M. & Ishiguro, M., 1991. A procedure for tidal analysis with a Bayesian information criterion, *Geophys. J. Int.,* **104,** 507–516.

Teuling, A.J. *et al.*, 2010. Contrasting response of European forest and grassland energy exchange to heatwaves, *Nat. Geosci.,* **3,** 722–727.

Van Camp, M., Williams, S.D.P. & Francis, O., 2005. Uncertainty of absolute gravity measurements, *J. geophys. Res.,* **110,** B05406, doi:10.1029/2004JB003497.

Van Camp, M., Vanclooster, M., Crommen, O., Petermans, T., Verbeeck, K., Meurers, B., van Dam, T. & Dassargues, A., 2006. Hydrogeological investigations at the Membach station, Belgium, and application to correct long periodic gravity variations, *J. geophys. Res.,* **111,** B10403, doi:10.1029/2006JB004405.

Velicogna, I. & Wahr, J., 2001. Potential problems with the use of gravimeter data for GRACE Cal/Val, *EOS, Trans. Am. geophys. Un.,* **82**(47), Fall Meet. Suppl., Abstract G51A-0242

Virtanen, H., 2004. Loading effects in Metsahovi from the atmosphere and the Baltic Sea, *J. Geodyn.,* **38,** 407–422.

Virtanen, H. & Makinen, J., 2003. The effect of the Baltic Sea level on gravity at the Metsahovi station, *J. Geodyn.,* **35,** 553–565.

Virtanen, H., Tervo, M. & Bilker-Koivula, M., 2006. Comparison of superconducting gravimeter observations with hydrological models of various spatial extents, *Bull. d'Inf. Marées Terrestres,* **142,** 11361–11368.

Wahr, J., Molenaar, M. & Bryan, F., 1998. Time-variability of the Earths gravity field: hydrological and oceanic effects and their possible detection using GRACE, *J. geophys. Res.,* **103,** 30205–30230.

Wahr, J., Swenson, S., Zlotnicki, V. & Velicogna, I., 2004. Time-variable gravity from GRACE: first results, *Geophys. Res. Lett.*, **31**(11), L11501, doi:10.1029/2004GL019779.

Wahr, J., Swenson, S. & Velicogna, I., 2006. Accuracy of GRACE mass estimates, *Geophys. Res. Lett.,* **33,** L06401, doi:10.1029/2005GL025305.

Wang, X., de Linage, C., Famiglietti, J.S. & Zender, C., 2011. Gravity Recovery and Climate Experiment (GRACE) detection of water storage changes in the Three Gorges Reservoir of China and comparison with in situ measurements, *Water Resour. Res.,* **47**, W12502, doi:10.1029/2011WR010534.

Warburton, R.J. & Goodkind, J.M., 1977. The influence of barometric-pressure variations on gravity, *Geophys. J. R. astr. Soc.,* **48,** 281–292.

Weise, A., Kroner, C., Abe, M., Ihde, J., Jentzsch, G., Naujoks, M., Wilmes, H. & Wziontek, H., 2009. Gravity field variations from superconducting gravimeters for GRACE validation, *J. Geodyn.,* **48,** 325–330.

Wenzel, H.G., 1996. The nanogal software: earth tide data processing package: ETERNA 3.3, *Bull. dInf. Marées Terrestres,* **124,** 9425–9439.

Widmer-Schnidrig, R., Duffner, P., Forbriger, Th. & Zurn, W., 2010. The new dual sphere superconducting gravimeter at the Black Forest Observatory, Abstract GG P02, p.64, German Geophysical Society Meeting, Bochum, Germany, March 2010.

Wilkes, D.S., 1995. *Statistical Methods in the Atmospheric Sciences,* Chapter 9, Academic Press, San Diego, CA.

Wziontek, H., Falk, R., Wilmes, H. & Wolf, P., 2009a. Precise gravity time series and instrumental properties from combination of superconducting and absolute gravity measurements, in *Observing our Changing Earth,* pp. 301–306, ed. Sideris, M.G., Springer, Berlin.

Wziontek, H., Wilmes, H., Wolf, P., Werth, S. & Güntner, A., 2009b. Time series of superconducting gravimeters and water storage variations from the global hydrology model WGHM, *J. Geodyn.,* **48**(3–5), 166–171.

Zerbini, S., Richter, B., Negusini, M., Romagnoli, C., Simon, D., Domenichini, F. & Schwahn, W., 2001. Height and gravity various by continuous GPS, gravity and environmental parameter observations in the southern Po Plain, near Bologna, Italy, *Earth planet. Sci. Lett.,* **192,** 267–279.

