Response to Matt King's review:

Thank you so much for your helpful review. Sorry it's taken me so long to get back to you on it. Let me see if I can respond to each of your main points.

1.) "Given the duplication of the C2,0 term, should not it be excluded from the comparison to GRACE?"

I think this is a really good point, since the influence of C20 is so large at the poles. You're right that using very similar C20 terms for GRACE and the Cheng SLR series might bias them toward each other for reasons that have nothing to do with GRACE itself. However, because the C20 terms are such a big part of the final signal, I didn't really want to produce this paper by totally excluding it. Instead, to answer your question, I decided to test what the impact of removing it was, to see if it was reducing the divergence I see between GRACE and the Cheng SLR series.

So I recreated each of the three main series (GRACE, Cheng 5x5 SLR and Sosnica 10x10 SLR) and totally omitted the C20 terms, then inverted each and took a look at the time series. If the C20 term was causing falsely alignment with GRACE, I would see a larger divergence between GRACE and the Cheng series, in which case, my paper would require revising.

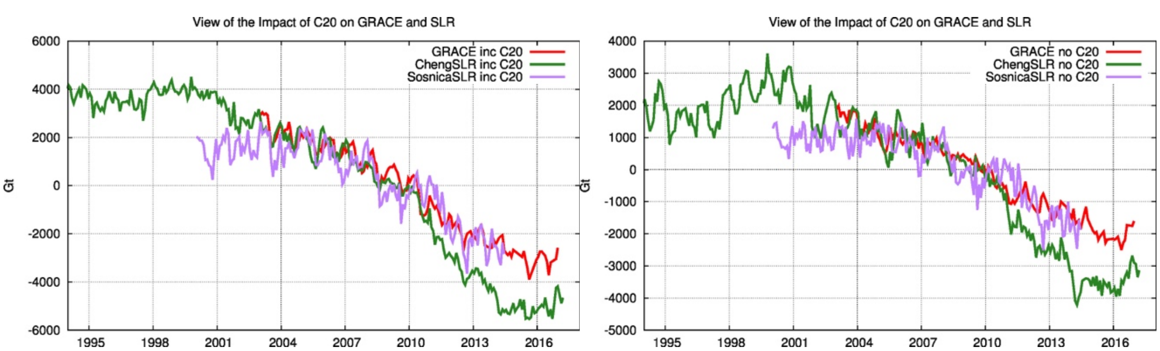


Figure 1: Left-hand image is the inversion with C20 included. Right-hand image is the inversion with C20 totally removed.

However, I see no notable changes in terms of divergence. There are three main effects of removing the C20 terms. First, the overall trend of all three of the series dropped like a rock. (No surprise, given the geometry of the situation.) Second, the month-to-month jitter in all three of the series changed. Third, most oddly, removing the C20 term from the Cheng series produced a large, visible annual signal before about 2007. The other series (including GRACE, using a similar C20) didn't show this impact. So that's bizarre. I assume that the C20 term in the Cheng series is coupled with some other term, to produce this (which wouldn't especially surprise its creators, since they're aware of the general coupling between harmonics caused by a barely solvable problem).

In any case, there was not any significant change in the interannual signal divergence. So in practice, the replacement of the GRACE C20 should bias GRACE towards the Cheng SLR series doesn't seem to have any major effect on the part of the spectrum that I'm worried about. That's a relief.

