Referee comments in italic
Responses in blue
"Quotes from the revised manuscript as plain text"

Review – ATTRICI v1.1 by Mengel et al.

This is a greatly improved version of the paper that was originally submitted – thank for. In my view, is nearly ready for publication. The main thing that I think remains to be done is to add some additional caveats that draw attention to nuances of interpretation that users will need to be aware of when applying the method and the dataset.

Thank you for these positive words! We address all comments in detail below. We include in the revision an update of the factual datasets GSWP3 (from v0.5b to v1.09) and W5E5 (from v1.0 to v2.0) that correct several errors and extend the covered period to 1901-2019. For more details see here. We apologize for the additional time this data update took. All line numbers refer to the clean version (no shown tracked changes) of the newly submitted manuscript.

The following, therefore, lists a few additional comments and suggestions for the authors to consider. The comments are referenced by line number.

11: The meaning of "counterfactual" should be stated in the abstract so that readers perusing GMD abstracts will understand what the method described in the paper is intended to produce.

Response: Thank you. We added the following explanation (line 11):

"... counterfactual baseline, which characterizes the system behaviour in the hypothetical absence of climate change, …"

This statement is from Stott et al. (2013), which fed into the IPCC WG2 AR5 definition of attribution.

73-76: I find this statement problematic on a couple of levels.

First, ironically, the methods used to produce the counterfactual climate are Bayesian, indicating that the parameters relevant to constructing counterfactual climate scenarios are described probabilistically. Shouldn't the fact that the intent of a Bayesian treatment is to quantify uncertainty at each stage of an analysis and propagate it appropriately to the next stage signal that a probabilistic treatment of impacts attribution is also needed?

More fundamentally, this suggests that the impacts do not feed back onto the climate – but often they do, both by affecting the evolution of the forcing and thus the forced response of the climate and by affecting its internal variability, and thus altering the climate forcing that the impacted system is experiencing. Perhaps this only happens locally, but it could also happen on a large scale with potentially large implications for the evolution of the forced component of climate change (e.g., carbon cycle feedback from climate impacts on forest ecosystems). The impacts themselves are also likely subject to their own sources of internal (not forced by *climate) variation that in turn might, or might not, affect the climate that is doing the forcing via feedbacks. One could think, for example, of a forest that is being impacted by the climate variations that it experiences, but that is also being impacted by insect disturbances, where insect population dynamics have their own internal variation that might not be entirely determined by climate. For these reasons, I would think that the attribution of impacts to climate would, in general, need to be treated in a probabilistic way, just as attribution of climate change to external forcing needs to be treated in that way.*

I suggest, therefore, that a bit more work be done to carefully nuance this statement.

Response:

The reviewer is correct that in reality a) impacts feed back to the climate evolution and could alter climate variability and b) impacted systems may have internal variability. The statement therefore indeed needs more nuance. Concerning a), handling climate as external forcing to impacts is one of the simplifications at the core of ISIMIP, which allows to run impact models stand-alone and not as part of larger and more expensive climate models. This approach is taken up in ATTRICI to make impact attribution feasible for as many impact modellers as possible. Handling climate as external forcing is analogous to classic climate attribution where human emissions (e.g. greenhouse gas concentrations by RCPs) are handled as external forcing and changes in climate do not feed back to them. Our statement was also fuzzy concerning b): the ATTRICI approach does not hinder impact modellers to include internal variability - for example through insect population dynamics - and to investigate the role of such variability versus the external climate forcing. We now write (lines 74-80):

"The dataset is derived from the observed realization of climate, excluding the analysis how climate variability could produce alternative realizations of factual or counterfactual climate. The attribution approach is thus deterministic and not probabilistic, focusing on the separation of climate change from direct human influences as potential drivers of changes in the impacted systems. Concerning the internal variability within impacted systems, impact models to date largely do not resolve such variability and model a deterministic response to external drivers. Our approach would however allow for probabilistic attribution to climate change once impact models resolve internal variability."

We also made small adjustments to Fig. 1 to make the difference between attribution to climate change and attribution to anthropogenic forcing clearer.

Concerning the Bayesian approach, we use this to fit our model to the factual climate data only for numerical stability reasons, see lines 303-305. In fact, the posterior distributions we obtain for our model parameters are very narrow. They are too narrow to represent the true uncertainty of our counterfactual climate data because those posterior distributions neither represent the fundamental limitations of our detrending approach nor the data problems we discuss. To avoid creating a false impression of certainty, those Baysian uncertainties are not propagated. Instead, we advise users to apply our detrending approach to multiple factual datasets in order to at least quantify the climate input data-related uncertainty (lines 136-138).