Zerbini, S., Richter, B., Rocca, F., van Dam, T. & Matonti, F., 2007. A combination of space and terrestrial geodetic techniques to monitor land subsidence: case study, the Southeastern Po Plain, Italy, *J. geophys. Res.,* **112,** B05401, doi:10.1029/2006JB004338.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Supplement. Empirical orthogonal function (EOF) properties and tests.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Electronic Supplement – EOF Properties and Tests

Empirical orthogonal function (EOF) analysis, or Principal Component Analysis (PCA), is widely known; Preisendorfer and Mobley (1988) give a comprehensive pedagogical development, and we use Wilkes (1995) for the basic theory. We first implemented EOFs for the ground-satellite comparison in Crossley et al, (2004) and profited from the FORTRAN code provided by Pierce (2003), as did Neumeyer et al. (2008).

S1. Basic expressions

The Europe fields (gravity, total water storage) can be represented as a space array (lat, long) at certain epochs or times (e.g. 10 days, 1 month). Each space array is packed into a column of a system matrix A, where columns represent sample times. Array dimensions (nx,nt) vary depending on the space and time sampling. We use the 7 ground stations without gridding and at CSR GRACE epochs (for example), in which case the dimensions of A are (7,65). For data are on a 0.25°grid of (65x49) at GRGS epochs, the dimensions of A are (3185,187).

The EOF routine performs an eigenvalue decomposition of the data covariance matrix $B = AA^T$, solving $Bu = \lambda u$ (equivalent to a singular value decomposition of A), for the eigenvectors U and eigenvalues λ ordered by decreasing size. The principle components (PCs) for each mode are found from $V = A^T U$. We take U as a 2-D spatial field (eigenvectors, or EVs) and V^T as a 1-D time series (PCs). A reconstructed field at each epoch can be generated from the first p eigenmodes $A' = \sum_p UV^T$.

S2. Normalization

We wish to interpret the PCs in this paper and so need to know their units. Orthonormalized eigenvectors are usual in SVD analysis, and the PCs are scaled accordingly. The squared variance of each PC equals its associated eigenvalue, but the units of the PCs are left arbitrary, provided the reconstruction retrieves the original data. For our purpose, however, this procedure did not reproduce the PCs in the original units (μGal) of the test data sets.

We therefore renormalized the outputs, scaling the eigenvectors and the PCs by the square root of the dimension of the input array. This scaling was arrived at by comparing results from stations with those from a gridded array. For 7 stations, each eigenvector has a norm of 7.0; for a grid of 65x49, the norm is the number of cells (3185). The PCs become independent of the spatial dimensions of the input data, and they can be interpreted in μGal (in our case).

S3. Annual wave tests

A second question concerns the phase of the seasonal component in the SG residual series, because some of the stations below ground clearly have a reversed amplitude component compared to those above ground, implying a phase shift of 180°. How does this affect the PCs and eigenvectors?

We used cosine functions with specific amplitude and phase information, as tests of our EOF code and also approximated the observed GGP residuals by cosine waves $g_{ann} = A\,cos\,2\pi\,[(date-2002)+C/365.25]+D$. The latter are shown in Figure 8 with their phases indicated with respect to 2003.0, and their parameters are in Table S1. Note that stations with mixed hydrology (MB, MO, ST, VI) have the smallest amplitudes, and the phase C is not correlated with geography.

Table S1. Fitted annual waves

station	code	A (μGal)	C (day)
1	BH	1.70	-65.7
2	MB	-1.63	-88.1
3	MC	2.15	-32.7
4	MO	0.73	-22.1
5	ST	-0.66	-123.6
6	VI	-0.91	-156.1
7	WE	2.77	-101.2

These annual waves are used as a test input (Test #5) to the EOF program and are sampled at 10-day (GRGS) times. The values for each test (the first 4 are synthetic) are as follows (giving only those that differ from the previous test, and with amplitudes A in μGal , phases C in days): (1) all A = 3.0, all C = -60, (2) for stations 2 and 4-6, A = -3.0, (3) for stations 2 and 4-6, A = -1.0, (4) C = -60, -30, -60, -90, -120, -30, -60, and (5) A and C as given in Table S1. For each of the tests we noted the number of modes found (p), the variance explained (V1, V2), the amplitudes and phases (C1, C2) of the first two PC's, and the eigenvector EV1 (Table S2). For Test #5 we add the second mode eigenvector.

