J Geod (2012) 86:1165–1177
DOI 10.1007/s00190-012-0572-5

ORIGINAL ARTICLE

High-frequency signal and noise estimates of CSR GRACE RL04

Jennifer A. Bonin · Srinivas Bettadpur · Byron D. Tapley

Received: 13 April 2011 / Accepted: 14 May 2012 / Published online: 3 June 2012

Abstract A sliding window technique is used to create daily-sampled Gravity Recovery and Climate Experiment (GRACE) solutions with the same background processing as the official CSR RL04 monthly series. By estimating over shorter time spans, more frequent solutions are made using uncorrelated data, allowing for higher frequency resolution in addition to daily sampling. Using these data sets, high-frequency GRACE errors are computed using two different techniques: assuming the GRACE high-frequency signal in a quiet area of the ocean is the true error, and computing the variance of differences between multiple high-frequency GRACE series from different centers. While the signal-to-noise ratios prove to be sufficiently high for confidence at annual and lower frequencies, at frequencies above 3 cycles/year the signal-to-noise ratios in the large hydrological basins looked at here are near 1.0. Comparisons with the GLDAS hydrological model and high frequency GRACE series developed at other centers confirm CSR GRACE RL04's poor ability to accurately and reliably measure hydrological signal above 3–9 cycles/year, due to the low power of the large-scale hydrological signal typical at those frequencies compared to the GRACE errors.

Keywords GRACE · Gravity · High-frequency · Error · Signal-to-noise ratio

J. A. Bonin (✉)
College of Marine Science, University of South Florida,
140 7th Ave S., Saint Petersburg, FL 33701, USA
e-mail: jbonin@marine.usf.edu

S. Bettadpur · B. D. Tapley
Center for Space Research, University of Texas at Austin,
3925 W. Braker Lane, Suite 200, Austin, TX 78759-5321, USA
e-mail: srinivas@csr.utexas.edu

B. D. Tapley
e-mail: tapley@csr.utexas.edu

1 Introduction

Since the Gravity Recovery and Climate Experiment (GRACE) was launched in 2002, it has successfully measured the mean gravitational field of the Earth and its seasonal variations, resulting in time-variable gravity estimates with accuracies several orders of magnitude better than those constructed from previous geodetic data (Tapley et al. 2004). However, with almost no exceptions, studies using GRACE data have focused on low-frequencies: means, slopes, and annual signals. Partly, this is because those are the largest signals in most places and thus the most important. But also, it is an effect of the product definition and release cycle of the primary GRACE analysis centers. The Center for Space Research (CSR) and the GeoForschungsZentrum (GFZ) create non-overlapping estimates of the gravity field typically using 1 month of data. This monthly pattern of solutions, chosen to contain a sufficiently long series of observations to produce stable products, means signals with frequencies above 6 cycles/year cannot be resolved by CSR and GFZ GRACE products.

In this paper, we use solutions developed with an alternative processing technique. We apply sliding window filters to the Release 04 (RL04) CSR data to increase the apparent sampling rate and prevent aliasing at high frequencies. We also apply windows of differing lengths and relative weights during the estimation process. The windows chosen (see Sect. 2) allow frequencies higher than 6 cycles/year (cpy) to be measured.

The ability to manufacture signal at a given frequency does not mean that the signal is meaningful or correct. The higher frequencies may instead be dominated by error. Two ways of estimating the high-frequency noise in the GRACE solutions are given here. The first (Sect. 3.1) is an upper-bound noise estimate determined from the high-frequency signal

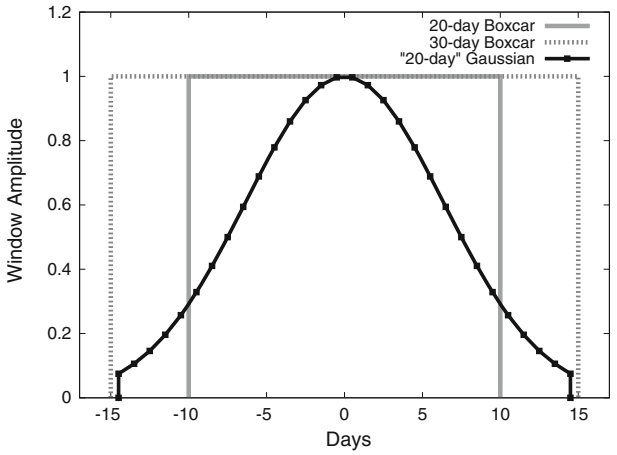
over a quiet part of the ocean. The second (Sect. 3.2) uses two other high frequency GRACE series (ITG-GRACE03 and the Goddard Space Flight Center 10-day 'mascon' solutions) as comparison cases. A measure of the divergence between the three high-frequency GRACE series marks a lower-bound on the average errors of the three series. These error estimates are verified using a hydrological model (Sect. 4).

Low signal-to-noise ratios (SNR) are found to be common above 3 cycles/year, so special care must be taken when drawing conclusions from high-frequency GRACE series. Though such series can easily be made, the day-to-day signal differences more often represent noise than true signal.

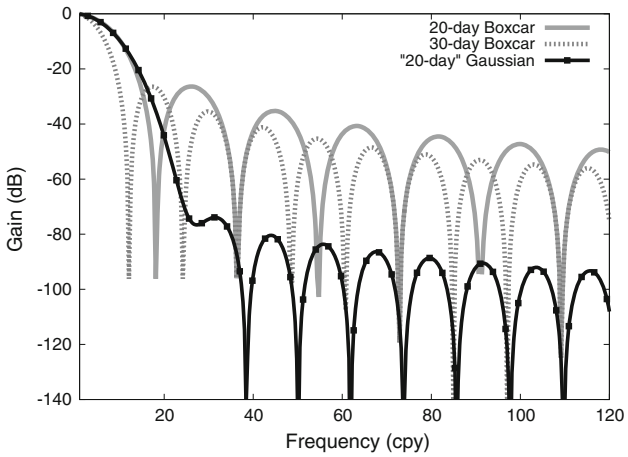
2 High-frequency windowed CSR data

CSR RL04 GRACE solutions are averaged products, with the amount of averaging determined by the number of days of data used in a single least-squares estimation process. The simplest way to achieve higher temporal resolution is to use fewer days in each estimation. Minimally acceptable solutions are possible down to a minimum of about 8 days of data (for temporally unconstrained solutions of degree/order 60). At that point, solution quality plummets (global and local measurements of RMS become unrealistic on an order-of-magnitude scale) as the number of orbital revolutions falls below twice the maximum spherical harmonic order (Bonin 2010; Kim 2000). In practical application, the lack of homogeneous groundtrack spacing and the presence of data gaps make it wise to increase the minimum solution length slightly. In this paper, solutions representing 10, 20, and 30 uniformly weighted days will be shown, capable of resolving signals up to 18, 9, and 6 cpy, respectively, without notable loss of gain.

In addition to changing the length of the averaging window, one can also change its shape. Estimating with uniform weights, as both CSR and GFZ typically do, is equivalent to applying a uniform or boxcar window to the satellite data. Such windows contain high-frequency artifacts in their transfer functions (see Fig. 1b), which allow signal at specific frequencies above the cut-off to propagate through to the windowed solution. This greatly reduces their value as averaging functions. Better transfer functions can be made by weighting the input satellite data by their order in time. Gaussian windows can be tailored to closely match the low-frequency spectrum of any uniform window, while greatly reducing the amplitude of the high-frequency sidelobes. This gives a cleaner cut-off between low frequencies (desirable) and high frequencies (averaged through) in the GRACE solution series, resulting in fewer aliasing problems due to windowing.



