Earth System Science Data Discussions

Earth Syst. Sci. Data Discuss., author comment AC5
https://doi.org/10.5194/essd-2022-104-AC5, 2022

Reply on RC2

Giacomo Grassi et al.

Author comment on "Carbon fluxes from land 2000–2020: bringing clarity to countries' reporting" by Giacomo Grassi et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2022-104-AC5, 2022

Thank you very much for the very useful comments. Please find below a point-by point reply (in bold) to your comments.

The study by Grassi et al. presents a database of LULUCF fluxes of CO2 based on national greenhouse gas inventories (NGHGI) reported to UNFCCC, as an update from Grassi et al. 2021. These data are publicly available but for non-Annex-1 (NAI) countries, different national sources exist, not necessarily consistent with each other. The LULUCF fluxes data are gap-filled by the authors, including by using expert judgement. More NAI countries are present in this database compared to Grassi et al. 2021.

NGHGI data are compared to data from the UNFCCC GHG data interface and to the forest resource data reported by countries to FAO. Global LULUCF is a sink of CO2 that ranges from 1.6 GtCO2 y-1 in NGHGI (gap filled), 5.4 GtCO2 in "new" data from UNFCCC GHG data interface and 1.3 GtCO2 y-1 in FAOSTAT.

The effort to collect, gap fill, and compare different nationally reported estimates is a valuable contribution to the Global Stock Take process and to scientific assessments of the carbon budget in general. Previously, less attention has been paid to differences between various country submissions to UNFCCC than to differences between 'scientific' estimates and country submissions. Differences between three data sources, all based on country submission are examined, and tentatively explained. I found the explanation of differences between TBFRA, FAOSTAT and NGHGI very interesting even though it remains partly qualitative.

In addition, this study brings tentative estimates of the area of managed lands based on available information from countries.

Thanks for the nice words and the constructive and useful comments below.

Main comments:

1- The data from NAI countries are a 'mixed bag' of national communications, Biennal

Update Reports, REDD+ submissions and NDCs. For NAI, a gap filling procedure was applied, with both interpolation and extrapolation. A gap filling rate of 48% indicates a large difference between the database of this study and original "raw" national reports.

Therefore, a comparison between the raw data reported by countries (from different sources) and the gap-filled data produced by this study would need to be provided, in order to be able to appreciate the effect of the data 'processing' applied to different NGHGI data sources.

For all countries, the raw data are included in Suppl. Table 4 online (original, not gap-filled data), and can be compared to the gap-filled ones in Suppl. Table 5 online (gap-filled data). This allows anyone to make checks and further analysis.

As explained in the text, "we filled the gaps using standard statistical methods, with the aim to maintain the levels and trends of the underlying, reported raw data". Furthermore, as now more extensively explained in the Methods, "we tested the potential impact of different gap-filling methods on the level and trends of carbon fluxes. Specifically, we compared the procedure described above with two alternative approaches: (i) i.e. the average 2000-2020 using the non-gap-filled data, and (ii) a gap filling where the interpolation between two data is done taking the most recent data to fill the missing years (while extrapolation backward and forward is done as described above)."

The Results section now also include this text: "To this end, we compared our results with two equally reasonable alternatives for gap-filling on the resulting level and trends of carbon fluxes. The first alternative, i.e., a simple average of the original non-gap-filled data in each country for 2000-2020, results in a global LULUCF net sink (-1.58 Gt CO_2/yr) very close to the one obtained with our gap-filling procedure (-1.64 Gt CO_2/yr); qualitatively identical results are obtained when the analysis is done at the level of specific land categories (forest land, deforestation). The second alternative, i.e. no linear interpolation between two data points (see Methods, produced a global net sink of -1.69 Gt CO_2/yr for 2000-2020 and a trend which is very similar to the one of our NGHGI DB (Supplementary Fig. 3). This indicates that the global levels and trends that are highlighted by the NGHGI DB data are robust across a range of credible gap-filling procedures." All this means that the real effect of gap-filling is much less than what the 48% in itself could suggest.

2- A point of concern is the lack of a detailed per country uncertainty assessment of the data compiled in the database: uncertainties arising from the gap-filling vs. averaging procedures, from different thresholds of forest sink being considered as "implausible" and from differences between estimates coming from two or more national sources (such as BUR, NDC and NC).

