What's New, What's Wrong: Response to Reviewer RC1

Roger M. Cooke and Willy Aspinall

May 24, 2020

Thank you for reading the paper and for your comments. The paper offers a simple perspective on analyzing paleo data, it is not concerned with the provenance of the data. This does not require a long exposition. We cited several authors pointing to "state dependency" in the paleo climate sensitivity parameter. Many authors suggest non-linear regression. Problems with some of these approaches motivated our study.

Since the reviewer finds nothing new, we call out the elements which are new, in our opinion.
(1) The first "new" aspect of our analysis -- i.e. new to the discussion of paleo climate sensitivity -- is to point out that partitioned linear regression is another way to explore state dependence with certain advantages, in our opinion. One well known feature of non linear regression is that the best fit can behave wildly out of sample. Indeed, a cubic fit to the data in our Figure 1 allows temperature <u>increase</u> for forcing below $-2.5Wm^{-2}$ (both with and without constrained intercept).
(2) Another "new" facet draws attention to reasons for not constraining the intercept to equal zero. It may, and in this case does, happen that the regression line through the origin is a worse predictor of the dependent variable than simply predicting the mean of the dependent variable for all values of the independent variable. We have not found this insight in the cited literature.
(3) Nor have we encountered recognition in the literature that $\mathbf{R^2}$ does not correspond to the fraction of explained variance if the intercept is constrained to zero. ("Fractional explained variance" is statistical parlance, meaning the fraction by which variance is reduced by the statistical model. In a physical sense, statistics explains nothing). I n this light, strong arguments are needed to force the intercept to zero. The arguments we found in the cited literature are weak. "*However, note that here a necessary condition for the calculation of* $S_{[X]}$ *over the whole range of* $\Delta R_{[X]}$*, but not for the analysis of any state dependency, is that any fitting function crosses the origin with* $\Delta R_{[CO2,LI]} = 0Wm^{-2}$ *and* $\Delta Tg = 0K$*, implying for the fitting parameters that a* [the intercept] *= 0. This is also in line with the general concept that without any change in the external forcing, no change in global mean temperature should appear.*" (Kohler et al 2015, p1808). Our Figure 1 shows only CO_2 forcing, but the following remarks also apply when land ice forcing is included ($\Delta F_{CO2LIVDW11}$ from Martinez-Boti et al 2015).
There is substantial noise in the data. Thus, focusing on $\Delta F_{CO2} \sim 0$, values of ΔT vary from $1K$ to $-3.5K$ (with $\Delta F_{CO2LIVDW11}$ this is $1K$ to $-2K$). If we constrain the regression line to pass through the origin, then we must explain why the deflections at $\Delta F_{CO2} = 0$ strongly tend to drive ΔT down, while those at $\Delta F_{CO2} = -2$ strongly tend to drive ΔT up. The attempt to circumvent the intercept issue leads to questionable mathematics: "*For the calculation of mean values of* $S_{[CO2,LI]}$*, we then analyse the* $S_{[CO2,LI]} - \Delta R_{[CO2,LI]}$ *space in a second step, where* $S_{[CO2,LI]} = \Delta T_g \times \Delta R^{-1}{}_{[CO2,LI]}$ *is first calculated individually for every data point and then stacked for different background conditions (described by* $\Delta R_{[CO2,LI]}$*). In doing so, we circumvent the problem which appeared in the* $\Delta Tg - \Delta R[X]$ *space that the regression function needs to meet the origin. Some of the individual values of* $S_{[CO2,LI]}$ *are still unrealistically high or low; therefore, values in* $S_{[CO2,LI]}$ *out- side the plausible range of 0–3* $K\ W^{-1}\ m^2$ *are rejected from further analysis.*" (Kohler et al 2015, p1808). Studying the dependence of random variables ΔT_g and $\Delta R_{[CO2,LI]}$ by studying the mean or distribution of their ratio $S_{[CO2,LI]}$ is problematic. Putting aside issues of stability and truncation, consider two independent uniform variables on $[-10,-1]$, called T and R. By definition there is no dependence, yet the mean of T/R is 1.41. Suppose we examine the state dependence of T/R on R. The conditional mean $\mu_{T\,|\,R=r}$ of T, given $R = r$, is -5.5, independent of r. However, the ratio $[\mu_{T\,|\,R=r}\,]\,/\,r$ increases from 0.55 to 5.5 as r goes from -10 to -1. Statisticians estimate the coefficient of linear dependence of T on R as $COV(T,R)\,/VAR(R)$, which has dimension $[T]/[R]$.