(a) Windows



(b) Transfer functions

Fig. 1 20- and 30-day boxcar windows compared to a "20-day" Gaussian window (15-day FWHM and 30-day length)

Through experimentation, we have determined two rules of thumb to match the low-frequency spectra of a Gaussian window with a previously determined boxcar window. Our chosen parameters for a Gaussian window are total window length (L_{Gaus}) and the full-width at half-maximum (FWHM), which is the width of the window (in days) at 50 % amplitude. Given a boxcar window of known length L_{Box}, we find that to match frequencies below the (non-overlapping) boxcar's Nyquist limit, a Gaussian window must have a FWHM given by:

$$\mathrm{FWHM} \sim \tfrac{3}{4}\, L_{\mathrm{Box}}. \tag{1}$$

In addition, to prevent sharp cut-offs at the edge of the Gaussian window from spoiling the clean high-frequency behavior of the window, a minimum window length is required. A rough approximation based on the width of the window is:

$$L_{\mathrm{Gaus}}^{\mathrm{minimum}} = 2\,\mathrm{FWHM}. \tag{2}$$

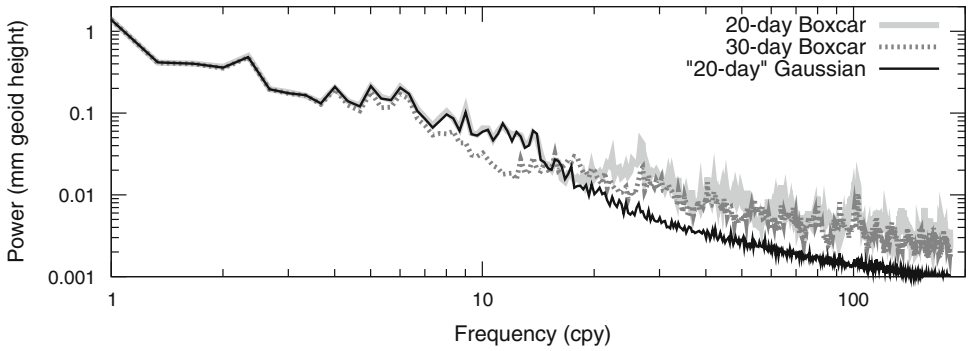


Fig. 2 Impact of windowing on GRACE's summed power spectrum

Experimentation has shown that increasing the Gaussian window length beyond twice the FWHM slightly decreases the window's gain at high frequencies but has no notable impact below the Nyquist frequency. Reducing the window length significantly below this limit results in boxcar-like high-frequency artifacts in the gain (Bonin 2010). We choose to use the minimum acceptable Gaussian length (twice the FWHM) in our experiments.

Using these two simple rules, Gaussian windows were created to best match the low-frequency spectra of 10, 20, and 30-day boxcar windows. Figure 1 shows the match for the 20-day case, using a Gaussian with a FWHM of 15 days and a length of 30 days. The nominal "monthly" 30-day window used in ordinary GRACE processing is shown for comparison. For simplicity, we will call such a window a "20-day" Gaussian window, where the 20 days represent the window's "effective" length, given by the nearest-matching boxcar window length.

To further reduce aliasing of the high frequencies, we use a sliding window technique. Solutions were made centered at each day during 2005–2007 and windowed with the above Gaussian technique. In places where GRACE data gaps exist (40 days total, with no more than 3 missing days in a row), linear interpolation is used between solutions.

The series used here have also been constrained in space (but not time) using regularization (Save et al. 2012). Only a weak constraint was used, so the north/south stripes are not completely removed from each solution. The regularization was instead designed to reduce their magnitude to an "acceptable" level, where large seasonal hydrological signals were made visible through the stripes. As such, post-processing remains valuable in most non-polar areas. Because of the different error characteristics of the unregularized solutions, the amount of constraint differs between the three windowed series. The regularization was roughly normalized to maintain a similar level of visibility of the seasonal hydrology, relative to the obvious stripes.

Figure 2 shows the power spectrum of the regularized GRACE results windowed with a 20-day boxcar window and the best-fitting "20-day" Gaussian window described previously. The power spectrum is defined as:

$$S(f) = \sqrt{\sum_{n=0}^{n\max}\left[\sum_{m=0}^{n}\left(A_{nm}^{f}\right)^2 + \left(A_{nm}^{-f}\right)^2\right]}, \tag{3}$$

where A_{nm}^{f} is the set of Fourier-transformed amplitudes of the GRACE time series, at frequency (f) and spherical harmonic degree (n) and order (m). The un-transformed set of complex spherical harmonic coefficients for each solution in time can be written as:

$$a_{nm}^{t} = a_e(c_{nm}^{t} + is_{nm}^{t}). \tag{4}$$

The factor a_e, the radius of the Earth, is used to convert into units of geoid height.

The comparison in Fig. 2 confirms both the low frequency match of the Gaussian and 20-day boxcar windows and the Gaussian's improvements at frequencies above the 9 cpy 20-day boxcar Nyquist limit. Gaussian windows tailored to match a 10-day boxcar and a 30-day boxcar were also constructed and applied to the 2005–2007 GRACE data for use in the error testing here.

3 Determination of CSR RL04 high-frequency error levels

Despite the recent increase in GRACE solutions released at sub-monthly sampling periods (i.e. Rowlands et al. 2005; Lemoine et al. 2007; Dahle et al. 2008; Kurtenbach et al. 2009), little analysis has been done to determine the high-frequency accuracy of those series. Formal covariance sigmas exist, but these only tell how well the observed data fits the final set of spherical harmonics, not how well those harmonics represent the true gravitational signal. In the past, most researchers have focused on long-term signals from GRACE, where it is appropriate to approximate the signal as the mean, trend, and annual (and maybe semi-annual) terms, and declare everything else "error". Wahr et al. (2006), Strassberg et al. (2007), and Swenson et al. (2008) all use this method of error estimation. This is sufficient for use with annual and longer signals, but clearly will not work when focusing on signals at the monthly and weekly level.

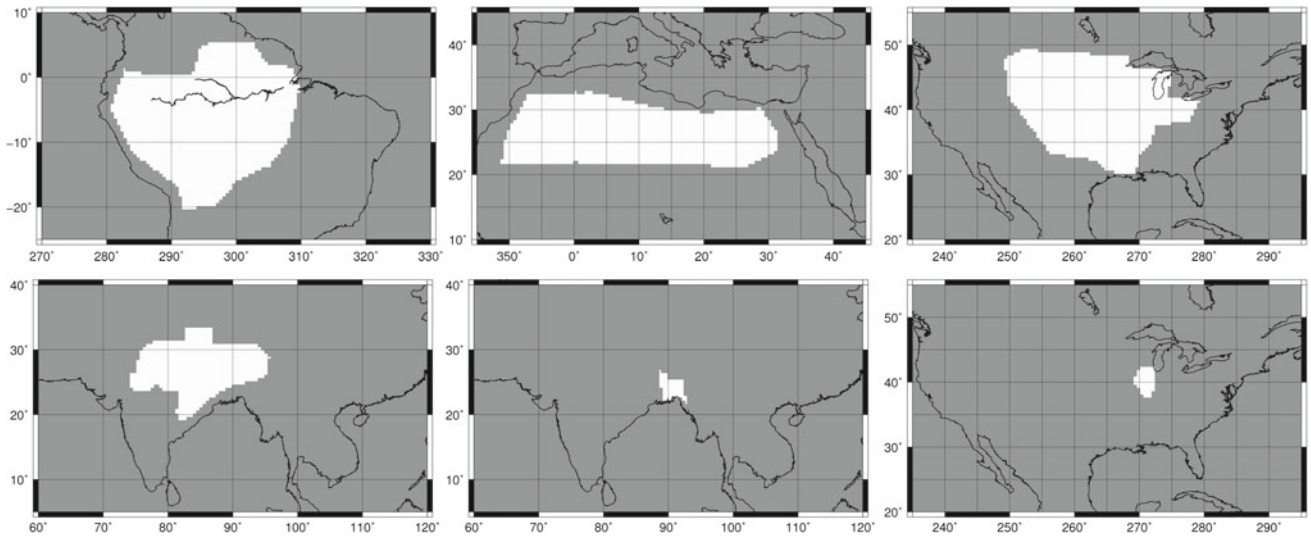


Fig. 3 Example basins used for averaging. In order of size: Amazon, Sahara, Mississippi, Ganges, Bangladesh, and Illinois

Others (Kurtenbach et al. 2009; Bruinsma et al. 2010; Liu et al. 2010, etc.) have used a more subtle method of error estimation, where the GRACE signal in the Sahara desert or oceanic areas is denoted as pure noise. Because the Sahara desert and similar areas are expected to have very little signal, this is perhaps a better approximation, and more suited to the needs here. Please note, however, that the errors associated with a basin average are not necessarily those of gridded data. Cancellation effects of a spatial average will tend to lower the error estimate of a basin, compared to the estimate of error within any individual $1° \times 1°$ grid cell, or the average of variances of many such cells.

