Review ESSD-2020-34

Substantial timely effort to resolve “range = uncertainty” “tangle”. Large challenging effort for which time and expertise required remain hidden or at least understated.

Overall, a remarkable effort of immense value. Everyone involved in emissions and budget calculations, as well as larger numbers of folks engaged in numerous (often divergent) political approaches and strategies around lowering those emissions, should read at least the cautionary Discussion (Section 8). Authors and advocates of various emission products, even if their particular product does not appear here, need to learn from this work. If pressed, I would recommend Figure 3 and Figure 5, but specific countries and specific analysts will need to explore carefully through all details of Section 6. I strongly recommend publication in ESSD!

Author several times notes the absence of a definitive reference emission data product. Add a statement, at least once, that extends that statement to the inevitable conclusion that no data product discussed here can claim fully-global all-sector coverage despite their titles. (Perhaps GCP comes closest?) Author several times notes benefit of well-staffed data compiling organizations (e.g. IEA) and of positive outcomes enhanced by training and consultation (e.g. in EU). Bring these positives forward into a recommendation, perhaps near the end of Section 8? Recommendations as they stand (mostly sotto voce) seem to imply more work to tease out differences. Call out at least a few good examples that might serve to minimize errors and differences? Set yourself or colleagues up for a good EU proposal to accomplish some of the needed rectification?

This reviewer remains a bit confused about which data sets the author evaluates and why. Early in the manuscript, the author disclaims attention to gridded emissions products. For good reason - this manuscript focuses on source (“system boundary”) uncertainties rather than disaggregation / interpolation techniques. But his list of proscribed gridded products includes many products that later appear in tables and discussion: CDIAC, CEDS, EDGAR, etc. In some cases I know that both national and gridded versions exist, but in other cases I would need to check (meaning that I do not understand offhand the distinction). Other readers may also not know nor understand this initial distinction. Figure 1 shows six primary and five secondary products (total = 11, including gridded products). Of these, the author goes on to provide details of all but ODIAC (one of the gridded products). Table 1 (useful, necessary!) lists nine sources, omitting ODIAC and CAIT. Sections 5.1 through 5.10, however, provide textual detail on 10 sources, now including CAIT. In Figure 3 the reader again encounters eight data products, now excluding ODIAC, CAIT and (surprisingly) UNFCCC CRFs. I assume the author has valid reasons for inclusion or exclusion of specific products in each of various tables, figures and paragraphs, but those reasons escape me. I have few doubts about the magnitude of the effort nor about the skill required, but the organization seems to distract?

<u>Specific comments</u>

Page 2, line 14: “reasons why estimates differ between datasets, but this requires”. Please resolve singular / plural: ‘reason … requires’ or ‘reasons … require’.

Page 2, lines 16, 17: “not all datasets attempt to be comprehensive either geographically or by including all emissions sources”. True, perhaps, but they all appear under a ‘global’ label with implicit ‘comprehensiveness’? A technical methodological fault clouded by mis-leading self-promotion? Not much the author can do if a data product claims global coverage but misses recent years or ignores key sectors? See note on possible recommendations, above.

Page 3, line 5: “This” change to ‘that’, referring to Macknick 2011

Page 4, line 1: “In so doing he presents of emissions estimates from the global combustion of coal …” Something wrong or missing in this sentence?

Page 4, lines 3 to 5: “Guy Callendar, investigating the influence of fossil …” This sentence would make more sense and show consistency with the prior sentence if author moved the citation ahead, to just after the name: ‘Guy Callendar (Callendar 1938), investigating the …’ Same for line 6, Gilbert Plass. By these changes (e.g. moving the citations and therefore the dates to the start of each sentence), readers can better follow the time sequence of these early estimates.

Page 4, line 19: “They” here refers to the advisory panel, to the Revelle et al citation, or to? Please clarify.

Page 5, line 4: “alternative source of energy data” - Alternate to what? What alternate? One of these sources, original or alternative, eventually evolves into the basis for CDIAC?

Page 5, lines 8,9: “… parameters, still constant in time, an assumption …” - the phrase ‘constant in time’ invites confusion here. You mean parameters such as carbon content of fuels, combustion efficiency, etc. were used by Marland and Rotty as fixed (invariant) through time and across countries? But CDIAC eventually and now used source-specific and country-specific combustion efficiencies? On line 11 you refer to these country-specific factors? You also write (lines 11, 12) about avoidance of the “use of global-average conversion factors”. What constant parameters were used in the original paper and what time-variant replacements (improvements) occurred as a consequence of Marland and Rotty (better that you should tell us based on your knowledge rather than that we each should apply our own interpretations)?

Overall comments, Section 2: I enjoyed historical accounting and consider it of immense value. To have it now in one place, close by all the current estimates, seems intensely useful. From my training, however, I missed discussion of the Suess effect related to bomb ^{14}C? Keeling certainly published on that by 1979 (I had to go back to look) and Hans Suess (cited earlier) must have done likewise. Perhaps not so much related to quantitative emissions but - given many side comments in this section about source assumptions - identification (first) and quantification (later) of the Suess effect pretty much nailed fossil fuels as the causative source of growing CO2 concentrations? Author discusses the (Seuss?) dilution effect with reference to Baxter & Walton 1970 but other authors, including Suess, had it earlier? Perhaps those earlier publications did not follow through to actual quantitative emissions? No specific changes here, only curious about how what I thought I learned fits with the author’s re-counting of events.

Page 5, line 28: “flue gas desulphurisation” need a valid reference here.

Page 5, lines 29, 30 - clarify by slight changes in wording and punctuation: “Datasets may exclude carbonate emissions entirely or include emissions only from cement production (e.g., CDIAC) or from all carbonate decomposition (e.g., EDGAR).”