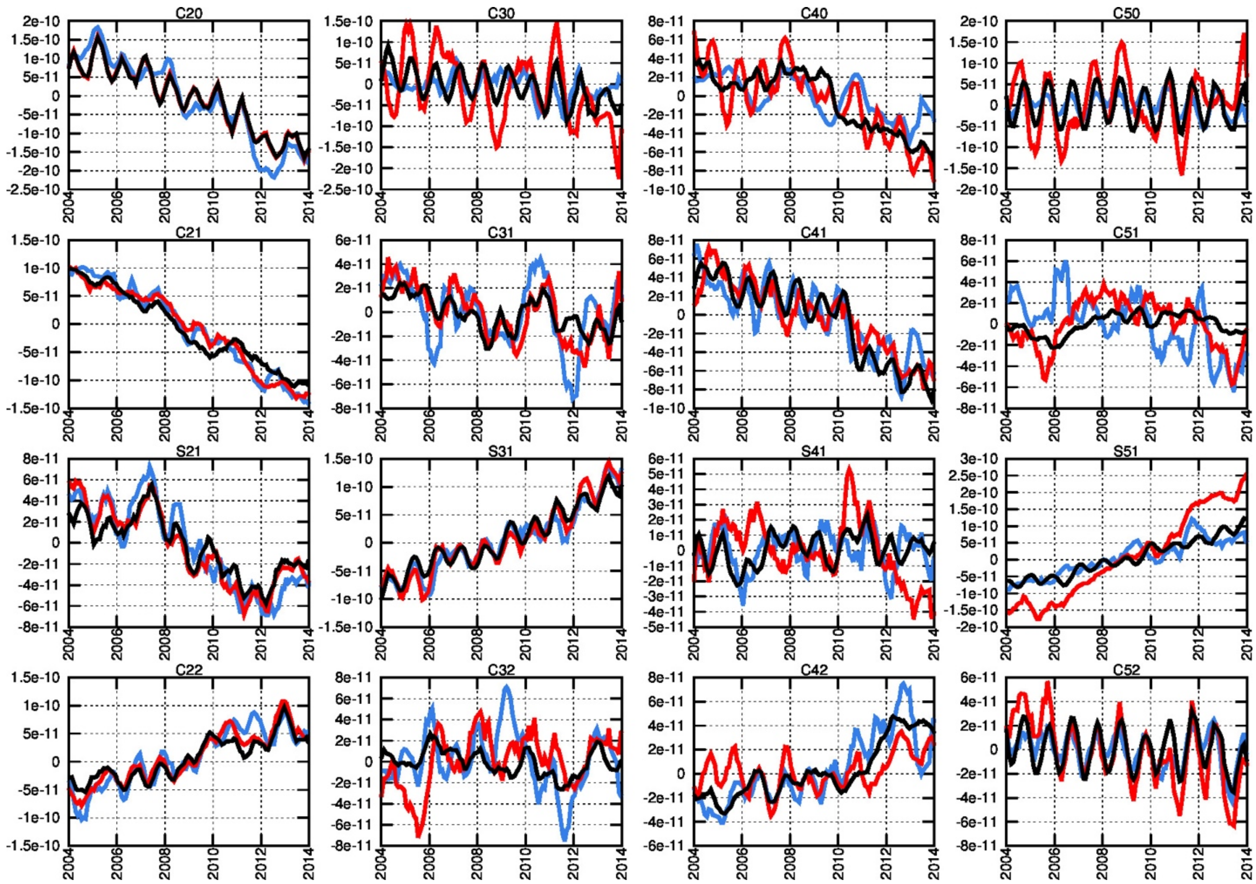
I have created the following commentary for the final version of the paper, briefly discussing this:

> We did consider the impact of replacing the GRACE C_{20} term with that from a series related to the Cheng 5x5 SLR data. To test whether this unfairly biased the Cheng 5x5 SLR results toward GRACE, we removed the C_{20} terms completely from all of the GRACE and SLR series, then inverted each of them again. Removing the impact of the equatorial bulge did greatly reduce the trend of each Greenland+Antarctica inverted series, but it did not significantly impact the interannual differences between GRACE and any SLR series. We thus conclude that the replacement of GRACE's C_{20} values is not a large contributing factor to these results.

Again, I want to thank you for this idea, since it was certainly a troublesome possibility.

2.) "It would be good to see in the supplement (degree, order)-specific time series comparisons for GRACE and SLR to see where the differences occur."