In this section, two separate methods are used to determine the errors in the Gaussian-sliding-window, regularized GRACE series at sub-annual frequencies. Both techniques focus on the errors of basin-wide averages (using the technique given in Swenson and Wahr 2002), rather than point-based or global errors. The six basins used as examples are shown in Fig. 3. The four larger basins are regions known to be resolvable by GRACE. The two smaller ones are designed to test GRACE's ability to retrieve high-frequency signal over small spatial scales. Note that the Bangladesh basin is a subset of the Ganges basin, as Illinois is of the Mississippi.

The variances of the windowed regularized GRACE results in each basin are listed in Table 1, for the full signal and for only those frequencies faster than 3 cpy (period of 4 months). All basin averages have a 300-km Gaussian spatial smoothing applied to them, which increases potential leakage errors in to or out of the basin but reduces north/south striping as well as spatial cut-off effects from the sharp edges defining the basin. While the full-signal variabilities differ greatly between basins, high-frequency variances in all basins fall between 0.5 and 2 cm of water layer. Recall that the regularization has been "normalized" in a rough fashion, to keep the annual hydrological signal visible above the remaining stripes. This generally results in slightly reduced seasonal amplitudes for the shorter windows, where the unregularized solutions contain proportionally more stripes. (This signal damping effect occurs in any constrained series.) Window length has a moderate effect on the variance of basin averages, with shorter-windowed solutions containing more high-frequency variability due to their lesser averaging qualities. The remainder of the paper examines whether that greater variability comes from the determination of more higher frequency signal or more noise.

Table 1 Variance of CSR-GRACE windowed basin averages at all frequencies (300 km smoothing)

Window	Amazon	Sahara	Miss.	Ganges	Bang.	Illinois
For all frequencies						
"10-day"	11.04	1.34	3.83	7.31	12.15	4.07
"20-day"	11.46	1.18	3.77	7.73	13.24	4.17
"30-day"	12.13	1.15	3.62	8.14	14.14	4.24
For frequencies above 3 cpy only						
"10-day"	1.18	0.79	0.91	1.12	1.76	1.32
"20-day"	0.96	0.58	0.68	0.92	1.65	1.04
"30-day"	0.84	0.47	0.69	0.99	2.55	0.85

Units: cm water layer

3.1 Open-ocean estimate of errors

The first method of high-frequency error determination recognizes that the expected satellite measurement errors of GRACE are almost independent of longitude (see Figure 2 of

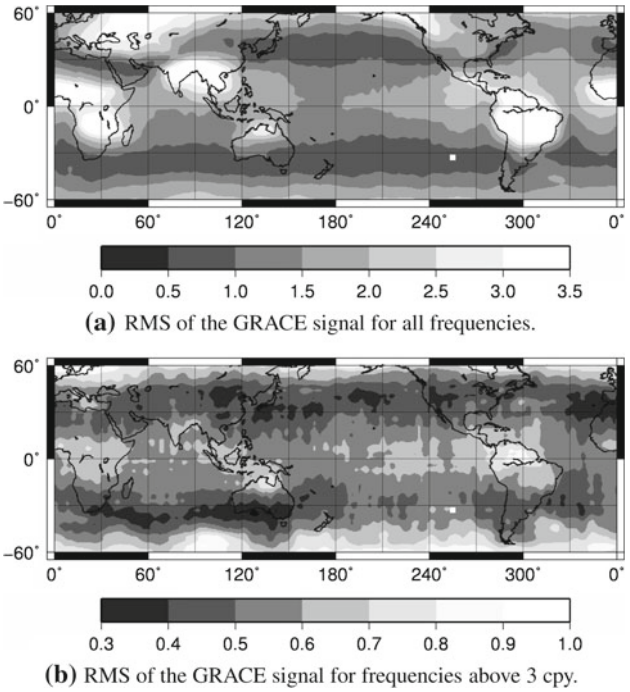


Fig. 4 The RMS of the "20-day" windowed GRACE signal. The *white square* in the southern Pacific denotes the central location of the open-ocean test area. Units are cm of water height. (Values outside the color range are shown in *white*.)

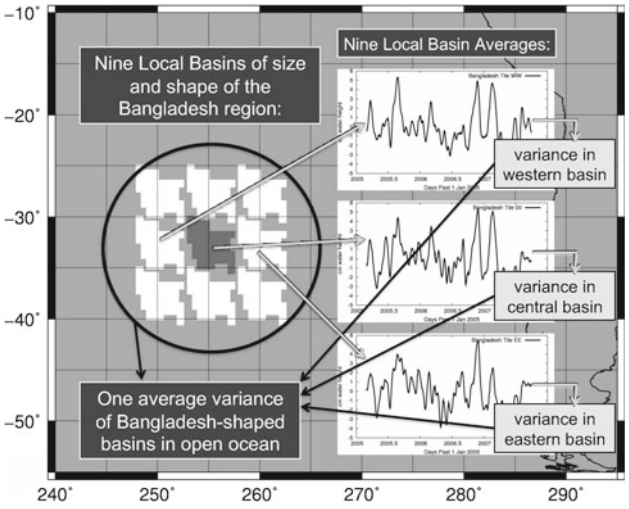


Fig. 5 Schematic of the open-ocean error estimation technique, using Bangladesh as an example

Wahr et al. 2006 or Figure 5.6 of Bonin 2010). Such measurement errors include both the random noise associated with the "formal sigmas" attached to each GRACE solution and the dominant correlated errors which show up as north–south "stripes". If only these errors exist, GRACE basin errors will be relatively constant across the globe as long as changes in latitude are recognized and the shape and size of the basin remain constant. In reality, additional errors due to leakage and inaccuracies in the background models also exist. These are neglected in this estimate, as we are focused on the satellite measurement noises instead, but they are significant in some areas.

For a first estimate of the GRACE measurement errors, a region of ocean off the coast of Chile is assumed to contain no true signal above 3 cpy, only noise. This place was chosen since it was far from land (where leakage would be more significant) and at a latitude close to that of most of the six test basins used here (except the Amazon basin). Near this region, the $1° \times 1°$ RMS of the "20-day" Gaussian-windowed GRACE series is 0.5–1.0 cm of water height (Fig. 4a), and 0.6–0.8 cm if only the high-frequency portion is considered (Fig. 4b). Results from the Sahara desert are similar to those shown here, but potentially corrupting hydrologic sources are closer and might introduce significant leakage errors.

We compute the open-ocean error bars for each region shown in Fig. 3, using the following process. First, the basin outline is placed at the center of the test region (at 33°S, 255°E) and a basin average computed from GRACE. The standard deviation of the highpass-filtered ($f > 3$ cpy) series is computed. The basin is then translated 5° in each compass direction (the white regions in Fig. 5) and the standard deviation of the GRACE series recomputed at those locations, resulting in eight additional measurements of variability in the ocean (though for large basins, the translation results in overlapping measurement sites and thus correlated variances). The average of the nine measurements is taken as the variability of a basin in the middle of this quiet ocean area. This averaging process has little effect on basins the size of the Amazon or Ganges, but larger effects on small basins where individual stripes may greatly affect the basin's signal. (A visually evident local stripe caused the 2–3 cm difference in the last strong peak of Fig. 5's 'eastern basin', relative to the other two basins, for example.) We interpret this quiet ocean variability as the GRACE uncertainty in a basin of the chosen shape and size.