Table S2. EOF test solutions

test	p	V1	V2	PC1	C_1	PC2	C_2	EV1						
		%	%	μGal	day	μGal	day	BH	MB	MC	MO	ST	VI	WE
1	1	100	0	3.0	-58	0	0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2	1	100	0	3.0	-58	0	0	1.0	-1.0	1.0	-1.0	-1.0	-1.0	1.0
3	1	100	0	2.1	-58	0	0	1.43	-0.48	1.43	-0.48	-0.48	-0.48	1.43
4	2	95.4	4.6	2.1	-58	0.46	-29	1.46	-0.41	1.46	-0.41	-0.41	-0.41	1.46
5	2	76.3	23.7	1.4	-80	0.84	-178	1.12	-1.12	0.96	0.24	-0.35	-0.20	1.83
							E V2*	-0.55	-0.13	-1.91	-0.71	-0.51	-1.05	1.07

* this line applies only to Test #5

Only two modes are necessary to represent the data in Table S1. Clearly PC1 matches the input data for tests 1 and 2, so the scaling of the PCs is in μGal , as desired. Reversing the amplitude of the annual wave (test 2) does not change PC1; it is the EV1 sign which is changed. Changing amplitudes (test 3) changes the PC1 and the EV1s, and when phases are shifted (test 4), a second mode appears. Test #5 gives a smaller PC1 than the higher-amplitude stations in Table S1, because of the mixing effect of the other stations.

Note that $C_1 = -58$ is only approximately the expected -60 day phase of the first 4 test series due to the 10-day sampling, and $C_1 = -80$ (end of March) of Test #5 approximates the mean C of the

strongest stations in Table S1. C_2 accounts for the secondary phases in the respective data series. Figure S1 shows the first two PCs for Test #5.

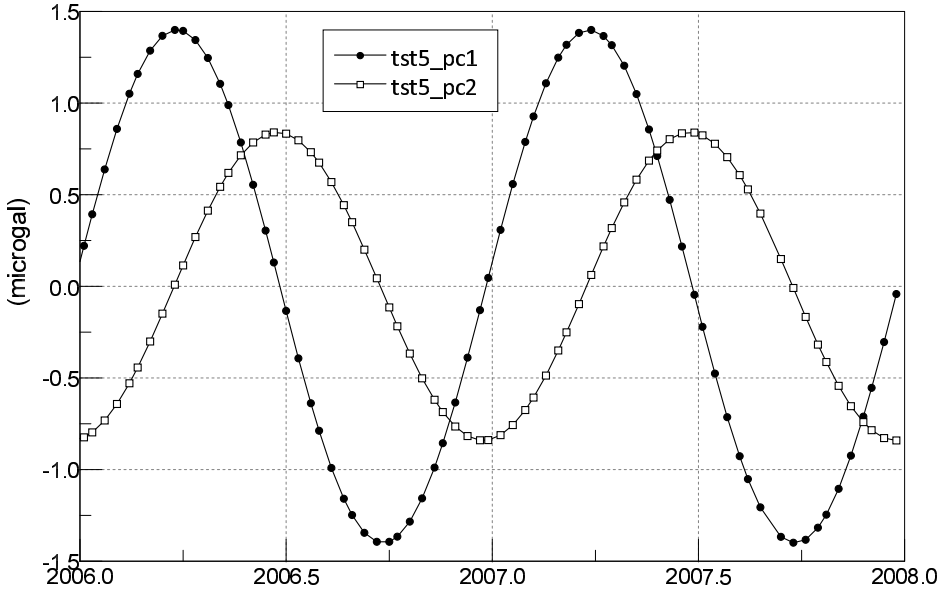

Figure S1. PCs for the annual wave test, showing PC1 (76%) and PC2 (23%)

With a station EOF, the eigenvectors are single values at the station locations, and could be interpolated to represent a spatial map without adding any new information. In Table S2, the EV1 for Test #5 displays a strong anti-symmetry between stations WE and MB, and stations ST and VI are negative as expected.

Figure S2(a) displays a comparison of the EV1s for the GRACE solutions (sampled at the stations from Figure 13), and grouped by dataset. Note that the sign of all the EV1s for the GRACE solutions are positive, and the only significant differences between the solutions are for stations VI and WE. In Figure S2(b) we compare the GGP station EVs from Figure 13(f) with the solution for the annual waves in Table S2. Test #5 generally gives larger EV1s than the actual stations (except for ST and VI) because the annual waves are simpler (Figure 8). Note the EV1 for MC is much lower than expected considering it is a ground-level station with a strong annual component. Also, the EV1 for MB is not as negative as expected from the annual wave considering that the SG is well below the ground surface. But these properties (MC large and positive, MB large and negative) are clearly seen in the second mode, EV2, for the actual GGP data; so EV1 by itself is insufficient to represent the GGP data. Station MO is ambiguous and has a difference of sign, in both EV1 and EV2, between the GGP and test solutions.