I have created this visual comparison (up to deg/ord 5) and will add it in the appendix for the final paper. For your immediate edification, here they are:


C20
C30
C40
C50
C21
C31
C41
C51
S21
S31
S41
S51
C22
C32
C42
C52

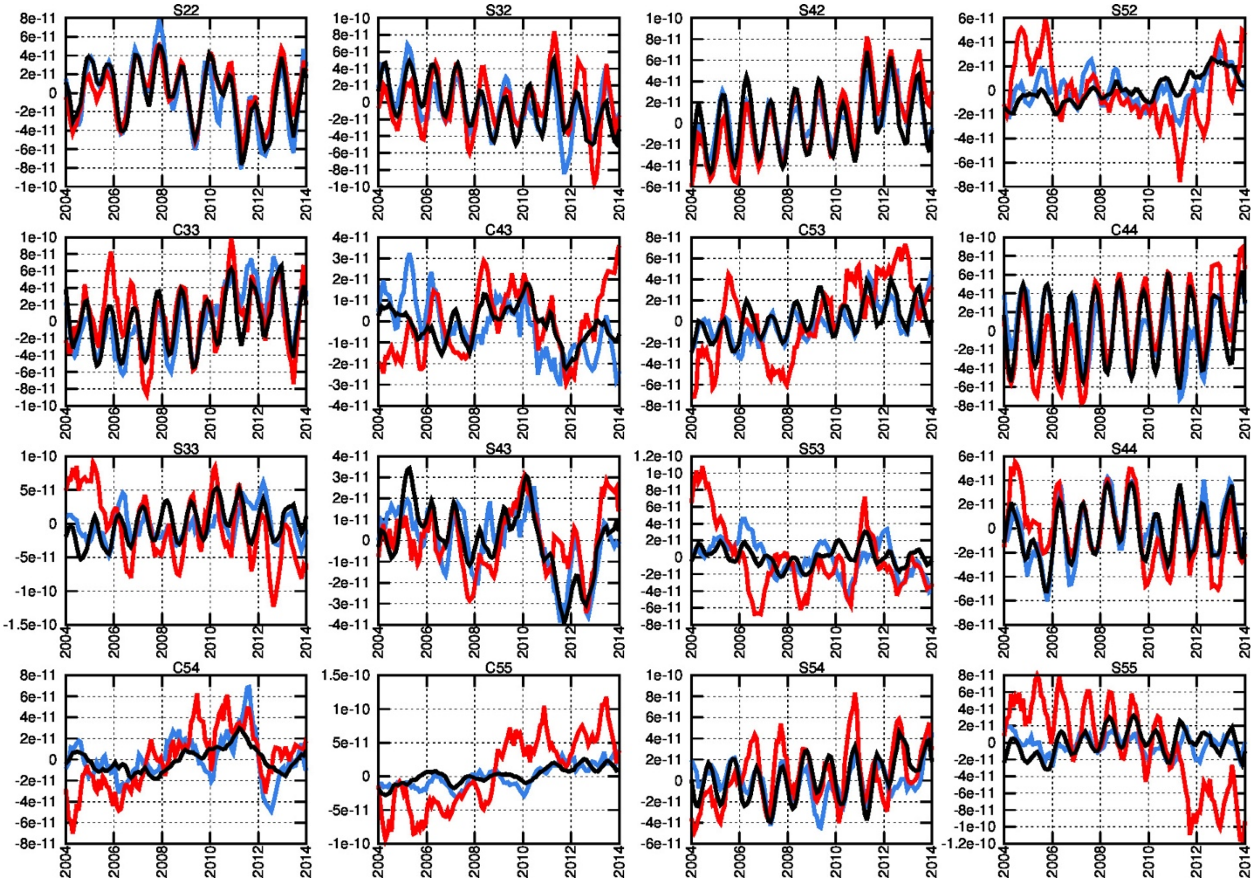


Figure 2: GRACE is in black, Cheng's SLR is in red, Sosnica's SLR is in blue.