Understanding causes of deflections from a trend line is important.

(4) We have not seen fractional explained variance used as a diagnostic in the cited literature. ΔFCO_2 accounts for *64%* of the variance in ΔT over the full Pleistocene data set. During deglaciation it accounts for *75%* and during glaciation it accounts for *48%*. Does that tell us something?
(5) Moreover, before *424 KaBP*, ΔFCO_2 accounts for *42%* of the variance of ΔT and after *424 KaBP, 73%.* This is also not found in the cited literature.
(6) ΔFCO_2 has low explanatory power on partitions into low, medium and high CO_2. Different physical situations with the same reconstructed forcing can have different global surface air temperatures. Perhaps these facts can help us understand those differences. (Parenthetically, we note that Martínez-Botí et al (2015) over-samples the recent past. Removing this feature did not materially affect our results, and similar results are obtained with the dataset of Snyder (2019), which used *1000y* time steps.)

Much of the literature emphasizes Land Ice forcing, and the fact that this must be removed for predicting the effects of doubling CO_2 when the land ice is vastly reduced. We looked at this and eventually decided not to use these forcing terms as predicting the future was not our goal. We take advantage of this opportunity to share the following:

In Martínez-Botí et al (2015), three versions of Land Ice forcing are considered: $\Delta F_{CO2LIVDW11}$, $\Delta F_{CO2LIR09E12}$ and $\Delta F_{CO2LIR14}$ which include CO_2 forcing (see Martínez-Botí et al 2015 for detailed definitions). When regressing ΔT on these Land Ice forcing terms, the climate sensitivity parameter is lower than regressing on ΔF_{CO2}. At the same time, these forcings account for more of the variance in ΔT. The lower values of S are explained by the wider range of values of the land ice forcing terms. The strongest effect occurs with $\Delta F_{CO2LIVDW11}$. The linear regression coefficient of ΔT on ΔF is COV(ΔT, ΔF)/VAR(ΔF). For $\Delta F = \Delta_{FCO2}$ these values are $0.85/0.42 = 2.04$. For $\Delta F = \Delta F_{CO2LIVDW11}$ they are $2.18/1.99 = 1.096$. Quadrupling the variance of the forcing term overwhelms the doubling of the covariance term, roughly speaking.

If we remove the ΔF_{CO2} and regress ΔT on $\Delta F_{CO2LIVDW11} - \Delta F_{CO2}$, something curious happens. $\Delta F_{CO2LIVDW11} - \Delta F_{CO2}$ yields a better predictor of ΔT ($\mathbf{R}^2$=0.94) than $\Delta F_{CO2LIVDW11}$ ($\mathbf{R}^2$=0.89). This suggests that $\Delta F_{CO2LIVDW11}$ may incorporate information on ΔT to the extent that $\Delta F_{CO2LIVDW11} - \Delta F_{CO2}$ becomes a proxy for ΔT.

Finally, we believe that one of the fruits of the arduous work that has gone into preparing these high value paleo climate data sets is that others, from neighboring disciplines, can perhaps bring new ideas and tools to bear in analyzing these data. As non-specialists in paleo climate we have benefited enormously from the inclusive and supportive atmosphere within the paleo climate community, and look forward to strengthening these collaborations.

References
Köhler, P., de Boer, B., von der Heydt, A. S., Stap, L. B., and van de Wal, R. S. W. On the state dependency of equilibrium climate sensitivity during the last 5 million years, Climate of the Past 11(12):1801-1823, doi:10.5194/cp-11-1801-2015, (2015).

Martínez-Botí, M. A. et al. Plio-Pleistocene climate sensitivity evaluated using high-resolution CO_2 records. Nature 518, 49–54; doi:10.1038/nature14145 (2015).

Snyder, C.W. Revised estimates of paleoclimate sensitivity over the past 800,000 years. Climatic Change 156, 121-138; doi:10.1007/s10584-019-02536-0, (2019).