104-111: I think a few caveats are needed here. An implicit assumption is made here that an impacts model that is calibrated for the factual climate will continue to work equally well for counterfactual climates. That's not something that is necessarily a given. Consider, for example, a hydrologic model that is deployed on the Elbe River basin. Most hydrologic models of the Elbe would be carefully calibrated using observed hydrologic quantities (streamflow, water temperature, etc) and observed meteorological drivers (air temperature, precipitation, wind speed and direction, solar radiation, etc) prior to using the model for prediction, historical reconstruction or future scenario development. This tuning is specific to the observational period that is used for calibration – with the result that the tuned model might not perform robustly in a different climate (e.g., with greater winter snow storage in the drainage basin in a cooler climate, or perhaps with greatly diminished snow storage in a future warmer climate). The point is that impacts models, just as climate models, are not entirely process based (as stated on line 119).

Response: That is a valid point, thank you. We have added the following condition to the end of the paragraph (lines 114-115):

"This assumes that the climate impact model calibrated to perform well in the factual simulation performs robustly also with counterfactual climate input data."

317-318: There are placeholders for subscripts that presumably should be removed.

Response: Thanks for spotting this, we removed them.

319-320: There are many black and orange dots, so should "dot" be plural? Some explanation of the small dots and the single larger dot of each colour would also be in order.

Response: The larger dots are to illustrate one particular example of quantile mapping of a daily temperature value. To make this clearer we have slightly adjusted and extended the explanation in the text (lines 321-326):

"The counterfactual daily data are generated by quantile mapping, i.e. an observed value x that corresponds to a certain quantile of the factual distribution A(T,t) is mapped to the counterfactual value x' that corresponds to the same quantile of the counterfactual distribution A(T=0,t). We illustrate this for an observed value x that corresponds to the 95th percentile of the factual distribution in Fig. 4: We first obtain the cumulative probability of the factual (i.e. observed) temperature (large black dot in panel a) from the factual cumulative distribution function (CDF) (black line in panel b). We then obtain the counterfactual temperature (large orange dot in panel a) from the counterfactual CDF (orange line in panel b)."

We also made the figure caption clearer (Fig. 4):

"The single large black and orange point on the dashed vertical line in panel (a) highlight the factual and counterfactual value on October 25th."

333: Say what is meant by "physical bounds".

Response: Thank you for that. We are more specific now, writing (lines 338-339)

"We use the Gaussian distribution to model these variables as their values are far from their physical lower bound of zero."

349-351: I imagine that this process of randomly turning dry days into wet days will result in some physical inconsistency with other variables. Perhaps a few words drawing attention to that possibility would be in order.

Response: That is a valid point, thank you. We have extended this paragraph as follows (lines 357-358).

"This random conversion of dry days into wet days may result in physical inconsistencies with other climate variables. These inconsistencies are small by design since the new wet days are the least wet of all counterfactual wet days."

372: Suggest using the IPCC AR6 regions, if possible.

Response: This is a good suggestion, and we should indeed have used AR6 regions from the very start. As it would at this stage need a new analysis and the rewriting of the discussion of example regions, we would like to stick to the current definitions.

491: It would be good to say something about what constitutes "attribution". A calculation of the difference in impacts between factual and counterfactual climates (assuming that impacts can be determined with similar levels of confidence in both climates) would be a start, but attribution – drawing a causal connection and quantifying the change due to that cause – presumably requires careful arguments to rule out other confounding causes. At minimum, I think would need to be convinced a) that the change calculated with the impacts model is a reliable estimate of the change in the real world, b) that observed changes (to the extent that there is data) agree with the model simulated changes and c) that this similarity is not inadvertently due to confounding factors that affect the observed world but are perhaps not taken into account in the factual and counterfactual data used to drive the impacts model.

Response:

Thank you for this. The idea is in the manuscript, see for example line 107-110:

"In a first step the climate impact model forced by observed climate and socio-economic drivers has to demonstrate to be able to reproduce the observed changes in natural, human and managed systems as measured by an impact indicator (comparison of black and blue solid lines in Figure 2). The attribution of the observed changes in natural, human and managed systems is built on a high explanatory power of the factual simulations."

See also the second sentence in the Figure 2 caption:

"First, in an evaluation step it has to be demonstrated that historical impact observations (black line) can be explained by the process-understanding as represented in the applied impact model and available knowledge about historical climate and socio-economic forcings."

However, it deserves more clarity and the use of causality language. We therefore added a short paragraph to the discussion that builds on your ideas (lines 505-510):

"Attribution draws a causal connection and quantifies the change due to the cause. An important part of the attribution work is thus to ensure that the cause-effect relationship is correctly captured in the model. This requires careful analysis and model evaluation to show that the change estimated by an impact model is a reliable estimate of the real-world change. Simulated changes need to agree with observed changes and it needs to be ruled out that this agreement is due to confounding factors that drive observed change, but are not part of the model simulations. The ISIMIP3a historical simulations serve to address these points and demonstrate the explanatory power of impact models as an integral part of the attribution work."

References

Stone, D., Auffhammer, M., Carey, M., Hansen, G., Huggel, C., Cramer, W., ... & Yohe, G. (2013). The challenge to detect and attribute effects of climate change on human and natural systems. *Climatic Change*, *121*(2), 381-395.