

## Response to Reviews

### Major comments

1. The additional discussion in the conclusions section helps to put this work into context of the current state of understanding in the field. The paper's conclusions are at times perhaps overly negative about the value of the study. That the inclusion of the parameter ensemble does not have a substantial impact on uncertainty in global average temperature estimates is a useful result. The paper's results indicate that parameter uncertainty estimates may be important at regional scales. The true impact of this will not be clear until the data set is used in other studies.

**Response:** This section is slightly revised in light of these comments.

2. There appears to be no reference to Figures A1 and A2 in the text. Figures C1 and C2 appear to be referenced in the incorrect location (where it seems that A1 and A2 would be relevant).

Some additional comment on what we see in the figures in Appendix A and Appendix B would also be beneficial as they appear to show that the parameter estimates differ for differing months. This appears to have an impact on uncertainty estimates in less well-measured regions.

**Response:** These issues are addressed now.

3. There are several typos and grammatical errors to be corrected (see detailed comments).

**Response:** These are corrected now.

### Detailed comments (line numbers corresponds to the version with tracked changes):

- Page 1 Line14 – I suggest removal of the word natural from this sentence. The cited references assess observed variability with no such assessment on whether this variability is natural or non-natural (i.e. anthropogenically forced).

**Response:** Many thanks, word natural is removed.

- Page 1 line 15 – A slightly odd choices of references with limited relevance to this study. The use of model here is rather broad, covering two assessments of NWP model output and an application of an integrated assessment model. Milton and Earnshaw (2007) compare NWP temperature output to individual observation sites and not global gridded data sets (although they do use spatially gridded data for precipitations). Edwards et al (2011) assess NWP temperatures against observations from only a single observing site (no spatial estimation or fields used). Glanemann et al. (2020) does use global temperature data from NASA GISTEMP.

**Response:** Milton and Earnshaw (2007) and Edwards et al (2011) citations are removed from this line.

- Page 1, line 19 – To say this data is “mostly derived” from these sources may not be accurate. While much modern data is indeed shared under WMO and GCOS umbrellas, the historical station data sets depend greatly on data compilation by research and data rescue efforts. This includes work by groups at NOAA, the Climatic Research Unit, the International Surface Temperature Initiative, and others such as the ACRE initiative.

**Response:** This line is modified.

- Page 2, line 29 – typo: does not.

Response: Corrected.

- Page 3 – line 10 – Could mention HadCRUT5 here, as this supersedes HadCRUT4 and includes spatial interpolation through kriging/gaussian process regression.

Response: Lines 1-4 on page-7 are shifted here. (previous track changes document)

- Page 3, line 26 – It has yet been mentioned that the paper uses HadCRUT4 so this line appears abruptly. Should this sentence appear later in the paper, in the first paragraph of Section 3 or at the end of section 3?

Response: This is shifted at the end of section-3.

- Page 4, line 27 – This additional description of the functions of the autoregressive weight and lambda parameters, with reference to Nychka et al. (2019) is appreciated.

Response: Many thanks.

- Page 5, line 2 – There are a few grammatical errors/typos in this added text: “\*The\* empirical semivariogram”, “the average distance \*between\* each bin”, “\*The\* semivariance”, “The semivariogram”.
- Page 7 line 20 – typos: “Therefore, it was was converted to an ensemble data \*set\* by Morice et al. (2012) using \*the\* Brohan et al. (2006) uncertainty model.”

Response: Apologies, these are corrected now.

- Page 8, line 25-28 – this is now clear that only the ensemble is used in this study and not any additional uncertainty information provided in HadCRUT4.

Response: No action needed.

- Page 9, line 13 – should omega be a symbol here?

Response: We used the same notation as it is used in the manual of the LatticeKrig package so that the readers could relate easily. However, it is converted to maths writing now.

- Page 9, line 25- the additional detail on why 1988 was chosen is useful. The additional technical details in this section improve the reproducibility of the study. This will also be of aid to readers that are familiar with the LatticeKrig method.

Response: No action needed.

- Page 9, line 32 – sentence construction: please change “Other two” to “Two other”.

Response: Thanks, changed.

- Page 10, line 1 – It is helpful to see the modelling assumptions now clearly stated here. The additional figures and discussion in section 4.2 are welcome additions that aid interpretation of the differences between the new results and those of Ilyas et al (2017).
- Page 14, line 15 – thank you for adding this discussion of the effects of lambda for different sampled values.

Response: No action needed.

- Figure 5 – It is rather hard to see the lines in this figure. Please consider increasing the contrast between the lines and shaded regions.

Response: Contrast is increased.

- Page 17, line 1 – This should mention the reason that the data set it is rather unwieldy – the large ensemble.

Response: Mentioned.

- Page 17 lines 11 – I recommend adding the word average here to be clear that this statement refers to global average temperature estimates. The case for regional temperature is not clear as it has not been investigated in detail.

Response: The word 'average' is added.

- Page 17, line 14 – Inconsistent use of capitals: HADCRUT4/HadCRUT4 (HadCRUT4 used elsewhere).

Response: This is corrected here.

- Page 17, line 15 onwards – That the inclusion of the parameter ensemble does not have a substantial impact on uncertainty in global average temperature estimates is a useful result. It indicates that the use of point value estimates in existing assessments does not have a substantial impact on total uncertainty estimates. The comment on use in studies of climate variability at line 15 is apt. While not investigated in detail in this paper, it is quite possible that the interpolation parameter uncertainty may be more important for regional climate studies than for the global average.

Response: No action needed.

- Page 17 line 17 – This statement here may be over critical of the value of the work. Spatial model parameter uncertainty estimates may be important for regional assessments, which is not ruled out by this work.

Response: This line is removed and two lines are added.

- Appendices – These figures may be better suited to a supporting information section rather than appendices, dependent on the Journal's style guidance.

Response: Noted.

- Figures A1 and A2 - Figures A1 and A2 do not appear to be cited in the text. Should the references to Figures C1 and C2 at page 9 line 32 refer to these two figures? These two figures with Figure 2 suggest that the model parameters do change for different fields/times. The corresponding uncertainty fields in Figures 3, B1 and B2 suggest an impact of uncertainty in spatial fields, particularly in unobserved regions. This does not appear to be commented on in section 4.2.

Response: These details are now added in section 4.2. Page-9, line-32, reference of the figures is corrected as well.

- Figures C1 and C2 – These appear to be referenced in the incorrect location in the text (see previous comment). They should be referred to in Section 6, around page 16 line 4.

[Response](#): Thanks, these are referred in this section now.