3.) You are correct that Figure 1 in the main document (the percent variance explained) was given in terms of the proportion of the signal from 0 to 1, not as a real percentage. That's been changed, so the values in the figure go from 0 to 100%. I'll update for the final version of the paper. Thanks.

4.) "I wasn't sure if autocorrelation was really treated correctly - the authors assume it is diminished by 13-month averages and reduce the degrees-of-freedom appropriately but I think the assumption the series is white noise after this averaging (ie, uncorrelated). Exploration of the noise model by examining the spectra and fit of various noise models could be worth considering although I see an argument here that an exact specification of uncertainty is not the key message but the bias magnitudes…"

> The error statistics given for our own solutions (ie: table 1) contain the assumption that the residual solution (after the mean, trend, annual, and semiannual terms are fit and removed) still contains a correlating signal. We assume an AR-1 method and estimate the errors based on that assumption. Based on the paper you recommended, this seems to be a reasonable assumption for Antarctica, and presumably Greenland as well.

I think, though, that maybe the statistic you were really worrying about was the comparison with the IMBIE data? Referring to that, we wrote: "The uncertainty here is based on the variance of the smoothed residuals about the fit, but also accounts for temporal correlation due to the 13-month smoothing already applied to the IMBIE data. This reduces degrees of freedom from 186 to 14, so inflates the error from the least squares fit by sqrt(186/14)."

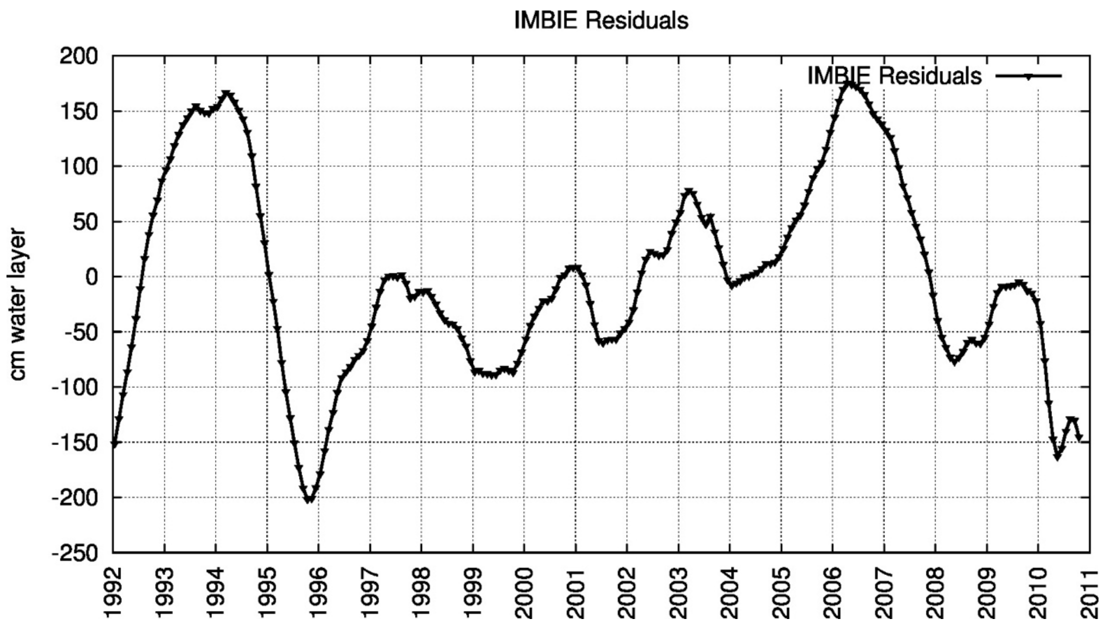


Figure 3: IMBIE inversion over GL+Ant, after removing a seven-parameter fit.