Page 6, line 4: “from land-use change and carbonates combined” - this phrase implies that you want combined emissions from LUC and carbonates. But, in fact, you mean ‘from land-use change and from combined carbonate sources’? With this minor change, sentence now fits better with 13% and 5% from most recent GCP budget?

Page 9, line 19: “Eurostat (IEA, no date-a).” Author uses this designation to link to distinct EIA and IEA URLs in the reference list. Copernicus / ESSD perhaps have a preferred format? Typesetters / proofreaders will pick this up?

Page 12, line 23: “because of high non-fuel use in oil” - non-fuel use of oil?

Page 12, line 28: “oil is fully disaggregated” - what does this mean? Oil use already fully disaggregated (e.g. no further detail) or oil fully disaggregated into source types, refined uses, non-combustion uses, etc?

Page 13, line 46 (line numbers turn continuous in Section 5?): “the other two datasets”. Which ‘ other two’ datasets? CDIAC and IEA? Acronym confusion here.

Page 13, line 48: any uncertainty data for CAIT?

Page 13, line 52: CEDS = gridded!? As expected if useful for CMIP6?

Page 14, line 94: “number of small countries (see Figure 1)” - Figure 1 shows CDIAC compiled from UNFCCC reports plus independent CO_3 data but shows nothing about inclusion or exclusion of small countries?

Page 15, line 128: “emissions as ±10% at 95%/2sd” - at least at first use, write this out as 95% Confidence Interval / 2 standard deviations?

Page 15, line 141: “more complete v5.0 provides” - v5.0 here refers to the v5.0 Crippa et al. source listed in Table 1. That source also publicly available, e.g. in reference to v5.0 FT version referred to a few lines earlier (on line 139). So the FT version has available and non-available versions while the definitive v5.0 always exists in available form? Possible confusion here?

Page 16, lines 152, 154: I realize author needs to adopt uncertainty units as provided by each of the sources but here one encounters “2σ” where earlier (for CDIAC) we had 95% CI and 2 standard deviations. Could the author apply and report uncertainties in a uniform set of units, perhaps with a prefatory note that 95% CI ~= 2 sd = 2σ?

Page 16, line 156: “EIA is a federal statistical agency formed” - write instead a ‘US federal statistical agency? Small change here will clarify reference to US data in subsequent paragraphs.

Page 17, lines 204, 205: “GCP reports uncertainties at 68%/1sd level (Friedlingstein et al., 2019).” Because you have, in prior sentence, referred to emissions uncertainties in GCP at (again) 95% CI / 2 standard deviation, here you should clarify that 68% CI, 1 standard deviation applies to the entire (net) global carbon budget (sources and sinks)?

Page 18, line 236: “41 flows.” Explain ‘flows’ in this context? Or, move the flows definition starting from line 239 to here instead?

Page 18, lines 242, 243: “usually released in October/November” - for the prior year? Because of independent specific reports and (presumably) skilled staff efforts, subject to the same year reporting uncertainty described earlier?

Page 19, line 249: “in addition to differing by statistical differences” - Not sure what you mean here? By statistically valid differences? Or by quantitatively useful statistical differences?

Page 19, lines 257, 258: I think this means that, in working with and responding to IEA, national staff develop knowledge and skills that then improve national ability to comply with other national reporting functions, e.g. to UNFCCC?

Page 19 line 267: At the start of this sentence you refer to the BUR in lower case but at the end of the same sentence you refer to them in upper case. Decide one or the other? Note also line 299, BUR capitalized.

Page 21, line 305: HYDE 3.2 (updated to 2015?) appeared in ESSD (https://doi.org/10.5194/essd-9-927-2017), with a much better description of sources and time extent.

Page 21, line 325: Figure 3 shows eight, not six, emissions data products.

Page 22, line 331: Legend refers to seven data sets but table shows eight?

General comment on Section 6: Packed with information but this reader finds text and graphics very useful!

Page 22, line 335: Data easily available and very easy to use. Thanks.

Page 23, Figure 4: 'WLD' refers to 'world'? (Given the title, replace 'world' with 'global'?) Because EDGAR data goes through 2018, this must be EDGAR v5.0 FT or v5.0 standard? Sub-categories of PRIMAP (Hist-TR and Hist-CR) were NOT described in PRIMAP section 5.9?

Page 24, Figure 5: similar questions to those for Figure 4 but convergence factors (e.g. Page 23 lines 344 to 349) very clear. If, taking GCP as reference, emissions in 2014 amounted to 35 Gt, then 1.7 amounts to only 5%. Substantial improvement? Does this deserve more mention / discussion? Note: I find 5% number repeated around line 655 but still without much emphasis? Author should, rightfully, claim or proclaim this number as a not-automatic outcome of sorting through a mass of information? Best case under current limitations?

Page 24, line 363: "in liquid fuels that IEA because" - 'than' instead of 'that'?

Page 25, line 368: "because of statistical differences" - again, what does the author mean by use of the term 'statistical differences'?

Page 25, line 385: "global energy consumption from three energy datasets" - list the three data sources (IEA, BP, EIA), otherwise readers needs to go to Figure 7 to know what the author refers to? Why include EIA here given its earlier outlier status?

Page 26, line 393: "BP's oil consumption numbers lie slightly below those of the IEA and EIA over the entire period". Not true in Figure 7 as presented.

Page 26, line 396: "Gross Calorific Value (GCV; Higher Heating Value)" - Why does reader encounter 'Higher Heating Value' here. Term only used on this page and in earlier usage (line 389) author did not capitalize. Not clear what value the term adds?

Page 39, line 614: "emissions through to over 6000 Gt CO2" - What does the author want to say here? Not clear.