Such a detailed evaluation of uncertainty per country is needed e.g., for assessing the significance of differences between this database and other 'scientific' estimates in further work. The authors are encouraged to further assess uncertainties at national level beyond a mean relative uncertainty value for AI and NAI, even if their value is deemed to be conservative. NAI countries are very diverse. Some of them have detailed inventories (e.g., China) and / or frequent communications and revisions, while others may have only one or two reports and a sheer lack of national data. A 50% uncertainty for all NAI country is thus not a credible estimate being 'good for all'.

Thank you. We agree that more country-level information of the uncertainly

would be useful, but we also feel that deep analyses on that would be a bit beyond the scope of our paper.

With regard to the uncertainty arising from the gap-filling procedures, we feel that the tests we did, for which results are provided in the main text (now with added explanations) and in the supplementary figure 3, should provide enough confidence on the fact that we did not alter the levels and trends in the original not-gap-filled data.

With regard to the uncertainty arising from different thresholds of forest sink being considered as "implausible", we improved the description of the five countries for which we consider the sink "implausible" and quantified its total effect (i.e. 0.9 GtCO2/y)

With regard to the uncertainty arising from differences between estimates coming from two or more national sources, all the original estimates from the various sources (NC/ BUR, REDD+, NDC) are included in the online dataset, even when a specific source was not used. This should allow anyone to "play" with the sources. Related to this, we added new text (at the end of section 3.4) on the sources being used by the Washington Post and the just-released UNFCCC Synthesis for the global stocktake reports a global LULUCF (https://unfccc.int/documents/461466). The latter source is relevant and was not available at the time of the original submission of our manuscript.

With regard to the assessment of the uncertainties at national level beyond a mean relative uncertainty value for AI and NAI, we highlight the importance of not focusing too much on country-specific values of uncertainty, due to a number of reasons: (i) while in principle all NGHGIs should follow the IPCC Guidelines in estimating the uncertainty values, in practice different and not-fully comparable approached may be used by different countries; (ii) in the context of country GHG reporting to UNFCCC, the uncertainty analysis should be seen, first and foremost, as a means to help prioritise national efforts to reduce the uncertainty of inventories in the future and guide decisions on methodological choice (IPCC 2006), not necessarily to be compared to other countries or independent estimates.

Country-level values of uncertainty for LULUCF have been already shown by McGlynn (2022), and despite additional work we did, we could not find a significant amount of extra information. Finally, we note that the good match for the mean uncertainty of AI (35%) and NAI (50%) countries used in our study (based on Grassi et al. 2017) and the (fully independent) study by McGlynn (see below) should raise confidence on the fact that the numbers we used reflect in a reasonably good manner what countries reported in the NGHGIs.

We added this new text: "Based on the values of uncertainty collected in Grassi et al. 2017, complemented by expert judgement, in this study the uncertainty on the net LULUCF CO_2 flux was estimated to be 35% for AI countries (where the dominating component flux is FL) and 50% for NAI countries (where the dominant flux component is deforestation). These values are similar to those collected by McGlynn et al. (2022) for the LULUCF estimates reported by six AI countries (average 33% uncertainty) and twelve NAI countries (average 47% uncertainty). It should be noted that the estimated % uncertainty has a broad range across countries (e.g., from 14% of Japan to 102% of Cambodia, McGlynn et al. 2022), and may be affected by the closeness to zero (i.e. when emissions and removals nearly balance out, the aggregated % uncertainty is likely to be higher). Given the incomplete information on the uncertainty of NGHGIs

(especially for NAI countries), the values used in this study should be considered as rough approximations. We then averaged this information at AI and NAI level and aggregated it at global level using equation 3.2 from IPCC (2006), Vol. 1, Chapter 3."

Furthermore, we added a new figure (Suppl. figure 5) with estimates of uncertainty for global LULUCF net emissions from 2000 to 2020.

With regard to assessing the significance of differences between our database and other 'scientific' estimates, we think that in many cases the uncertainty values from NGHGIs are incomplete and not fully comparable with those reported e.g. by global models. For example, while NGHGI express uncertainty as 95% CI, in global models (e.g. Friedlingstein et al. 2022) uncertainty is expressed as +/-1 standard deviation. In the latter case, "the uncertainties combine statistical analysis of the underlying data, assessments of uncertainties in the generation of the datasets, and expert judgement of the likelihood of results lying outside this range", an approach not necessarily always comparable to the one in the IPCC Guidelines (see below discussion on this point). In addition, the % uncertainty in LULUCF is strongly affected the by the NET balance between emissions and removals: if close to zero, the % uncertainty will appear bigger. This balance, in turn, is affected by the system boundaries set in the estimation of the CO2 flux. To this regard, global models and NGHGI use different system boundaries to define the "anthropogenic sink", and this further complicates the comparison of their estimated uncertainty values. Identifying ways to make the uncertainty estimation from NGHGIs and global models more comparable remains a very important aim, which however goes beyond the scope of this study. We take the good reviewer's comment as a stimulus to further research needs.