Latitude will have an effect on the accuracy of this error estimate. The closer the true basin is to being located at 33°S north or south latitude, the more exact this will be. Figure 4b suggests that the latitude effect will be small, however, for basins between about 15°−45° north or south latitude. As all of our basins except the Amazon fall within this range, we choose not to apply an explicit latitude ratio. Based on Fig. 4b, we would expect the errors in the Amazon to be larger than we estimate due to this latitude effect. It is difficult to say how much larger, though, since much of the variability of Fig. 4b is caused by short-scale features, which tend to cancel out in a basin average.

We use the average open-ocean variance as a one-sigma error bar of the GRACE basin averages. Because this method defines all signal in the test region above 3 cpy as noise, it

Table 2 GRACE high-frequency measurement error variances (excluding leakage)

Window	Amazon	Sahara	Miss.	Ganges	Bang.	Illinois
Open-ocean errors						
"10-day"	1.05	1.08	1.23	1.22	1.50	1.50
"20-day"	0.80	0.85	0.88	0.94	1.21	1.20
"30-day"	0.63	0.68	0.68	0.75	0.97	0.98
Multi-series errors						
"10-day"	0.76	0.58	0.75	0.76	1.36	1.09
"20-day"	0.69	0.50	0.66	0.71	1.32	0.99
"30-day"	0.66	0.47	0.65	0.71	1.31	0.98

Units: cm water layer

is an upper-bound estimation of the satellite measurement errors. However, the total errors actually seen in each basin will be the sum of these measurement errors and errors from leakage in to and out of the basin. Leakage effects are not uniform across the globe, but will instead be dependent on the magnitude, shape, and time-variability of the geophysical signal in nearby regions, and can be quite large. Leakage error is not included in the estimates of this section, as it cannot be measured by any technique which geographically moves the basin. In regions with significant leakage errors (i.e. small basins near large external signals), the approximate measurement errors here may thus understate the total errors. Temporal aliasing, due to imprecision in the background models may also influence this approximation of errors. In addition, 3 years of GRACE data is not long enough to rule out unusually large errors or ocean signals from biasing the error results. These uncertainties may cause an over- or under-statement of the true errors and are one reason we also choose to make a second error approximation using a different method.

The open-ocean error variability based on the high-frequency ($f > 3$ cpy) oceanic signal is shown for six basins in the top part of Table 2. Shorter windows produce 24–45 % more error variability than longer ones, due to their greater high-frequency gain. (This is despite the regularization's tendency to remove more signal from the shorter, noisier windows.) Typical one-sigma error levels found from the open-ocean technique are 0.7–1.2 cm of water height.

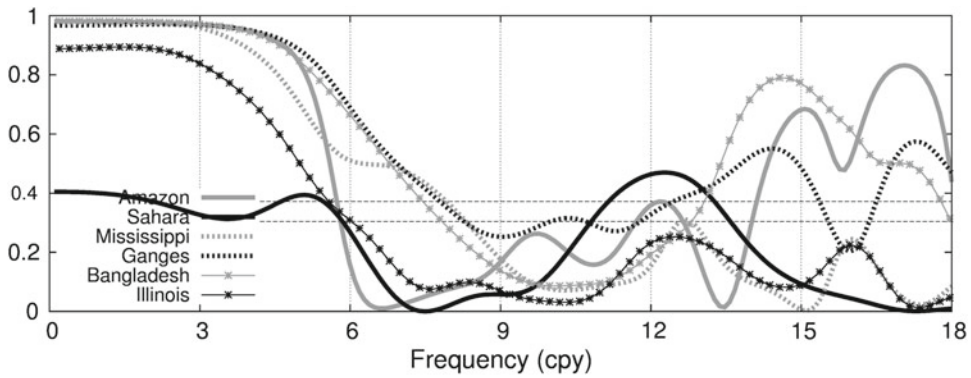
3.2 Divergence of alternative high-frequency GRACE series

A second method of high-frequency error determination uses several different GRACE series, created by different centers using different processing techniques. One is the regularized, 300-km smoothed, Gaussian-windowed CSR series described previously. The others are the University of Bonn's ITG-GRACE03 and the Goddard Space Flight Center's Mascon solutions. These series are based on the same raw GRACE data. It is assumed (perhaps incorrectly) that similarities imply true signal, while divergence surely means uncertainty.

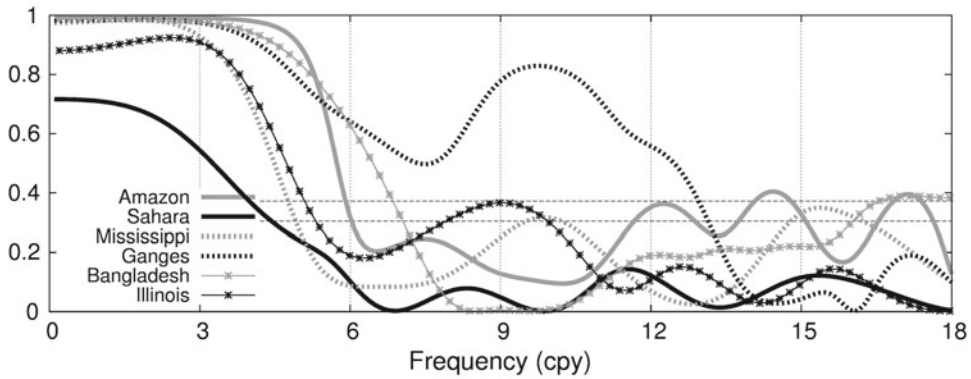
ITG-GRACE03 is a spline-based spherical harmonic field to degree 40, created by the Universität Bonn's Institut für Geodäsie und Geoinformation (Mayer-Gürr et al. 2009). To create solutions centered on any day they use quadratic spherical harmonic splines as basis functions in time. To approximate a solution occurring between nodes, data from three nodes spaced a total of 30 days apart are required. A degree autocorrelation for frequencies above 3 cpy shows that points less than 24 days from each other remain correlated due to overlapping data, making ITG's cutoff behavior most like a 24-day boxcar (Bonin 2010). The ITG-GRACE03 solutions are constrained in space and time during the estimation process such that they contain no power above degree 40 and minimal power above order 30. This results in a constrained series which needs no post-processing, so neither smoothing nor destriping has been applied. The ITG-GRACE03 data set ends in April 2007, so only the 2005–2006 portion is used here.

The second set of high-frequency comparison GRACE data comes from the Goddard Space Flight Center (GSFC) (Rowlands et al. 2005). Their non-overlapping solutions are released every 10 days. Rather than report their results in terms of spherical harmonics, they use "mass concentration blocks" or "mascons", are $4° \times 4°$ regions on the continents only. GSFC's estimation scheme uses only local KBR data and no GPS during the final solution of each mascon block. They apply a simultaneous temporal and spatial constraint to their mascon series, keeping mascon blocks neighboring in time or space more similar than they would otherwise be. While global comparisons are impossible, the mascon technique lends itself well to basin averaging. The mascon cells are subdivided into $1° \times 1°$ bins, to fit the basin mask resolution. A spatially weighted average of the Mascon data over the desired region gives a result comparable to a spherical-harmonic-based basin average. No post-processing is applied to this constrained data series.

