Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2020-145-RC1, 2020

Geoscientific Model Development Discussions

Interactive comment on "ATTRICI 1.0 – counterfactual climate for impact attribution" *by* Matthias Mengel et al.

Anonymous Referee #1

Received and published: 26 August 2020

This paper describes the construction of so-called counterfactual datasets, using parametric statistical models to describe high frequency variability in the observed climate, removing trend from the parameters of those distributions via what is essentially a low-pass filtering approach, and then mapping observed quantiles onto the quantiles of the distributions with trend removed. Some basic results are show illustrating the removal of trend and the retention of variability. The method is applied to two gridded datasets.

Things that I found to be missing from this paper include 1) Justification for the various choices that are made in setting up the probability model (e.g., why Gaussian for several variables, Weibull for wind speed, and why the particular choice of low-pass filtering approach based on singular spectrum analysis). 2) Motivation for the decision to fit the model using a Bayesian approach, which then requires that priors of various

kinds be chosen – and justified. 3) Discussion about how one goes about ensuring that dependence between variables is respected (which needs to be well represented for many impacts models). 4) Discussion about maintaining physical consistency between variables (also required by many impacts models, e.g., to ensure closure of energy and water budgets) and how it is maintained in the heavily processed counterfactual datasets that are produced. 5) Sufficient discussion of the homogeneity (or lack thereof) of the underlying factual dataset and its impacts. The authors attempt to minimize potential problems, for example, by pointing to a basis in a long reanalysis (line 70) and by trying to argue that early data have limited influence (lines 282-283), but I don't find these arguments very convincing. 6) Cross-validation using references from the literature concerning differences between factual and constructed counterfactual climates that are found. For example, have others written about changes in South American wind speeds, and have causes been explored and tested in models? Surely, this kind of validation is the minimum that should be expected to ensure that the datasets that are produced are fit-for-use. 7) Discussion of what applications might, or might not be suitable.

There are also some aspects that I found to be somewhat confusing. An example is the consideration of daily temperature skewness – which I found confusing given that daily mean temperature is modelled as a Gaussian random variable (and thus has zero skewness, by assumption). What is being discussed using some kind of shorthand that is known to the authors, but perhaps not others, is the nature of the diurnal temperature cycle – but how that diurnal cycle, and its variation in time, is represented by daily minimum temperature, daily maximum temperature and a measure of "skewness" of some kind is not made clear. Another example is the choice to represent relative humidity (which has a strong diurnal cycle that can be important for some impacts models, and is confined to values between 0 and 100%) as a Gaussian variable – and then to clip that distribution if quantile mapping happens to produce values outside the 0-100% interval.

Finally, it seems to me that this paper is not a terribly good fit for GMD; it might be a better suited for a data journal in my view.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2020-145, 2020.