3- The notion of 'uncertainty' is not defined in this study. If the difference between two NGHGI datasets is for instance larger than 200% for a given country, then an "uncertainty" of 50% is clearly underestimating the uncertainty of the whole approach.

From IPCC GLs the following text has been added to clarify what uncertainty refers:

"As per IPCC Guidelines, Uncertainty is here defined as the lack of knowledge of the true value of a variable that can be described as a PDF characterising the range and likelihood of possible values. It does refer to random errors, although the central value of the pdf may be affected by unknown/unquantified biases. As per IPCC good practice systematic errors (biases), once identified/quantified, shall be removed while uncertainties are to be reduced so far as practicable"

4- A comparison of TBFRA vs NGHGI for the NAI countries that have a stock change based approach should be provided to test the hypothesis provided by the authors.

We added examples of countries using the stock-difference approach in the NGHGI : USA, many EU countries, India and China. Furthermore, for these and many other countries we added new figures (Supplementary figure 6, 7 and 8) where the trends in CO2 fluxes in FL are compared between NGHGIs and FAOSTAT. These figures are discussed in the main text (section 3.5).

Other comments:

L265. How do AI and NAI estimates of uncertainties compare with the detailed analysis of Glynn et al. 2022 of LULUCF CO2 fluxes uncertainties (https://doi.org/10.1007/s10584-021-03254-2).

The values from McGlynn compare very well with the values we used, based on Grassi et al. 2017. We added now this text:

"Based on the values of uncertainty collected in Grassi et al. 2017, complemented by expert judgement, in this study the uncertainty on the net LULUCF CO_2 flux was estimated to be 35% for AI countries (where the dominating component flux is FL) and 50% for NAI countries (where the dominant flux component is deforestation). These values are similar to those collected by McGlynn et al. (2022) for the LULUCF estimates reported by six AI countries (average 33% uncertainty) and twelve NAI countries (average 47% uncertainty). It should be noted that the estimated % uncertainty has a broad range across countries (e.g., from 14% of Japan to 102% of Cambodia, McGlynn et al. 2022), and may be affected by the closeness to zero (i.e. when emissions and removals nearly balance out, the aggregated % uncertainty is likely to be higher). "

L305. The compilation of UNFCCC country reported data by FAO is intriguing. Could more information / comparison be given here about this dataset ?

The UNFCCC dataset disseminated in FAOSTAT is an independent effort but consistent with the work presented via the UNFCCC GHG data interface. As discussed in the relevant FAOSTAT metadata, available online at the dataset webpage, it too contains data collected directly via the UNFCCC data portal, individual CRFs, NDC and REDD+ reports publicly available. A part of our efforts are jointly made with PIK within the construction of the PRIMAP database. The main difference is that the FAO UNFCCC data are not gap filled. Furthermore, the FAO UNFCCC data extends to AFOLU and also makes available AFOLU activity data in addition to emissions.

L374. Should be clarified in the abstract that what is called "NGHGI data" also combine (although for a small fraction) FAO data

We can add it, but we think it is not strictly needed. On CO2 (the main focus of the paper), FAO is not used at all; on the area of managed forest, FRA data are used for 2% of forest area: we believe it is not worth adding a new sentence in the abstract just for this.

L390. Please precise the rationale for comparing NGHGI "managed land" only with forests area from FAO. There are likely to be managed "other wooded land" and cultivated lands. Why should both variables be comparable ?

The comparison we do (e.g. in Figure 1) is on forests only. Other wooded land is not considered as "forest" by FAO.

L420. What is done for averaging / gap filling when the same year has two different estimates, e.g., from a national communication and a BUR

As explained in the Methods (and now improved), for each country, only one type of submission is used in the NGHGI DB (i.e. for NAI countries, national communication or BUR, or REDD+ or NDC), prioritising the most recent one (assuming that either the data or methods, or both, have been updated) but also taking the completeness of information into account. In selecting the source of data for NAI countries, expert judgement is applied in a few cases, e.g. whenever a NC/BUR is clearly more complete than a slightly more recent NDC, the former is used (see country-specific explanation in the Supplementary Table 1, online).

