Verbatim text from two reviews compiled here.

Review #3, Tom Oda. Page and line numbers refer to track changes version; reviewer assumed line number =1 at the top of each page.

Dear the authors of the manuscript,

It is my pleasure to have an opportunity to review this manuscript. I do my review for ESSD manuscripts following the guideline summarized by Carlson and Oda (2018) ESSD. The 2018 guideline should tell you where I/my comments came from. Given the nature of ESSD, most of my comments go to the method section pointing out the information that I feel missing. Below is a summary of my comments. You will also see a list of my line-by-line comments. Hope my review is useful.

Sincerely,
Tomohiro Oda (toda@usra.edu)

1 Data description
While I imagine the amount of effort dedicated to compile this budget/manuscript, I feel the level of the information is not meeting the ESSD's standard. First, please improve the data/model tables. Tables should include name, version, data source (where we can get the data, w DOI), data citation (this is different than data source), etc of the items (e.g. data and models) used in this manuscript. For the data that are not yet published, the authors should provide enough information to allow readers to understand the data. This is the basis requirement. ESSD will be hosting this manuscript/paper not for carbon cycle scientists, but also for people in the broader Earth science community. Perhaps also for someone who randomly pick up this manuscript w/o any strong Earth Science background. ESSD papers need to be useful for such readers and the described data needs to remain available (see more in the 2018 guideline).

2 Summary/Rationale/Justification of the approaches
This is partially done in some sections, but we would like to see the summary of what are covered and what are not. I see the authors did a wonderful job to collect the all of the results and synthetize them. However, probably the authors would agree with me that there should be something not included in this manuscript. I use as atmospheric inversions. It is totally fair to use a subset of inversion models available in the community. I would like to suggest the authors to just list model inclusion criteria and expand discussions including what are not covered in this manuscript. For example, some highly relevant papers such as Crowell et al (2019) (OCO-2 inversion comparison) and Schuh et al. (2019) must be worth discussing in this manuscript. I thought the authors did a great job in the DVGM section. It was concise and short (given the amount of underlying info). Yet it has the basic info that I wanted to see. For some sections, I also thought it would be great to have a short review of what types of the approaches are taken to examine a flux component and tell us if you have them all in this report or not. That would help us reader to recognize the position of this synthesis report in the big picture of the carbon cycle science.

3 Uncertainty analysis
The authors did great job in listing the potential sources of uncertainties. I would like to suggest the authors to discuss the sources in more quantitative way. I can imagine some of the uncertainty sources are tough to assess. However, at least, the authors should give a try to convince us that the uncertainties discussed should be small enough and the results remain robust.

4 Summary of advance in our understanding of the global carbon cycle
The authors kind of did this, but I thought it would be great to have a summary section (or as a part of the abstract/conclusion) of the advance in our understanding of the global carbon cycle (excluding the methodological differences) beyond the budget breakdown (I understand this is the main thing, though). For example, the authors mentioned the confidence levels for the budget components. Are the levels are changing? Since probably the authors are placed in the best position to summarize the current understanding of the carbon cycle science, I would like the authors to highlight a bit more of the state-of-the-art science results collected. I believe it would be a benefit for the general scientist readers. Also, it would make this ESSD manuscript accessible to the random people who just googled "global carbon budget" and found this manuscript.

P5, L5: Describe -> Describe and synthesize?
P5, L6: five major components, such as fossil fuel emissions (FOS), Land-use Change (LUC), G_atm, S_ocean and S_land, This is pretty obvious for us carbon cycle scientist, but maybe not for the ones who just randomly picked up this manuscript w/o any carbon cycle background.
P5, L24: COVID-19, spell out.

P6, L5: and gas flares?

P7, L20: introduce ppm

P8, L1: any other references? Maybe RECCAP reports/papers?
P8, L1: The IPCC methodology, which IPCC methodology do you refer to?
P8, L21: Citation for AR5
P8, L14: our estimates, which are mostly annual basis? (w some exceptions)

P9, L1: detailed description of the data sets, the data table does not seem to meet ESSD's standard. For example, data description, data source(s) and data citation(s) are only insufficiently provided.
P9, P7: citation for AR6
P9, P17: I believe this should be the best synthesis effort. However, it is still true that this was done by a subset of scientists in the research community. I feel this manuscript should mention the criteria for data and model collection.
P9, L20: ESSD requests to provide data sources in order to ESSD data users/ to examine original data. In principle (ideally), the data sets used in an ESSD paper needs to be available with a DOI (see Carlson and Oda, 2018). Many of the model output do not seem to be accessible for ESSD data users.

