Earth Syst. Dynam. Discuss.,
doi:10.5194/esd-2016-10-AC10, 2016
© Author(s) 2016. CC-BY 3.0 License.

Interactive comment on "Comment on "Scaling regimes and linear/nonlinear responses of last millennium climate to volcanic and solar forcing" by S. Lovejoy and C. Varotsos" *by* K. Rypdal and M. Rypdal

K. Rypdal and M. Rypdal

kristoffer.rypdal@uit.no

Received and published: 21 June 2016

We thank the reviewer for illuminating comments, which had been incorporated in the revised paper. Below follows a detailed response to the comments.

[1] On the hypothesis testing for subadditivity. We perform two tests for the subadditivity of the ZC model. One, described in Sect. 2.2 and 2.3, is a different test from the one performed by L&V, since it includes internal variability (which L&V do not). The reviewer is quite correct in pointing out that this test does not represent a direct rebuttal of the L&V test, if one assumes that internal variability can be ignored. However, in Sect. 2.4

we repeat the L&V test (i.e., without including internal variability) and present the result in our Fig. 2. We don't find the factor 1.5 claimed by L&V on the longest time scales, so this figure presents a rebuttal of the L&V test. See also our response to Point [7].

On the language - personal and subjective. Lines 164-165. We think it is important to point out that the approximation in question is unnecessary. This is neither personal nor subjective. Approximations are justified if they simplify things, and do not introduce biases. In this case, the approximation does not simplify anything (computations are just as easy without it), and it introduces a bias towards subadditivity. In their revised paper L&V present results both with, and without, this approximation (see their Fig. 3), which demonstrates this bias. We take that as an admission that the analysis based on this approximation is wrong (biased). However, they don't draw the obvious conclusion and omit the approximate analysis and results, but present both as two alternative approaches, and in the concluding section they present the difference between the approximate and exact result in a way that misleads the reader to interpret it as an uncertainty range. We think it is appropriate to point out these facts, but in the revision we have reduced this paragraph to pointing out the nature of their approximation, a brief description of the results they have presented, and a description of our findings.

Line 318: In the revision we have removed the the offending phrase, which we agree is unnecessary.

[2] Line 31: In 8 out of 16 models studied by Geoffroy et al. (2013) one can observe a very small overestimation of the transient response in the 1 pctCO_2 scenario when parameters in the two-box model are estimated from the 4xCO_2 step-function scenario. This discrepancy does not have to arise from nonlinearity, however. It is just as likely a result of the simplicity of the two-box model. It is well known that a long-memory response will lead to a slower temperature rise under transient forcing than a short-memory response (Rypdal and Rypdal, 2014, Rypdal, 2016). The physical reason is that a long-memory response is associated with energy transport from the surface into the abyss and hence slower temperature rise at the surface. Hence, if the GCMs

contain a response on even longer time scales than the long scale in the two-box model the result would be a slower temperature rise in the GCMs than in the two-box model for the 1pctCO2 forcing.

As we understand Merlis et al. (2014), volcanic forcing and abrupt CO2 change yield similar values for the fast component of the climate sensitivity in GCMs, but 5-15% smaller than the transient climate sensitivity. For the same reason as explained above, long memory in the response will give rise to a lower transient response and an underestimation of the sensitivity. Hence, these effects do not necessarily imply nonlinearity in the response.

[3] Agree, see e.g., Fig. 8 in Rypdal and Rypdal (2014). We have decided to omit the reference to Andrews et al. and include several others that are more relevant.

[4] For global GCMs we know that a two-exponential, or a power-law, response function work quite well, and we have a pretty good ide why. It has to do with the different thermal inertias of the mixed layer and the deep ocean, and the rate of heat exchange between the two. We have much less clear ideas about the response function of the ZC-model. As Reviewer #2 pointed out. The ZC model is very different from a GCM. The 25 yr time delay response to the slow solar forcing is solely based on visual inspection of the forcing and response time series is admittedly very crude, but we have no reason to believe that a more sophisticated response function is any better. Since the purpose here is just to find an estimate of the variance of the internal variability, we think the approach makes sense.