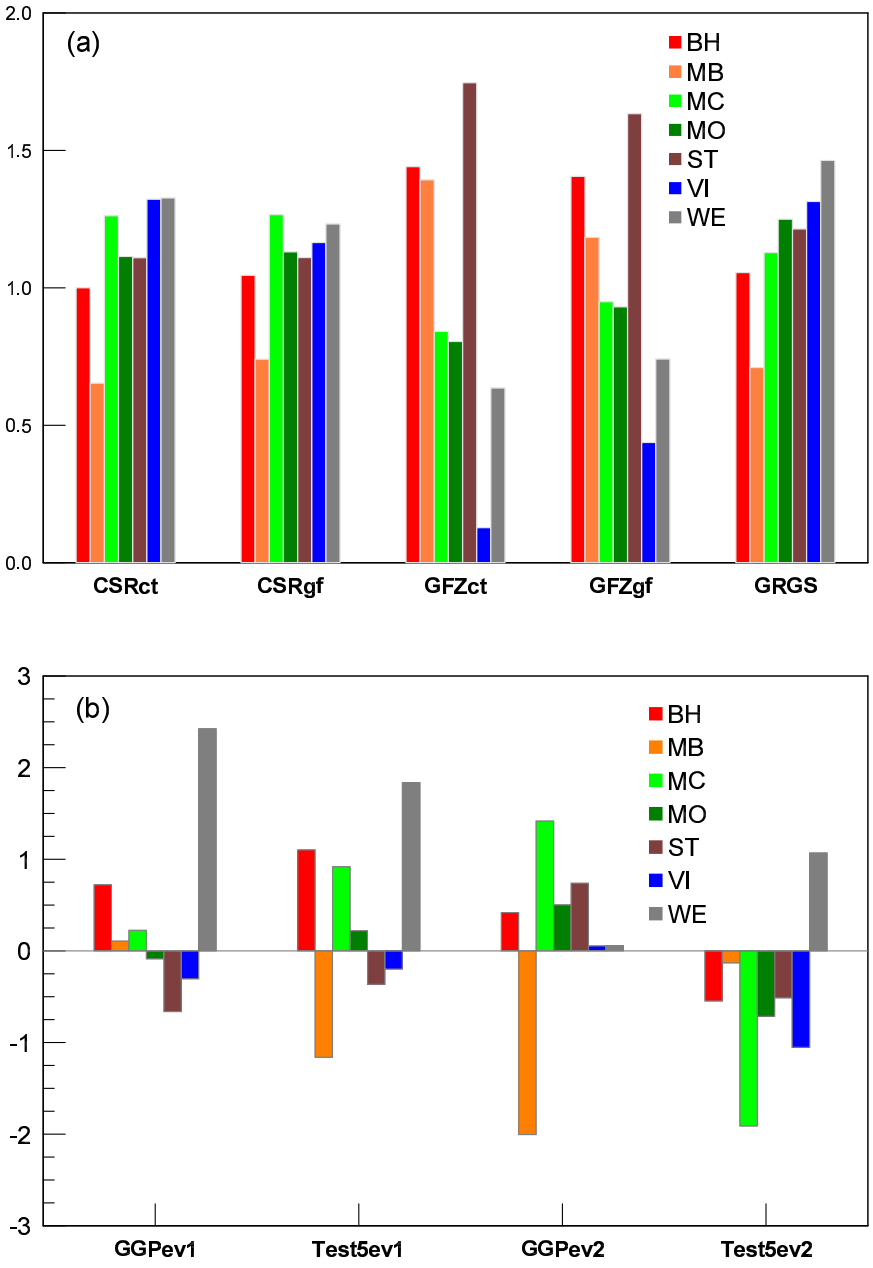

Figure S2. EVs for the GRACE and GGP solutions (a) EV1 for the GRACE solutions extracted from the gridded eigenvectors shown in Figures 13(a)-(e) and (b) the GGP values are from the 7-station solution in Figure 13(f) and the Test #5 solution in Table S2. Note that EV amplitudes are uncalibrated, unlike the principle components that are scaled to μGal . Thus while relative magnitudes between data sets for a mode can be usefully compared, magnitude differences between EV1 and EV2 are not meaningful.