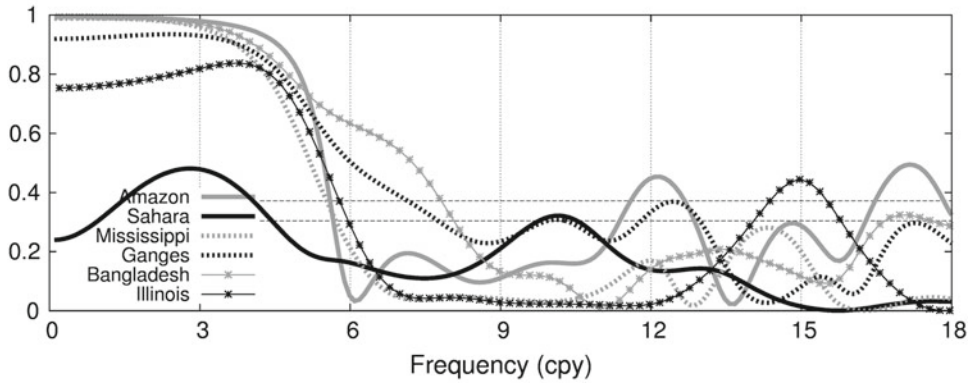
The frequency-based similarities and differences of the three centers' GRACE series can be better understood through their coherence spectra. If the phase difference of two series is constant at a given frequency, then their coherence at that frequency will be one. Series have coherences near zero at frequencies where they are uncorrelated or random. Figure 6 shows the basin average coherence of CSR, ITG, and the Mascons in pairs, using only every tenth CSR and ITG solution to match the 10-day spacing of the Mascons. The 90 % confidence level for the "20-day" Gaussian case, based on a Monte Carlo simulation, is at a coherence of 0.30, while the 95 % confidence level is at a coherence of 0.37.



(a) "20-day" Gaussian CSR vs ITG-GRACE03.



(b) "20-day" Gaussian CSR vs GSFC Mascons.



(c) ITG-GRACE03 vs GSFC Mascons.

Fig. 6 Coherence spectra of GRACE series over six hydrological basins. *Dashed horizontal lines* represent the 90 and 95 % confidence levels

These values have been marked with a dashed line on all coherence plots. Coherence between any pair is very high at the lowest frequencies, but drops sharply between 3 and 9 cpy. (Above 15 cpy and especially at much higher frequencies, coherency often increases. This is not necessarily a sign that the signals are similar at high frequencies, but rather that the small size of the signal (see Fig. 9) makes the technique used to smooth the periodograms more significant than the actual signal, when measuring coherence. The transfer function of the smoothing is identical for any input, resulting in meaningless high coherences when the power of the input signals is very low. Since longer windows repress more high-frequency signal, coherences made using the "30-day" Gaussian-windowed GRACE series approach 1.0 faster than those shown here, while the opposite is true of the "10-day" windowed version. The coherences at lower frequencies, where significant signal power exists, are unchanged.)

The source of the satellite and instrument errors must be identical between the three GRACE series. However, those noises may be convoluted differently based on the different centers' processing schemes and choice of background

models. In places where the three series disagree, at least two of the three must be wrong. We define multi-series error bars as the variance of the three series' basin averages away from their ensemble mean at each time, as given by:

$$\text{Error} = \left[\frac{1}{(3-1)T} \sum_{t=1}^{T} \bar{x}^{\text{ITG}}(t)^2 + \bar{x}^{\text{Mascon}}(t)^2 + \bar{x}^{\text{CSR}}(t)^2 \right]^{1/2}. \quad (5)$$

Here, T represents the total period of time considered. The denominator of $(3-1)$ is used to account for Bessel's correction around 3 series. For each series (ITG, Mascon, and CSR), $\bar{x}^*(t)$ represents a mean-removed version, defined by:

$$\bar{x}^*(t) = x^*(t) - \frac{1}{3}\left[x^{\text{ITG}}(t) + x^{\text{Mascon}}(t) + x^{\text{CSR}}(t) \right]. \quad (6)$$

Because the Mascon data exist only every 10 days, the other two types of GRACE series are reduced to contain only solutions centered at the same points (i.e. every tenth solution).

The divergence between the three GRACE series means error, but convergence does not guarantee a true signal, as systematic errors in modelling, techniques, and raw data could be consistent between the series. We chose three series created using very different processing methods to reduce the number of occasions when similarity exists because of consistent errors. Nonetheless, this method is likely to understate the true measurement errors, due to correlations between the series noises. We anticipate mistakes caused by differences in signal leakage between the series to be small, since the series are all filtered to similar spatial scales and defined across regions neighboring each basin. Thus, leakage effects should be similar in all cases and not add to this error measurement.

The bottom half of Table 2 lists the 1-sigma error bars found from these multi-series comparisons, which vary by basin. (See Bonin 2010 for the different solutions' basin average time series themselves.) Shorter windowings cause larger errors, but the effect is slight since the ITG and Mascon series use constant windowings. Typical one-sigma error levels found from the multi-series technique are 0.5–1.3 cm of water height.

3.3 Signal-to-noise ratios

The SNR over a region is a measure of the quality of the GRACE signal there. If the signal is the variance of the windowed CSR basin average and the errors are the one-sigma estimates discussed previously, the GRACE SNR can be estimated as:

$$\text{SNR}^{\text{basin}} = \frac{\text{var(GRACE basin average)}}{\text{var(errors)}}. \quad (7)$$

Table 3 Signal-to-noise ratios for all frequencies

Window	Amazon	Sahara	Miss.	Ganges	Bang.	Illinois
Open-ocean errors						
"10-day"	10.51	1.24	3.11	5.99	8.10	2.71
"20-day"	14.39	1.39	4.28	8.22	10.94	3.45
"30-day"	19.25	1.69	5.32	10.85	14.58	4.33
Multi-series errors						
"10-day"	14.54	2.31	5.13	9.63	8.94	3.73
"20-day"	16.71	2.34	5.71	10.89	10.01	4.20
"30-day"	18.35	2.46	5.57	11.46	10.79	4.33

Table 4 Signal-to-noise ratios for frequencies >3 cpy

Window	Amazon	Sahara	Miss.	Ganges	Bang.	Illinois
Open-ocean errors						
"10-day"	1.13	0.73	0.74	0.92	1.17	0.88
"20-day"	1.20	0.69	0.77	0.98	1.36	0.87
"30-day"	1.33	0.69	1.01	1.32	2.63	0.87
Multi-series errors						
"10-day"	1.55	1.38	1.22	1.47	1.29	1.21
"20-day"	1.40	1.14	1.04	1.30	1.25	1.05
"30-day"	1.27	1.01	1.07	1.40	1.94	0.87

Note that this definition includes the true errors as part of the "signal", which limits the minimum SNR value to about 1.0.

Table 3 lists the SNRs computed from the two types of errors and the full-signal variances of Table 1. Because the open-ocean errors are typically upper-bound estimates, the upper half of Table 3 will tend to underestimate the SNRs. The opposite is true for the multi-series errors on the bottom half of the table. The full-field SNRs are large, especially in basins with large annual signals. As many others have shown (Wahr et al. 2004; Swenson et al. 2006; Strassberg et al. 2007, etc.), GRACE is able to distinguish long-period signal, especially in larger basins. The only basin where the errors approach the level of signal is in the Sahara desert. The Mississippi and Illinois basins are noisy, but the primary signal should be readily apparent. In the Amazon, Ganges, and Bangladesh regions, low-frequency signal dominates the errors. Overall, longer windows give higher full-signal SNRs than shorter windows (typically a 50–200 % increase between the shortest and longest windows).