P10, L7: The estimates of EFOS -> The estimates of E_FOS in this study, there are variants of E_FOS emission calculations. I assume this sentence refers to the reference approach (if we define it using the IPCC terminology)
P10, L10: peat fuel?
P10, L21: CDIAC, spell out
P10, L13: UNFCCC (spell out) national inventor reports, did you take estimate based on the reference approach (or sectoral approach)?
P14, L14: Most accurate -> the best estimates? Inventories can't assure their accuracy by themselves. Also, UNFCCC inventories from countries could be challenging to use together as they are not compiled in a globally systematic way like EDGAR does. Please not following the IPCC guideline does not assure the accuracy.
P10, L17: BP, yes, but the BP stats does not offer data transparency and traceability that we would like to see in ESSD papers.

P10, L24: UN -> UNSD?
P10, L27: Do those values remain the same as defined in Marland and Rotty (1984)? How much differences from other reference approach estimates (UNFCCC if used) should we expect to see due the values specifically used in the CDIAC estimates?
P10, L30: WWII, years?
P11, L3: It sounds like these are ones calculated by the sectoral approach. How exactly did the authors reallocate IPCC-defined sectoral emissions to fuel-based emissions?
P11, L8: CDIAC, is this originally due to the data collection by UNSD (or submission by India)? I believe India is not the only country that reports data using their own fiscal year period. Why India?
P11, L20: which values exactly did you use for projecting 2017 emissions? There are several values you could use to do the projection.

P12, P10: how about oversea territories for other counties, such as UK? How did you spatially allocate those emissions?
P12, P18: Could you provide information more quantitatively? How large the impact would be?
P12, L21: Why can you mitigate the issue by adding international bunker emissions? Even adding the international bunker emissions to the sum of country totals, you would still have ~3% or so emission discrepancy (e.g. Oda et al. 2018 ESSD).
P12, L24: Can you provide more quantitative information (e.g. sink size, sink rate, lifetime..) in order to let us know how significant this could be in the global carbon budget?

P13, L7: How did you assess the uncertainty? Is the method independent from Andres et al. (2014)? Note the uncertainty range for the global total (6-10%, 90% confidence interval) mentioned Andres et al. (2012) is based on the assessment by Marland and Rotty (1984). So you would not bump up the uncertainty for the 2018 and 2019 while those were projected. Any comments?
P13, L18: Also see Andres et al. (2014) for regional uncertainty estimates
P13, L20: medium confidence, what does this mean?
P13, Did Ciais et al. (2013) confirm consistency between E_fos and direct observations? This seems to sound stretch.

P14, L3: Are those export and import terms consistent with ones used in the E_fos calculation?
P14, L21: Does GTAP produce fuel-based estimates rather than sectoral-based estimates?
P14, L14: Is it fair if we say the uncertainty assessment might not fully capture a systematic part of the uncertainty as well as the model structural errors?

P15, L15: three separate <global> studies?
P15, L19: Provide some mode details of projection models in order to allow readers to assess the results.
P15, P22: proxy data, such as?
P15, P24: What exactly? This is important as you are comparing different projections in the single figure.
P15, L25: absolute <daily> emission changes?

P16, L24: calibrated to the traffic data in Paris?

P17, L26: How do you assess the uncertainties associated with individual projection models? We do not see the uncertainty estimates and can't tell how we should interpret the agreement and disagreement among the estimates.

P18, P2: Are these consistent with the UNFCCC inventories? At least, some compatibility with the UNFCCC inventories? How can we assess these estimates in relation to the UNFCCC reported values?

P19, L4: Short model/data description and model output data source (DOI) seem to be missing in this manuscript.

P23, L1: Inclusion criteria for other components (data and models) should be listed, too.
P23, L25: Given the challenges mentioned, how would you assess the uncertainty associated with projected ELUC?

P25, L23: Ok, so this is relative confidence levels used in this study. Then, what would be the confidence level for E_Fos?

P26, L14: Brief text regarding the inclusion criteria for the GOBMs? We want to hear the justification for the use of the model ensemble for this carbon budet assessment.

P29, L4: Is 101% suggesting an ocean area definition difference among models?
P32, L2: Introduce xCO2
P32, L8: Inclusion criteria (like you did for the DGMV section)?
P32, L23: While focusing on large scale analysis, it is surprising that the authors did not discuss the use of satellite data using a detailed report by Crowell et al. (2018) ACP. I'd expect more discussion regarding the use of satellite data here (e.g. in-situ vs. satellites, GOSAT/OCO-2, land/ocean glint…). Probably the authors would agree with me that this subset of inversion model calculations might have failed to capture something. Also, I feel Table A4 is not providing enough details while we all agree that the results do have sensitivity to the settings.
P32, L28: transport, Schuh et al. (2019) GBC

P33, L3: Need more details since this is not publicly available yet. Grid resolution (0.1deg?)?, temporal resolution (monthly? hourly?)? Data source?
P33, L10: not GCB, but Crowell et al. (2018) used the same prior info.

