

## ***Interactive comment on “Problems with solar, volcanic, and ENSO attribution using multiple linear regression methods on temperatures from 1979–2012” by T. Masters***

**J. Nielsen-Gammon (Referee)**

n-g@tamu.edu

Received and published: 24 December 2013

General Comments on “Problems with...attribution...”

This paper addresses some simple techniques to extract the effects of known forcing mechanisms from time series of global temperatures. The paper is not clearly written, does not go far enough in its testing, and is incomplete in its analysis. After rewriting and additional investigations, the paper would be likely to make a worthy contribution to the journal.

Specific comments

C642

1. Several specific comments and numerous technical comments will deal with the writing. In this comment, I'll spend some time dissecting the abstract, since it's the most important text of a paper. The first sentence is overly complex and wrong as written. It should say “The effectiveness of two multiple linear regression approaches in removing solar, volcanic, and El Niño Southern Oscillation (ENSO) influences from global mean temperatures obtained from simple energy balance models (EBMs) and global climate models (GCMs) is examined.” The next two sentences present information in the wrong order, since it's much easier to understand an error if you already know what the correct statement is. Also, the use of the phrase “in fact” is inappropriate, since everything in the paper is supposed to be factual. The paper also uses a different term for EBMs than the one already introduced in the first sentence, leaving it to the reader to figure out whether they're the same thing. Better would be: “EBMs and GCMs driven by observed solar and volcanic forcing for the period 1979-2012 produce a slight warming trend after the year 2000, while multiple linear regression attribution with an assumed lagged response incorrectly diagnoses net cooling from those forcings. The assumption of an exponentially decaying response leads to better attribution of EBM and GCM output but still incorrectly attributes too much post-2000 cooling to combined volcanic and solar forcings in the presence of an imposed deceleration in the underlying warming.” At this point it would be appropriate to insert a sentence explaining why these errors arise, and these reasons are discussed below in point \*\*\*\*\*. Then, the abstract text jumps directly to the results of a different method without a transition sentence. Such a transition could be “To the extent that responses of the actual climate system are properly represented in these models, a similar linear regression analysis of the observed global temperatures would overestimate the role of solar and volcanic forcing in the observed post-2000 global warming hiatus.” Then, leaving out unnecessary information, “An alternative approach that applies the solar and volcanic GCM response to observed temperatures leaves an unexplained deceleration in the surface temperature increase after 2000 of  $-0.06$  to  $-0.12$  K dec<sup>-2</sup>, possibly caused by internal decadal variability.” However, the abstract might be changed beyond this as

C643

the paper itself is modified in response to this and other reviews.

2. The paper identifies various things as “Alternate A”, “Alternate B”, “Method 1”, “Method 2”, and “the primary configuration”, leaving it to the reader to keep things straight. Better would be “fast response”, “slow response”, “lagged response fit”, “exponential decay fit”, and “medium response”, respectively.

3. Testing lags between 0 and 10 months is fine as long as the best lag never comes out to 10 months. Report the lags, and if a lag is ever 10 months, expand the bounds of allowed lags.

4. 1068: Equation 3 is not derivable from Equation 2 because  $T_a$  does not appear in Equation 2. The LHS of Equation 3 should be  $\Delta T(t)$ , and line 22 should say “change” rather than “anomaly”

5. 1069, 1-3: Combining solar and volcanic into a single forcing series with units  $W/m^2$  assumes that the primary impact of solar variability is through its impact on total top-of-the-atmosphere forcing. However, there is considerable evidence that the impact of solar variability on climate is much more complicated than that. See, for example, the review by Gray et al. (2010, Rev. Geophys, doi:10.1029/2009RG000282). This assumption should be explicitly discussed in the text. Consider a scaling for solar forcing relative to volcanic forcing as an additional model variable or as part of a third multiple linear regression method that tests the exponential assumption but not the forcing equivalence assumption.