L455. Please explain why, without gap filling, the global CO2 sink is less than with gap filling for NAI.

The difference is very small, and is due to how gap-filling was applied (see section 2.1): linear interpolation between two points and/or through extrapolation backward (till 2000) and forward (till 2020) using the single closest available data. Depending on the data extrapolated, one may obtain slightly different results compared to the non-extrapolated average. Anyone can produce the same results from the Suppl. tables 4 and 5 (online).

L450. If there is a 50% error on each NAI report and each report is independent, what is the uncertainty propagated on the global CO2 sink of these countries. Please explain what are the countries that dominate the global uncertainty

Thank you, we added a new figure (Suppl. figure 5). Since LULUCF net values are close to zero in some year (with emissions being close to removals), expressing global LULUCF uncertainty in % may be misleading, because very small or large numbers may occur. For this reason, the new figure shows the aggregated uncertinaty values in absolute terms (as 95% confidence interval, following IPCC guidelines). In the main text we added "Based on the values of uncertainty used in our study (i.e., 35% for AI countries and 50% for NAI countries, see Methods), we estimated an aggregated uncertainty at global level (Supplementary figure 4) of about ±0.8 GtCO2/yr (average 2000-2020)."

Please note that, because uncertainty refers to random errors, its propagation -e.g. from national to global- tends to average out errors so that the global estimate is less uncertain than the national data.

L465. Central African Republic was filtered due to an implausible sink rate. Here using NGHGI DB, a sink of 1.7 GtCO2 is given for this country, which also seems 'implausible'. Please clarify what is done with CAF – is it kept of removed? Does this single country (+ few others) explain most of the difference between the 'old' and 'new' GHGI database ? or is this difference really due to numerous data updates from many other countries

In table 6 we have now clarified the following:

"This Central African Republic reports very diverse and contradicting estimates. The NC2 2015 reports a sink of -1.7 GtCO_2/yr, which we consider biophysically impossible (see Methods). The most recent NDC 2021 reports a sink of -0.7 Gt CO_2/yr), which we also consider biophysically impossible given the relatively small forest area (20 Mha of secondary forest), which as noted in Line 159 would be equivalent to a mean area-specific sink of -35 tCO2/yr/ha. In our NGHGI DB, we used the value from the NDC 2016 (see Fig. 1 in that document, including both emissions and removals: -0.2 GtCO_2/yr) despite considered "implausible" according to our criteria (see Methods)".

Thus, the sink of -1.7 GtCO2 in NC 2015 is used in the UNFCCC GHG data interface, which may lead to misleading global analyses, but not in our database,

In addition, Figure 7 disaggregates the reasons of the difference between our NGHGI DB and the UNFCCC GHGDI. As illustrated in detail in Table 6, relatively few countries (17) explain most of the large difference in global LULUCF values between our NGHGI DB and the UNFCCC GHGDI.

Note that what we consider "implausible" or "impossible" is explained in the methods (with improved and new text now):

"For the purpose of our analysis, we introduced two indicative thresholds to assess the plausibility of the net forest sink reported in the NGHGIs, selected on the basis of various considerations, including the distribution of the forest sink per unit of area among countries (see Supplementary figure 1), the typical range of IPCC default factors and expert judgement. In particular, we considered the net forest sink as "biophysically impossible" - and therefore not included in our NGHGI DB - when the average for the period 2000-2020 is greater than -30 tCO_2/ha*yr at country level (if occurring over >1Mha). The only case that could be potentially included in our NGHGI DB and that fell in this category was the Central Africa Republic. In this case, the forest sink reported in the most recent country submission (i.e., -0.7 Gt CO_2/yr, from the NDC 2021, corresponding to an area-specific sink of about 35 tCO_2/ha*yr) was excluded from the NGHGI DB, and the value from the NDC 2016 (forest sink of -0.3 Gt CO_2/yr, or 15 tCO_2/ha*yr) was used instead.