P34, 2.7 Would you be able to provide more quantitative assessment for these missing components? Convince us that the GCB assessment is still robust regardless of those missing components.
P34, L25: See Nassar et al. (2010) GMD and Wang et al. (2010) ERL

P35, L7: What about gas flare emissions? Not as Ch4, but as CO2?

P37, 3: Would you be able to summarize new findings/revisions/changes from what previous publication reported (improvements of the GCB product/report/summary)?

P46, P12: Are Crowell et al. (2018) and Schuh et al. (2019) irrelevant here??

P50, P10: This is a very interesting statement. For E_eos, the use of fuel consumed is a very straightforward way to get CO2 emitted (fuel burned). It still does not assure the accuracy by itself. However, from methodological perspective, it seems to be fair to give more trust to monthly fuel stats than proxy data-based approaches given many assumptions made. In fact, emission seasonality is often constructed using fuel stats (e.g. Andres et al. 2012 Tellus). Also, as the name suggests, proxy data are "proxy" for CO2 emissions. Proxy data would beat fuel consumption while we are not entirely sure how the proxy data get collected and how the indices are developed. We do have a history of the use of monthly stats for estimating

monthly emissions, but we've not even examined the single use of proxy data for estimating emissions.
P50, L17: Here is a preprint by Zeng et al. that shows some model simulations using the Carbon Monitor emissions.

P50, L1: I thought it would be a good idea to start with the same emission value at the beginning of 2020. YTD emissions include the differences at the beginning of 2020. The projected emissions further include differences from the YTD emissions. At least, the authors could remove one of the sources of the differences to make the interpretation of the results a little bit easier (while it is not perfect by no means). I believe the authors are not very confident with the accuracy of the absolute emission values, but the relative changes.
P51, L20: But you could assign higher confidence to Chinese YTD emissions from the carbon monitor given the amount of the data collected/used?

P53, L19: And also future perspectives?

P56, L18: First, the authors should provide names, versions, data sources (where we can get them), and data citation (w DOI) of the items used in this manuscript as a list. This is the basic requirement for ESSD papers. I do see some tables, but those are missing key information mentioned above. Note data source and data citation are not the same thing. Many of the citations listed are data/model citations. Second, data policy. I am not sure this is in line with ESSD's philosophy. Why not? Please refer to Carlson and Oda (2018).

References: Please check the links if they are still active. Especially ones w/o a DOI.

Table 2 & 4 : Data sources are missing.

Table 9: In addition to Korsbakken et al. (2016), Guan et al. (2012) Nature CC, gita-ton gap. That has demonstrated potential systematic errors in emission estimates due to poor energy stats data.

Appendix A. These tables need to be improved. Can't read some of the contents at all.

Appendix B. Maybe better to change the X axis range to 0-15 (and 0-0.3 for Y axis) in order to show the points better? Points are heavily overlapping.

Figure B3. How can we interpret this evaluation?

Figure B5. As mentioned earlier, maybe it would be better to start from the same Jan emission. This manuscript does not provide enough text to explain why those lines started from different estimates. Readers need to go back to the papers published. ESSD wants papers to carry enough info to sufficiently understand the contents.

Review #4 (anonymous):

~~~~~~~~~~

Review of GCB2020 (Friedlingstein et al. 2020)

Abstract line 65-66 — not clear why 6% is the median of [6,6,7,13] … add decimal point to numbers?
~~~~~~~~~~

Lines 103-105 — glad to see this, but perhaps you could be a bit more clear that the term "global carbon budget" is explicitly different and should not be confused with "remaining carbon budget" and it variants. Could add a citation here to [Matthews et al (2020) Opportunities and challenges in using remaining carbon budgets to guide climate policy, Nature Geoscience, in press.] who also articulate this distinction.

Lines 339-341 — unclear to me how uncertainty in consumption-based emissions would not be larger than territorial, since you are adding additional uncertain components like trade flows and sectoral emissions factors.

Lines 1047-1052 — this sentence is extremely long, and somewhat grammatically problematic. Also, I don't see mention of how gross fluxes were calculated in the methods section.

Lines 1368-1373 — I have noticed a general trend the past few years of overestimating the projected growth rate in fossil fuel emissions (e.g. last year's projection of 0.6% has been downgraded to 0.1% ... the projection for 2018 in LeQuere2018 was 2.7% and subsequently revised to 2.1% in F19 … 2017 was projected at 2.0% then subsequently updated to 1.6% …). Not sure if this is significant or not, but the relative overestimate in the projected value has been pretty consistent for the past few years, so maybe speaks to some bias in the projection methodology that is worth considering.

Lines 1454-1455 — why report the median here (and in the abstract) rather than the mean? The relative independence of the different estimate is also not terribly clear — on first glance, the UEA and GCB methods seem quite similar (also given the overlap in authorship), whereas the Priestly Centre estimate is more independent in terms of methodology. Given the various results presented here (including also the IEA estimate), a mean decrease of 8% seems equally plausible as a median of 6%.