[5] Along the same lines. The reason why we cannot do this in a meaningful way is that we have so poor knowledge about the response function for the ZC model on the short time scales. The reason we chose a harmonic oscillator model in Fig. 4 is the apparent enhanced ENSO oscillations after major volcanic eruptions. If we use a certain response function and we get different fits for the sum of responses from the responses to the sum of forcing we can always blame the incorrect response function.

Hence, this will not construe another test.

[6] In principle we agree that we should put confidence intervals on these two curves to demonstrate that they are not significantly different from each other. This could easily be done as we do in Fig. 2 of the revised paper by Monte Carlo simulation of 1/f processes. However, in Fig. 1d the two curves to compare are so much on top of each other that they cross each other several times. We have explained this in the revision.

[7] This point was discussed in our response to point [1], but let us elaborate on it here. The "invalid and completely unnecessary approximation" would be apparent by reading Sect. 3.4 in the L&V paper. The approximation is the basis of their Eq. (5), which assumes that one can neglect a cross term which is the product of the solar response and the volcanic response on a given time scale Δt. L&V argue that one can do this because solar and volcanic forcing are statistically independent processes. The approximation would have been OK if we had a large ensemble of realisations of solar and volcanic forcing to average over, but in this case we have only one realization of each (the historic forcing over the last millennium). One of us (K. Rypdal) was a reviewer of the paper and pointed out this weakness in the first review. The result was that L&V kept the old results, but added a paragraph at the end of page 8 where they admit that "the cancellation of the cross terms assumed by statistical independence is only approximately valid on single realizations, especially at low frequencies where the statistics are worse."

The source of this error is probably rooted in the sloppy notation of using angular brackets <> for averages which are really not ensemble averages (or expectations) but rather estimates in the form of time averages. If two quantities X and Y are statistically independent their the expectation E[XY]=0, but the time-average estimate <XY> is normally nonzero, and on long time scales Δt we have have virtually no statistics, so there is no reason to believe that <XY> is a good estimate of E[XY]=0.

To us it seems clear that L&V have understood the error, and the appropriate response would be to omit this approximation and replace the blue curve in their Fig. 3b with the one computed without this approximation. But then this Fig. 3b would look similar to our Fig. 2, and obviously be much less convincing. Instead they present the "correct" curve as a ratio given by the lower curve in their Fig. 3c, along with the "incorrect" ratio (the upper curve). The "correct" ratio is probably more or less the same as we would get if we compute the ratio between the red and the blue curve in our Fig. 2 (our results are not completely identical to L&V, which may be due to slightly different steps between the values of Δt where the Haar fluctuation is computed – but we have used codes downloaded from Shaun Lovejoy's web site). We also find that the red curve is higher than the blue for 200< Δt <1000 yr, but on these time scales the fluctuation level is estimated from 5 effective data points for Δt=200 yr, and for only 1 effective data point for Δt=1000 yr. Actually the number of effective data points is even smaller because the time series is not white noise, but exhibits dependence on all scales. Hence, it is obvious that these differences are not statistically significant.

So there are two issues here. One is scientific; the results without the approximation are not significant. The other is the way L&V are presenting their results. In the revision we have decided not to dwell too much on L&V's presentation and focus on the results.

[8] We included the reference to MacMynowski et al., Geoffroy et al., and Fredriksen et al., which have presented spectra for a large number of CMIP5 models.

[9] One should keep in mind here that the harmonic oscillator response was employed for comparison with the ZC model, which responds very differently from GCMs. As stated in the paper, the purpose of this demonstration was not to present a realistic response model for either of the model results analysed by L&V. It was simply to demonstrate that the effects which L&V attributes to nonlinearity is easily produced in linear response models with internal noise. And as a pedagogical tool, we think a driven, damped harmonic oscillator is an excellent choice.

Interactive comment on Earth Syst. Dynam. Discuss., doi:10.5194/esd-2016-10, 2016.