However, the primary interest of this paper is the short-period signals which are not determined by the monthly GRACE products. Table 4 shows the SNRs after a 3 cpy high-pass filter has been applied to the CSR windowed series. In this frequency regime, the variability of the signal is typically the same size as the open-ocean estimates of the errors. There is no clear dependency on window length and only limited

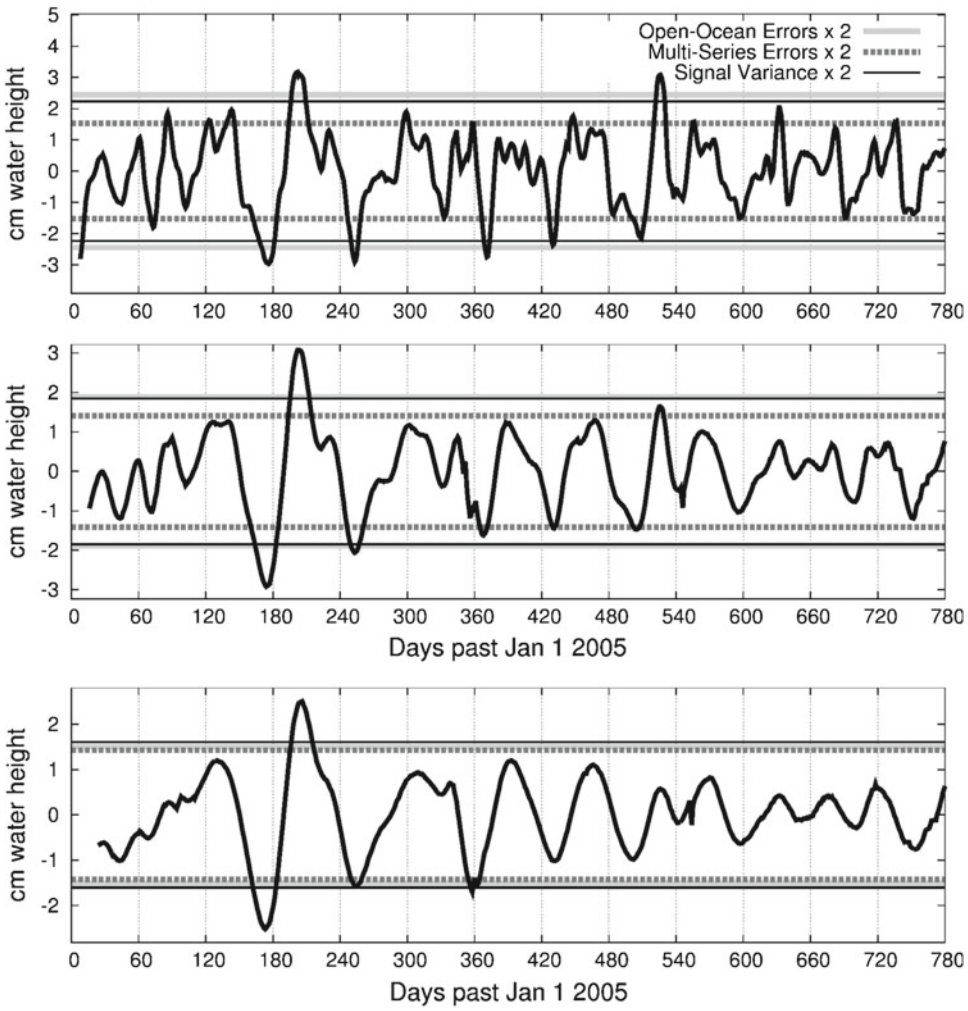


Fig. 7 CSR basin average of Ganges for $f > 3$ cpy. Series from *top* to *bottom* use Gaussian windowings with FWHM of 7, 15, and 23 days

impact from basin choice. The SNRs are in the 1–2 range when using the multi-series error estimate and 1–1.5 when using the open-ocean technique.

An example of the level of estimated error compared to the actual high-frequency signal is given in Fig. 7, for the Ganges basin. The top plot shows the "10-day" Gaussian-windowed CSR GRACE solution, the middle shows the "20-day", and the bottom shows the "30-day" solution. On each are sketched the two-sigma (95 % confidence) error levels about the mean, as well as dotted black lines denoting the equivalent level of variability in the actual basin average (twice the GRACE high-frequency RMS).

This basin was picked intentionally, since it and the overlapping Bangladesh basin are the only locations from 2005 to 2007 where we have seen the GRACE signal clearly emerge from behind both error bars (days 160–220). This is part of a 4-month-long signal, the longest possible given the 3 cpy high-pass filtering. In all other basins we have seen, and at all other times even in this area, most of the high-frequency signal lies within both 2-sigma error bars, and almost all lies within the larger, open-ocean line (Bonin 2010). Overall, high-frequency signal seems to be buried in noise. In addition, though the amplitude of the high-frequency GRACE signal retrieved grows as the window length decreases, the errors grow proportionally. Shortening the window length neither improves nor degrades overall signal recognition.

Three general types of basins emerge from the SNR calculations. The first includes basins like the Sahara desert, where even the annual signal is on the level of the satellite errors. In the majority of basins, like the Amazon and Mississippi, the low-frequency signal is clearly discernable by GRACE, but higher-frequency signal is near the level of the satellite errors. The third type of basin is one like the Ganges region, where even above 3 cpy, some signal is recognizable over the noise. In these areas, not all of the GRACE high-frequency signal is necessarily true, but the largest peaks or dips are likely to be.

GRACE's ability to resolve true time-variable signal is not dependent on the frequency of the signal, but on the amplitude of that signal relative to the local noise level. It is important to take into account the background noises, which are typically 1–3 cm in amplitude at the 2-sigma (95 % confidence)

level. Because long-term, annual, and near-annual geophysical signals tend to be large, they are represented well by GRACE. Faster-moving signals will only be visible over the errors if they are large (typically greater than 5 cm peak-to-peak change). Experience within these six basins and other hydrologic regions suggests that most of the large amplitude signals which GRACE sees typically happen on the time scale of several months, not faster.

4 Validation using the GLDAS hydrology model

The previous section concluded that basin-wide errors are at about the same level as basin-wide high-frequency signals. To validate this, we use NASA's Global Land Data Assimilation System (GLDAS) hydrologic model (Rodell et al. 2004). This version of GLDAS uses the Noah land surface model and is released as a $1° \times 1°$ grid every 3 h. The modeled terrestrial water storage (TWS) is the sum of the four soil moisture layers (maximum depth: 2 m), the snow-water equivalent (SWE), and the canopy water storage (Mirador Mirador Earth Science Data Search Tool 2010).

Unfortunately, the SWE component has a known problem where the snowpack in a few grid points incorrectly accumulates without melting for long periods of time (Matthew Rodell 2010, personal communication). This results in anomalously large values for TWS over localized areas. To correct this, the SWE grids were assessed during the local summer (August for the northern hemisphere and January for the southern hemisphere). Any grid cells which registered more than 10 cm of SWE during those months were considered suspect and their data eliminated for the entire span of time used. These holes were filled using a weighted average of the surrounding data, with Gaussian weightings and a halfwidth radius of 3°. The result is a SWE map with no unreasonable long-term accumulations of snow. This correction applies only to the SWE; the soil moisture and canopy water components have no smoothing or corrections applied.

Gaussian-windowed versions of the GLDAS TWS were created to match GRACE. The GLDAS series have been converted from their gridded format into 60×60 spherical harmonics, which has the effect of spatially smoothing them. As none of the GRACE series contain degree 0 or 1 terms, those have been removed from GLDAS.

Because the GLDAS model excludes surface water and ground water, it sometimes underestimates the amplitude of TWS (Niu and Yang 2006; Güntner et al. 2008). Lower seasonal amplitudes are common in regions where significant mass transport is conducted by surface water (Wahr et al. 2004). This results in a loss of signal amplitude in such regions (largest in the Amazon and Bangladesh, in this study) and occasionally a phase shift (seen here only in the Amazon) compared to GRACE. Despite this, basin-averaged GLDAS and the windowed GRACE series have a 92–99 % low-frequency correlation in all basins except the Sahara desert, where it is 75 % (upper part of Table 5). Windowing has no noticeable impact on these $f < 3$ cpy results.

