

Overall, this seems like a careful and constructive effort toward a very useful product. I applaud the authors for thinking to make this data publicly available and for the substantial effort to provide a reliable and well-documented description.

The manuscript as submitted suffers from some intellectual and compositional sloppiness. For one glaring example, without searching separately this reader still would have no idea what the acronym PRIMAP refers to?

Several of the Figures have serious presentation deficiencies that render them almost useless. These deficiencies subtract substantially from the overall quality of the product. Find a communication officer at PIK and use their help to change, revise and overall improve the graphics.

The overall discussion - very necessary - on data uncertainties seems to present serious logic failures. See detailed comments below.

Page 1, line 24: We need definition of the EDGAR acronym here, where it first appears?

Page 2, line 10: We have no definition of the acronym PRIMAP?

Page 2, lines 13,14: "For details and sector names we refer to Table 1 9." Confusing! This does not seem to point to Table 1 on page 14 but rather to Table 9 in Appendix B. Rather than referring the reader to an end-of-the document appendix, why not put the Table directly here where readers would find it useful?

Page 2, lines 29,30: "In the supplementary information we present a list of territories included in the emissions of UNFCCC 30 Parties as well as information on the territories that are treated separately (Section C)." Again confusing. 'Supplementary information' to this reader implies a supplement but this paper has no supplement and in any case ESSD and Copernicus do not archive supplements. Does these refer to Table 10 in Appendix C? Again, why not put it here if useful, or at least provide a hot link.

Page 3, line 15: "... preprocessing is available in Section D." This refers to Appendix D?

Page 5, line 8: The authors defined KP (Kyoto Protocol) in a footnote but not yet in the main text?

Page 5, line 11: which convention?

Page 6, line 27: Here we read that, under UNFCCC CRF, the 2015 edition used IPCC 2006 categories. But a few lines earlier (line 15) we read about IPCC 2006 categorisation as a disqualifying factor. Do we need more-exact definitions of the IPCC 2006 categories? Do we have a terminology problem? We need clarification. From subsequent text, e.g. lines 2 and 3 on page 7, we get the impression that 2006 IPCC categories represent a future and desirable standard not yet implemented in this data set? Clarification needed! Perhaps Table 9, which in the caption refers to IPCC 1996, provides some hints? If so, we need a stronger declaration here, we should NOT need to look at Table 9 and guess.

Page 7, line 15: "United Nations energy statistics" these presumably come from a different source than United Nations FCCC but from where? Do we get a reference or we need to go to the source materials?

Page 7, line 20: "... as well as other substances." Other regulated substances?

Page 7, line 22, 23: Confusing use of punctuation in this section. Presumably UneP means UNEP?

Page 8, line 1: FAOSTAT - does this source report annual data?

Page 8, line 22: "past 1990". Prior to 1990? Since 1990? Figure 1 implies that one can use Houghton et al. data back to 1850?

Page 8, line 24: implies that CMIP5 comes from IPCC but in fact it comes from World Climate Research Programme?

Page 8, line 28: give explicit credit for the MAGICC6 chemistry model (NCAR etc.)?

Page 10, line 24: Although referenced together in this sentence, Table 6 occurs several pages after Table 5.

Page 12, Figure 2: A useful but graphically fairly simple figure. Could the authors redraw it to show exactly the sequence used by the CSG for this data product?

Page 13, Figure 3: We need much more information and explanations about the steps, the sequence and the labels (or absence of labels) to find this figure useful!

Page 14: Presumably one could match the Table 1 steps to Figure 3. Table 1 caption could contain additional useful information? E.g column categories refers to IPCC 1996 categories? Likewise for Tables 2, 3, and 4?

Page 15, lines 5 to 16: A link to Table 3 in this section?

Page 16, lines 1 to 13: A link to Table 4 in this section?

Page 23, Figure 5: axis labels, category labels, panel labels all unreadable or missing. Need a substantial revision of this figure and with a substantial caption! Likewise for Figure 6, page 24!!

Page 21 and 22, Section 6: This section has no impact or utility because of the severe deficiencies in Figure 5 and Figure 6!

Page 25, line 9: Does Table 7 come from Andres 2014 or Andres 1996? Confusing.

Pages 28, 29, Figures 7 and 8: Again basically useless due to lack of panel labels, series labels, axis labels. Did no one look at these figures before submission?

Pages 22 to 27, the entire Section 7: (Why do we have sub-section 7.0.1, 7.0.2, etc?) The authors correctly describe the extreme challenges of extracting, compiling and reporting composite uncertainty estimates. Table 8 illustrates the problem to the extreme. But then the authors fail to follow their own advice and cautions and build this section largely - to the extent we can only guess from the unreadable Figures 7 and 8 - on comparison of the generated data set to the individual source data included in the composite. They know this is invalid, they say this is invalid, then they do it anyway! The entire section needs revision and rewriting to clarify how much they can not conclude because of this lack of independence of source and product and how much - if anything - they conclude without violating this fundamental requirement.

Section 8: This reader found Section 8 much more realistic and cautious, even if less quantitative, than section 7.

Appendices: Some of this information could go more usefully directly in the text. Only Appendix A, D and E seem useful in separate form.