Furthermore, we considered the net forest sink as "implausible" when the average for the period 2000-2020 is greater than -10 tCO_2/ha*yr at country level (if occurring over >1Mha). Five countries were included in this category (with a forest sink between -14 tCO_2/ha*yr and -18 tCO_2/ha*yr), collectively covering about 70 Mha of forest: Central African Republic (using the NDC 2016), Mali, Namibia, Malaysia and Philippines. For these countries, data are included in the NGHGI DB but are considered separately in the discussion (i.e. numbers are considered unlikely, but not impossible)."

How can the data change from a 1.7 GtCO2 sink to 0.2 GtCO2 source between a 2015 NC and the 2016 NDC.

This is a relevant question for the GHG inventory compilers of the Central African Republic, and underlines that there is further work required, especially for NAI, to improve estimation and reporting. To clarify: the value is our NGHGI is net LULUCF sink of -0.2 Gt CO2/yr (from NDC 2016), not a source.

L 505. I am a bit confused, please explain what is the gap filling applied by the Washington Post, if any – and which countries were removed

We added this new/improved text:

"The Washington Post (2021) estimated a global net LULUCF sink of -3.6 Gt CO_2/yr in 2019 (excluding data from Central Africa Republic), while for the same year our NGHGI DB estimates -1.9 Gt CO_2/yr (Supplementary Table 5, online). Most of the difference is due to the different sources used, i.e. the Post used only NCs and BURs, while our study included also REDD+ and NDC submissions if more recent than NCs and BURs. By using the same criteria as the Washington Post, we would obtain a global net LULUCF sink of -3.3 Gt CO_2/yr in 2019. The rest of the difference between our NGHG DB and the Post is linked to the more

updated data we used and the slightly different gap-filling procedures."

The gap filling applied by the Post is explained here https://www.washingtonpost.com/climate-environment/2021/11/07/invisible-methodology-measuring-emissions-gap/,

And here https://www.washingtonpost.com/climate-environment/interactive/2021/greenhouse-gas-emissions-pledges-data/methodology/?itid=lk_inline_manual_52

Since using the same sources used by the Post we obtain a nearly identical value (-3.3 vs -3.6) we argue that the gap-filling procedure has a small effect

L539. It is the first time that the "NHGHI' forest sink is called "implausible' ? Should clarify in section 3.4 which version of NGHGI data is more or less plausible. A global sink of 5.4 GtCO2 y-1 in forests does not look at first glance 'implausible' when compared to the land sink of the global CO2 budget.

The term "implausible" appears also in the Methods (see answer above). It is very true that a global sink of 5.4 GtCO2 y-1 (in LULUCF, not only forests) does not look at first glance 'implausible' when compared to the land sink of the global CO2 budget. Here the scope, however, is not to compare our numbers with the global C budget but rather to compile the most complete/recent information from country GHG inventories. This information is needed in the context of the Global Stock-take.

To this regard, we added this text:

"Overall, while the global LULUCF values from other datasets (-5.4 Gt CO_2/yr from the UNFCCC GHGDI for the period 2000-2020, -3.6 Gt CO_2/yr from the Washington Post for the year 2019, -3.1 Gt CO_2/yr from the UNFCCC Synthesis for the global stocktake for the year 2015) are not implausible when compared to the estimates from global carbon budget (e.g., around -6.0 CO_2/yr of net sink from all terrestrial ecosystems, Friedlingstein et al. 2022), we believe that the NGHGI DB presented here is the most complete, updated and disaggregated collection of LULUCF information based on NGHGIs."

L575 and p 22. This text makes the discussion of differences between FAO and NGHGI easier to understand. Could it be placed before

Thanks, we moved this box before as suggested

L585. Please name NAI countries that measured C stock changes in FL and how much of the NAI sink do they contribute.

We added that India and China make use of the stock-difference method in their NGHGIs, and in the context of REDD+ Mexico, and partly Nicaragua and Dominica Republic. Several countries use the NFI cycles mostly to improve the C density per strata estimates and increase its accuracy / precision.

L645. Comparison to Donegan et al. which is not published should be explained in details or removed. Unclear how a deforestation / degradation map can be estimated to be 'consistent' with FAOSTAT

Donegan is now published and the full ref. is included https://www.iufro.org/fileadmin/material/publications/iufro-series/ws40/ws40.pdf

With regard to the deforestation / degradation map can be estimated to be 'consistent' with FAO-FRA, we clarified that it is the regional forest loss statistic that is compared against.

Online tables are available here: https://zenodo.org/record/7034483#.Yw3otOxBxm8