The sqrt(186/14) assumption described here only refers to the treatment of the IMBIE data, and is based on their claims of a 13-month temporal smoothing. A quick look at the IMBIE data after removing the acceleration, trend, annual, etc (above) shows a 3-4-year quasi-periodicity remaining, so I definitely agree that the signal left isn't actually white noise. To test whether the degrees-of-freedom reduction we used (based on a 13-month averaging) is "close enough", I computed the autocorrelation of the monthly IMBIE data (black line below). At a 13-month lag, the autocorrelation is 0.2. It actually crosses zero at about 16 months. I also checked to be sure that decoupling the "monthly" data points from the neighboring ones by using only every 12^{th} point (colored lines) doesn't impact the autocorrelation significantly – and it doesn't.

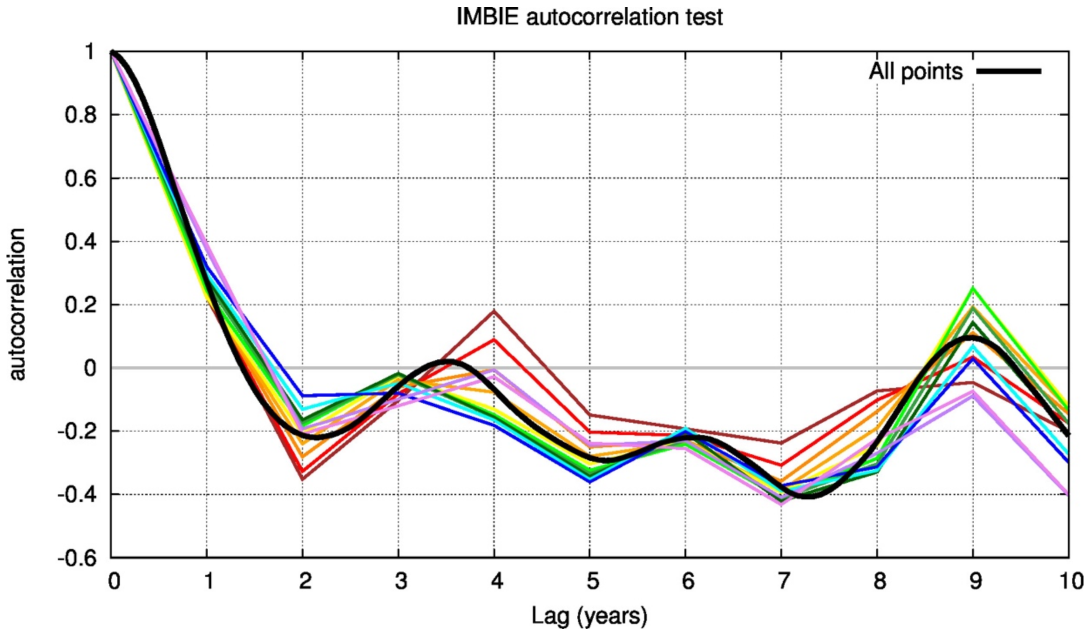


Figure 4: The colored lines are the autocorrelation using only every 12th point, which should not have been made dependent on each other due to the temporal smoothing. The black line is the autocorrelation using all the points.

So, technically, we should probably use a ratio of sqrt(186/11.6), leading to a weighting of the errors of 4.0 rather than 3.6. But since increasing the IMBIE errors won't impact the overall results of the paper, I doubt the detail is worth explaining the added complexity to readers (as you noted). If you feel strongly about this, though, we can change it.

(Overall, by the way, I agree with the paper you recommended: we too often assume that everything other than the mean, trend, annual, semiannual, and tidal aliases is "noise". Some sort of assumption for a low-frequency correlation seems more logical to me. Thanks for the link to the paper. I found it pleasantly clear to read, for a stats paper. Nice work.)

Thank you again for your excellent review. We appreciate the help!

-- Jennifer Bonin