Table 5 Correlations of the windowed GRACE series with similarly windowed GLDAS, and percent of model variance explained by GRACE

Window	Amazon	Sahara	Miss.	Ganges	Bang.	Illinois
Correlations for frequencies below 3 cpy only						
"10-day"	0.92	0.76	0.95	0.99	0.96	0.93
"20-day"	0.92	0.75	0.96	0.98	0.97	0.96
"30-day"	0.93	0.75	0.96	0.98	0.97	0.96
Correlations for frequencies above 3 cpy only						
"10-day"	0.34	-0.055	0.42	0.64	0.69	0.46
"20-day"	0.44	-0.077	0.57	0.70	0.70	0.47
"30-day"	0.56	-0.19	0.65	0.68	0.71	0.50
Percent of GLDAS variance explained by GRACE (freq > 3 cpy) (%)						
"10-day"	4	0	10	23	27	8
"20-day"	10	0	19	29	26	8
"30-day"	21	0	26	29	27	7

For faster-moving signals, the correlations are not as good (middle part of Table 5). Correlations are typically between 40 and 70 % (and effectively zero across the Sahara desert) and vary by location. The Ganges and Bangladesh regions show fairly good high-frequency agreement with the model. The Amazon, Mississippi, and Illinois basins are less well-correlated. Moreover, the percentage of the high-frequency GRACE variance explained by GLDAS is only 0–30 % in all basins (lower part of Table 5). Slightly more signal is sometimes explained in the longer-windowed cases, due to the decrease in GRACE errors, but the vast majority of the GRACE high-frequency signal is not seen by GLDAS.

The high-frequency comparison of GLDAS and GRACE is visible in a coherence spectrum of their basin averages. The cases shown in Fig. 8 use the "20-day" Gaussian windowing, but the drop-off characteristics of the coherence spectra do not change appreciably with window length (not shown). For frequencies lower than 4–5 cpy, coherence between the windowed GRACE series and GLDAS is above 0.8 in all but in the Sahara desert. With the exception of the Ganges basin, the coherence falls below the 90 % confidence level between 6 and 7 cpy, then remains low. In most places, the drop between high and low coherence is sharp, suggesting that either GLDAS or GRACE (or both) is no longer measuring much reliable signal above about 6 cpy. In the Ganges basin and to a lesser extent in Bangladesh, coherence remains relatively high through 12 cpy. In these areas, GRACE is seeing portions of the same signal that

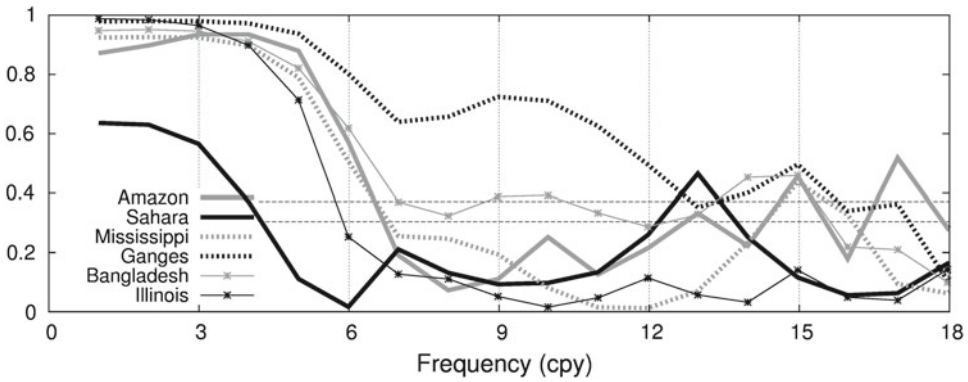


Fig. 8 Coherence spectra of CSR series versus the GLDAS model, over six hydrological basins. A "20-day" Gaussian window was used on both

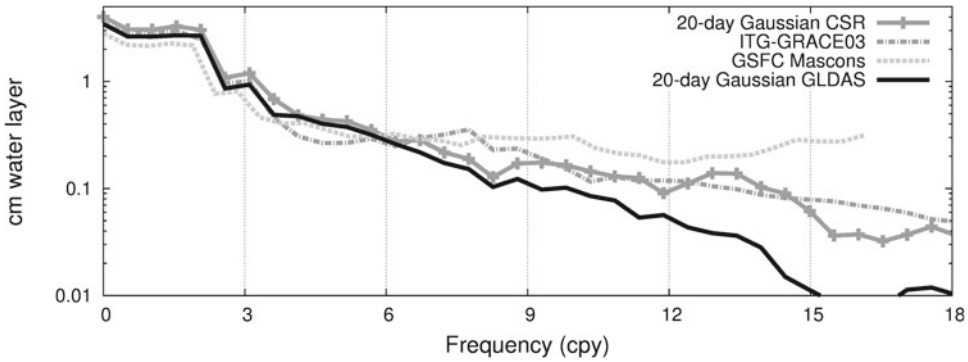


Fig. 9 PSD Comparison of GLDAS, CSR, ITG, and GSFC Mascon estimates over the Ganges basin

GLDAS models predict, suggesting there may occasionally be value in localized GRACE measurements through monthly frequencies, at least during limited portions of the time series.

These coherences with GLDAS hold broadly for the CSR series, ITG-GRACE03, and the GSFC Mascons (Bonin 2010). The declining correlation with the model between 3 and 9 cpy supports the previous finding that GRACE SNRs are high at low frequencies and decrease quickly as frequency increases. It is important to note, however, that the majority of the signal strength measured by either the GLDAS model or any of the GRACE series occurs at frequencies less than 3 cpy, and continues to decrease as frequency grows. Figure 9 shows the power spectral density (PSD) of the windowed GLDAS series as well as the comparable GRACE series over the Ganges basin. There is almost an order of magnitude drop in the hydrology model's power between 3 and 9 cpy, much of which is also seen by the GRACE series. This signal drop-off occurs because many higher-frequency hydrological effects (i.e. storms) occur over spatial scales too small for GRACE to resolve. As such, it might be said that GRACE and GLDAS are highly coherent in the frequency bands where both contain much energy, but grow incoherent as the power of the hydrological signal quickly decreases with frequency. Surprisingly, there is no observable correlation between the absolute magnitude of the GLDAS PSD between 3 and 12 cpy and the correlation or coherence of GRACE and GLDAS at those same frequencies. For example, the coherence of GLDAS to CSR (Fig. 8) is high at 6–12 cpy in the Ganges basins and low at 6–12 cpy in the Amazon basin. But the drop in the GLDAS PSD amplitude (such as Fig. 9) between 6 and 12 cpy is similar in both basins. The coherence cannot be predicted based on the strength of the hydrological model.

5 Conclusions

With the use of a sliding window technique, it is easy to create high-frequency GRACE series of equal quality as the official monthly series. The complicated question is whether the new, higher-frequency signal is real. Two separate estimates of the GRACE RL04 errors were computed, resulting in 2-sigma (95 % confidence) error levels of 1–3 cm. Though there are exceptions, SNR values tend to fall into two categories based on frequency. Low-frequency hydrology signals ($f < 3$ cpy) are generally visible and reliable, whether computed from ordinary monthly GRACE solutions or more frequent windowed GRACE series. High-frequency signals are generally unreliable, though in specific areas and at limited times real variability may be visible. Frequency-based coherence with the GLDAS model and non-CSR GRACE series supports this, falling from around 90 % at 3 cpy to 20 % at 6–9 cpy, in most of the basins considered.

Windowing with 10-, 20-, and 30-day Gaussian filters does not appreciably change the GRACE high-frequency SNR or GRACE's correlation with GLDAS. At a single frequency, applying any transfer function during the GRACE processing will multiply both signal and error by an identical gain, resulting in no change of SNR at that frequency. Even when summed over the 3–18 cpy frequency band where the windows' transfer functions noticeably differ, a shorter window will retrieve both more signal and more noise than a longer window, resulting in similar SNRs. For a similar reason, high-frequency SNR values are insensitive to basin size. There is more estimated error in the smaller basins and shorter windows, and also more retrieved signal.