6. 1070, 6-8: The Climate Prediction Center defines the NINO 3.4 Index as the average value of sea surface temperature in the Nino 3.4 region relative to a normal value calculated from a prior 30-year period. Is this what you did, or did you use a different definition?

7. 1071, 9-11: “it complicates things” is not a sufficient excuse. Discuss the real reason: a linear underlying trend is known to be a bad assumption pre-1975 because

C644

a quasi-steady growth in anthropogenic forcing as required by the multiple regression methods only occurs post-1975.

8. 1071, 14 to 1072, 6: Response and sensitivity both vary in your EBMs. Explain how you obtained these numbers. Are they constrained by data? Is one constrained by the other so that they must both change at the same time? Is there a reference for your statement that these values include both extremes of climate? Also, note that it only makes sense to talk about potential values for the “true” response time if the Earth behaves like an EBM.

9. The terminology for the underlying signal is confusing and inconsistent. For example, 1072 12-14 says that what’s underlying is a linearly increasing radiative forcing, and 1072 17 refers to an underlying linear trend. Then there’s an underlying oscillation added to this. Is there one underlying thing, or two? 1072 8-9 defines it as one, but the rest talks of it as two. Besides this problem, the word “signal” has a meaning exactly opposite its meaning in conventional linear regression. In conventional linear regression, you detect the signals (of volcanoes, ENSO, and solar) and what’s left over is the residual, including noise. Here the signal is supposed to be what’s left over. Choose different terms and define them consistently.

10. An important concept in model testing is that of a “perfect model”. With a perfect model, the model used to generate the time series has the same structural form as the model one is attempting to fit to the time series. The test of Method B on an EBM with a linear “signal” is a perfect model case. One expects the method to reproduce the EBM output very accurately. The test of Method B on an EBM with a decelerating “signal” uses an imperfect model, and should perform more poorly. Both tests of Method A involve imperfect models. These concepts should be explained at the beginning of Section 3.3 to help guide the reader’s understanding of the results that follow. You would then delete 1074, 8-13.

11. Following up on 10: any part of the time series that is not explainable within the

C645

structural framework of the statistical model (“method” as used here) will either be simply undiagnosed by the statistical model or, to the extent that this residual part of the time series partially projects onto signals that are included within the structural framework of the statistical model, part of the residuals will be aliased onto the detectable signals. This principle explains ALL the linear regression shortcomings that you detect! Discussion of your results within the context of this principle should be a key component of the manuscript. Here are a couple of examples:

11a. Method 1 is structurally constrained to have too quick a recovery time from the two volcanic eruptions in the EBM. The residual (EBM volcano response minus Method 1 volcano response) is strongly negative a few years after the two major volcanic eruptions, 1985-1988 and 1993-1998. This happens to coincide with two of the three minima in the solar forcing. As a result, Method 1 will compensate for its inability to model the recovery from volcanic forcing by overestimating the response to solar forcing. That is, the slow volcanic recovery is aliased onto Method 1’s solar response. This in turn leads to too high a residual (i.e. estimated “underlying signal”) during the third solar minimum after 2005. (For fun, try time-shifting the solar forcing by half a cycle. You should obtain a negative response according to Method 1.)

11b. Method 2 is structurally constrained to be unable to correctly model the deceleration in the EBM deceleration case. The residual (decade-scale oscillation) is positive in the mid-1990s and negative near the beginning and end of the period. This residual projects onto the true response from volcanic forcing. As a result, Method 2’s volcanic response is a combination of the true volcanic response and the residual oscillation. Since the residual oscillation tends to be positive when the true volcanic response tends to be negative, Method 2 underestimates the true volcanic response.

12. It would be helpful if all the figures had the same scale on the y-axis.

13. 1073, 25 to 1074, 2: Save this for the discussion, and note that for Method A to work even approximately, the decay time for the climate response to forcing must be

C646

as small or smaller than the decay time for volcanic aerosols to be removed from the stratosphere.

14. 1074, 18-19: Discuss in the context of comment 11.

