

## Review of “The Berkeley Earth Land/Ocean Temperature Record” by Robert A. Rohde and Zeke Hausfather

The paper is a dataset paper describing a new merged global surface temperature product arising from a merge of the Berkeley Earth land product with the Hadley Centre’s HadSST3 product. There is a clear need to document the merged product and the paper is well within the scope of the journal. The fundamentals of the method are sound. The paper thus should be published following revisions. I append comments differentiated by major and minor below for the authors to address.

### Major comments

1. The biggest issue is not one of the authors own making, but rather indicative of renewed interests in surface temperatures and a proverbial race by dataset producers to create new and improved versions of products, presumably for inclusion in the upcoming IPCC assessment report. The implications are two fold for the present study as follows.

Firstly, the Hadley Centre have recently updated their SST product to HadSSTv4 which builds upon the new ICOADS R3.0 which has considerably better coverage in several critical periods. It also does a better job of handling modern era biases. The dataset is available from <https://www.metoffice.gov.uk/hadobs/hadsst4/> and my feeling is that it would future proof the current analysis to use the HadSSTv4 product rather than HadSST3 as the marine basis. I am assuming that updates to HadSSTv3 will stop when HadCRUTv5 becomes operational so this decision will be enforced onto the team sooner or later. Changing now would save the need for another paper / the situation where there is a mismatch between the paper and the operational product. Changing would also, presumably improve coverage in the historical eras and thus improve the analyses. Unless there is a compelling technical impediment to doing so I would urge the authors to switch over SST source to HadSSTv4 now.

Secondly, the new versions of datasets and presence of new products means updates are likely warranted to the comparisons section. These include:

- HadCRUTv5 (in final review) – contact Colin Morice for details
- <https://essd.copernicus.org/articles/11/1629/2019/essd-11-1629-2019.pdf> - the Chinese merged product recently extended back to 1850
- <https://www.nature.com/articles/s41561-020-0582-5> which provides spatially complete estimation based upon HadCRUTv4
- <https://journals.ametsoc.org/bams/article/doi/10.1175/BAMS-D-19-0095.1/348446/The-EUSTACE-project-delivering-global-daily> - new global surface air temperature estimates
- <https://journals.ametsoc.org/jcli/article/33/4/1351/346368/Uncertainty-Estimates-for-Sea-Surface-Temperature> - substantially updated uncertainty estimates on the NOAA product

Not all these need be used but clearly the HadCRUTv5 and updated NOAA estimates are key to include in revisions to at a minimum figure 7.

2. I am not entirely convinced there is merit in persisting with a version that uses SST under sea-ice as this clearly is not a surface temperature. The true surface temperature in such regions is either the ice (near-)skin temperature or the air temperature at some nominal height above the surface. There is a potential risk of mis-use of a product that considers sea surface temperature under ice as a reasonable estimate of a true surface temperature. It would possibly be better to discuss this but provide only the air temperature over ice version as this is the only realistic product over these regions. It would be good to back up with references the contention made line 277-281 in this regard.
3. The discussion in Ln 30 to Ln 34 is dated. The ICOADS release 3.0 is described in a newer manuscript (Freeman et al., 2017, doi:10.1002/joc.4775) and forms the basis for HadSSTv4 and ERSSTv5. The GHCN dataset has been updated in Menne et al., 2019 (<https://doi.org/10.1175/JCLI-D-18-0094.1>) and now includes very many more stations arising from the efforts of Rennie et al to improve land holdings. NASA don't, any longer, therefore use additional stations. This whole passage could be expanded and made a little more clear to provide a better and more accurate context for the reader here.

More generally the discussion of others efforts is somewhat perfunctory and in several aspects significantly dated. The whole introductory section requires substantive updates including several additional new products and new versions of products as noted in major comment 1. In particular, with the move to HadCRUT5 all products will employ some form of interpolation.

4. The temperature to distance correlation assumption in the ocean interpolation step is probably reasonable in ocean interior gyres. But it presumably breaks down in vicinity of upwelling, downwelling, coastal shelf seas and boundary currents. A little more justification / discussion is required than is given in Ln 77-79. I suspect that you will need to provide a specific caveat about likely location-specific performance in such regions.

### Minor comments

1. It feels dangerous to claim on Ln 11 of the abstract that the product is more homogeneous. Without an absolute benchmark it is impossible, sadly, to say whether any given product is more homogeneous than another and aspects such as spatial smoothness can be misleading as pointed out in Sherwood et al., 2009 (<https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.1825>). I don't see an explicit justification for such a statement from the underlying text.
2. There are several newer analyses than Cowtan et al. 2015 on Ln 25 regarding SST/SAT. Richardson et al., 2018 plus newer in press papers by Lea Beusch, Nathan Gillett, Gareth Jones and others. It would also be worth being explicit how these measures may be expected to differ with time e.g. that SAT would be expected to warm / cool a little more than the underlying SST were to warm or cool.
3. Ln 28 the product is a joint effort between the Hadley Centre and UEA's CRU
4. Ln 28 Note newer references exist for the NOAA product and its uncertainty estimation (see major comments)

5. Please double check with GISS colleagues whether they continue to apply an additional night-lights based adjustment as my reading of Lenssen et al ended up ambiguous in this regard.
6. Both NOAA and NASA have switched to using ERSSTv5 so the text and reference ln 38-39 needs changing accordingly.
7. Globaltemp (ln40) does include some limited interpolation over land and is complete over the oceans. This needs to be corrected accordingly.
8. ln 47-49 NOAA and NASA now use closer to 20 thousand stations following the GHCNv4 update and this should be reflected here.
9. ln 233-236 makes little sense as written. I think you mean to say that in the more recent past coverage uncertainty diminishes in importance and bias uncertainty becomes increasingly important?
10. ln 273-274 or Antarctic coastal stations, surely? There are two hemispheres with sea-ice ...
11. Can a reference be given e.g. to the ERSSTv5 paper to back up the assertion on ln296-298? Note also that NASA and NOAA use ERSSTv5 and not v4 as noted in a prior comment.
12. In figure 3 are the trends OLS fits? You need to be explicit not just about the ARMA model applied but also the trend calculation method.
13. As noted in a prior comment NOAA is interpolated to an extent so ln 307 and ln313-314 require revision accordingly.
14. HadCRUT and not HadCRU in ln 338