GRACE SNRs are low above 3 cpy in frequency not because of frequency resolution difficulties, but because few natural signals are large enough on GRACE-sized spatial scales to clear the error levels inherent in GRACE RL04. In basins with low signal (like the Sahara desert), the noise makes it difficult to discern even the annual signal. In basins where significant 3–9 cpy signal is present (like the Ganges basin), the CSR windowed series, ITG-GRACE03, and GSFC's mascons can all see it, as can the GLDAS model. However, as the amplitude of hydrological spectra typically decreases with frequency, true signal is harder to resolve at higher frequencies. Across the same range of frequencies (3–9 cpy) where coherency between GRACE series and/or GLDAS falls from 90 to 20 %, the PSD of the GLDAS hydrological model falls by a factor of 5–10, depending on the basin. In the hydrological basins looked at, all recognizable signal occurred in the 0–9 cpy frequency range, and most of it was in the 0–3 cpy range. This lack of significant signal results in poor SNRs, poor series correlations, and generally poorly trusted behavior in the windowed GRACE series for most applications.

A sliding window series might still be used for high-frequency analysis in specialized situations. If a signal is expected at a specific frequency with a specific spatial pattern to it, modeling and fitting to that signal might bring it out from behind the GRACE errors. For example, it might be possible to tease out information of the aliasing of the diurnal/semidiurnal tides, or other longer-period, non-aliased tides (the M_f fortnightly (26.7 cpy) tide was demonstrated to be a likely candidate possible by Bonin 2010). Another potential use of the high-frequency GRACE products could be to pinpoint one-time extreme events, particularly large coseismic events. These examples are outside the scope of this paper, but they demonstrate that high-frequency GRACE series might still have use in studies where the signal is larger than the corresponding noise at one particular time or one particular frequency.

However, such signals have been difficult to find on a recurring basis, either on land or in the ocean. The gravity signal does change significantly between the beginning and end of a month, but most of the change is the continuation of long-term (usually annual) signals. Creating a low-pass fit to monthly data and interpolating to sub-monthly time steps determines this part of the sub-monthly change adequately. Improvements in GRACE-based hydrology coming from the use of sub-monthly sampling must be more than this. Unfortunately, such improvements are not typically seen in the sliding-windowed CSR RL04 cases, regardless of the window length picked. Nor are they found in the ITG-GRACE03 or GSFC Mascon GRACE series. Hopefully, as refinements to the GRACE processing technique continue, error levels will drop further and more sub-monthly hydrological signal will become visible. Due to the decaying spectral power of the hydrological system on spatial scales of appropriate size for GRACE, this will become a case of diminishing returns, however, where ever more effort will be needed to pick up ever smaller signals at ever higher frequencies.

Until processing improvements are made, it is critical that users of most "daily" GRACE series realize that weekly or even monthly scale variability in the series is primarily noise. Similarly, little should be read into month-to-month changes in the official monthly sampled series where they diverge from a low-frequency fit. As a general rule of thumb, CSR GRACE RL04 does not reliably separate signal from noise unless the signal is more than 5 cm in amplitude and/or propagates for at least 4 months. Signal above 3 cpy should be looked at with caution, and signal above 9 cpy doubly so. Without external knowledge of the physical system, only the long-period and near-annual frequencies are reliable.

Acknowledgments This research was supported by NASA under Contract NAS5-97213. The Texas Advanced Computing Center at the University of Texas at Austin provided high performance computing and database resources that greatly contributed to the research results reported here. The authors wish to thank Himanshu Save at the Center for Space Research for his regularization of the windowed GRACE data. The authors wish to thank Don Chambers at the University of South Florida College of Marine Science for his assistance in editing this paper.

References

Bonin JA (2010) Improving the observation of time-variable gravity using GRACE RL04 Data. Dissertation, University of Texas at Austin, Austin

Bruinsma S, Lemoine JM, Biancale R, Valès N (2010) CNES/GRGS 10-day gravity field models (release 2) and their evaluation. Adv Space Res 45(4):587–601. doi:10.1016/j.asr.2009.10.012

Dahle C, Fletchtner F, Kusche J, Rietbroek R (2008) GFZ EIGEN-GRACE05S weekly gravity field time series. Grace Science Team Meeting, San Francisco

Güntner A, Werth S, Petrovic S, Schmidt R (2008) Calibration analysis of the global hydrological model WGHM with water mass variations from GRACE gravity data. GRACE Science Team Meeting, San Franscisco

Kim JR (2000) Simulation study of a low-low satellite-to-satellite tracking mission. Dissertation, University of Texas at Austin, Austin

Kurtenbach E, Mayer-Gürr T, Eicker A (2009) Deriving daily snapshots of the Earth's gravity field from GRACE L1B data using Kalman filtering. Geophys Res Lett 36:L17102. doi:10.1029/2009GL039564

Lemoine JM, Bruinsma S, Loyer S, Biancale R, Marty JC, Perosanz F, Balmino G (2007) Temporal gravity field models inferred from GRACE data. Adv Space Res 39:1620–1629. doi:10.1016/j.asr.2007.03.062

Liu X, Ditmar P, Siemes C, Slobbe DC, Revtova E, Klees R, Riva R, Zhao Q (2010) DEOS mass transport model (DMT-1) based on GRACE satellite data: methodology and validation. Geophys J Int 181(2):769–788. doi:10.1111/j.1365-246X.2010.04533.x

Mayer-Gürr T, Eicker A, Ilk KH (2009) ITG-Grace03 Gravity Field Model. Universität Bonn Institut für Geodäsie and Geoinformation. http://www.geod.uni-bonn.de/itg-grace03.html (accessed April 11 2009)

Mirador Earth Science Data Search Tool (2010) Goddard Earth Sciences Data and Information Services Center. http://mirador.gsfc.nasa.gov (accessed 21 July 2010)

Niu GY, Yang ZL (2006) Assessing a land surface model's improvements with GRACE estimates. Geophys Res Lett 33:L07401. doi:10.1029/2005GL025555

Rodell M, Houser PR, Jambor U, Gottschalck J, Mitchell K, Meng CJ, Arsenault K, Cosgrove B, Radakovich K, Bosilovich M, Entin JK, Walker JP, Lohmann D, Toll D (2004) The global land data assimilation system. Bull Am Meteorol Soc 85:381–394. doi:10.1175/BAMS-85-3-381

Rowlands DD, Luthcke SB, Klosko SM, Lemoine FG, Chinn DS, McCarthy JJ, Cox CM, Andersen OB (2005) Resolving mass flux at high spatial and temporal resolution using GRACE intersatellite measurements. Geophys Res Lett 32:L04310. doi:10.1029/2004GL021908

Save H, Bettadpur S, Tapley BD (2012) Reducing errors in the GRACE gravity solutions using regularization. J Geod. doi:10.1007/s00190-012-0548-5

Strassberg G, Scanlon BR, Rodell M (2007) Comparison of seasonal terrestrial water storage variations from GRACE with groundwater-level measurements from the high plains aquifer (USA). Geophys Res Lett 34:L14402. doi:10.1029/2007GL030139

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. doi:10.1029/2007WR006057

Swenson S, Wahr J (2002) Methods for inferring regional surface-mass anomalies from gravity recovery and climate experiment (GRACE) measurements of time-variable gravity. J Geophys Res 107(B9):3. doi:10.1029/2001JB000576

Swenson S, Yeh PJF, Wahr J, Famiglietti J (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

Tapley BD, Bettadpur S, Ries JC, Thompson PF, Watkins MW (2004) GRACE measurements of mass variability in the earth system. Science 305:503–505. doi:10.1126/science.1099192

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

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