15. The CMIP5 models are used differently in Section 4 and 5. Some of the material needed for section 5 is presented prematurely in Section 4, such as Figure 4 and most of Section 4.2.

16. As described, part of the ENSO signal from the single ensemble member will be incorporated into the calculated forced response (10% in the case of GISS and 17% in the case of CNRM. Should one ensemble member have been withheld from the calculation of forced response?

17. The calculations here assume that (a) the intensity of the model’s response to ENSO is properly measured by the NINO 3.4 index, and (b) the model’s short-term global temperature variability is mainly driven by ENSO. These aspects are not well encapsulated by the phrase “ENSO processes”. Indeed, GISS-E2-R is worse than average at simulating the atmospheric response to ENSO, according to the metrics tested by Bellenger et al. 2013. To investigate whether NINO 3.4 has any hope of capturing the connection between GCM ENSO and GCM global temperatures, you should investigate the statistical relationship between the two in the GCMs, perhaps by comparing first differences.

18. 1075, 17-19: This assumption is untrue and should not be made. Rewrite this section, using the knowledge from the test described in comment #17.

19. An important experiment should be performed: fitting Method 1 and Method 2 to the means shown in Fig. 4 to see to what extent either method is structurally able to represent the GCM’s response to combined solar and volcanic forcing.

20. 1076, 4-24: This results section is rather unclear and should be rethought and rewritten in the context of comments 11, 17, and 18.

C647

Section 5 proposes a new method. This method should be tested to the extent possible. For example, you can withhold one model ensemble, recomputed the forcing, and use that recomputed forcing to try to diagnose the underlying signal from individual runs from the withheld ensemble. It seems quite inconsistent to do extensive testing on two methods to show that they don't work well, and then advocate a third method without any testing at all.

#### Technical comments

1066, 19-20: The last half of the sentence is unneeded; the information can be conveyed much more simply in the next sentence. However, any discussion of a slowdown needs a time scale. So replace “despite the continued. . .” with “relative to the previous few decades.”

1066, 21: See previous comment. Change “. . .impact of these greenhouse gases” to “impact of prior and ongoing greenhouse gas emissions.”

1066, 21: “which” as written refers to greenhouse gases. The sentence is very long anyway, and the simplest fix is to start a new sentence at “which”, but replace “Which” with “These”.

1067, 3: Another unnecessary phrase. Delete “In particular,”

1067, 11: For clarity, add at the end “. . .in the remaining adjusted temperature series.”

1067, 12: More than one method. Change the second “this” to “two” and change “method” to “methods”

1067, 12: You should provide more of a roadmap of the paper here, so that the reader knows what's coming and will be able to organize the information properly.

1067, 20-21: The information in this sentence has been said already. Delete.

Equation 1 needs either an alpha or an explicit description of how the y-intercept is constrained.

C648

1068, 10-11: Change “. . .a specific variable to represent the anthropogenic forcing. . .” to “. . .a variable anthropogenic forcing. . .”

1068, 17: Provide a reference for this form of EBM.

Equation 3 or the text thereafter should explicitly note that  $k = c/\lambda$ .

1069, 4-9: The cart is before the horse. Put this discussion of the synthesis of forcing data into Section 2.3 after the sources of the data are identified.

1069-1070: References are needed for the TSI data and the GISS Model E calculation of effective volcanic forcing.

1070, 9-10: “results” should be “index” and “in global models.” should be “from global model output.”

1071, 6: The second sentence is not a sentence. Change to “First, this matches the period analyzed by Foster and. . .”

1072, 26: It seems strange to reference Kosaka and Xie here (and only here) and to reference other papers attributing the pause to natural Pacific variability in the Introduction (and only the Introduction).

1075, 5-11: Don't describe what would be wrong. Replace “Using the CMIP5. . .by using the” with “We use the”, replace “we can find an” with “to”, delete the next “of”, and replace “on top of” with “to”.

---

Interactive comment on Earth Syst. Dynam. Discuss., 4, 1065, 2013.

C649