AUTHORS' RESPONSE

1. REVIEWERS' COMMENTS and OUR REPLIES

REVIEWER 1 (Clemens Schwingshackl)

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

This study provides an important step towards better and more reliable assessments of CO2 fluxes from LULUCF, as it presents a comprehensive dataset of LULUCF fluxes from country reports to UNFCCC, and it compares this dataset to LULUCF fluxes from other databases (such as the FAOSTAT emission database) with special emphasis on explaining the differences between the investigated datasets. The methods described in the manuscript are clear, the results and explanation of differences are convincing, and I think that the study will be an important contribution for future assessments of CO2 fluxes from LULUCF.

Thank you for the nice words and the useful and constructive comments

While reading the manuscript, I noted a few points and questions that might be of interest to the authors when revising the manuscript:

- To my knowledge, complete details about the applied methodology is not available for all countries, which further complicates the assessment of the reported fluxes (lines 96-98). It might be worth to note that as well.

  Thanks. We added "This highly heterogeneous and fragmented reporting, together with sometimes unclear methodologies, complicates the assessment of the LULUCF fluxes reported by several NAI countries."

- The third reason indicated for the differences between the investigated datasets (lines 99-101) sounds rather general. Would it be possible to explain a bit more in detail what this point encompasses?
- We modified as follows: "Third, carbon fluxes are associated with complex and highly dynamic biological systems, characterized by a marked spatial and temporal variability. Estimating these fluxes in a complete, accurate and consistent manner is very difficult, and different approaches may capture differently the various natural and anthropogenic drivers."
- Line 107: Does "reporting" refer here to GHG fluxes? Or to other variables as well?

  Yes, we essentially refer to GHG fluxes, but ancillary data (managed area) are also expected to be reported. This is now more clear in the text.

- I am wondering why gap-filling was also necessary for some AI countries (5%, line 167), given that they seem to report a complete time series (as indicated in line 163). What is the reason for the necessary gap-filling for AI countries?

In the originally submitted version, we used data from 2021 GHG inventories for AI countries, including data until 2019. Thus, 2020 was gap-filled. In the updated dataset used now, from the 2022 GHG inventories (including data until 2020), this residual gap filling is not needed anymore. The text now explains that "For AI countries, all information is sourced from the GHGI 2022".

- When reading the details about NC/BUR, REDD+ and NDC reports (lines 178-196) I was wondering how these reports were used to compile the data presented in this study. Although the information is given before (as I saw when checking the manuscript again) and Table 3 gives a nice and useful overview of the origin of the data, I think it might be worth to start the paragraphs in lines 178-196 by informing/reminding the reader again about why and how the NC/BUR, REDD+ and NDC reports are used here.

  Thank you. As suggested, we moved previous lines 178-196 few rows above, close to the explanation on why and how the NC/BUR, REDD+ and NDC reports are used in our study.

- Lines 203ff: Do the countries explicitly report the location of managed land or just indicate the total area of managed land?

  We added "most countries simply indicate the total area of managed land per each land use category, and only few countries (e.g. Canada, USA, Brazil) explicitly show maps of managed lands"

- Lines 230-231: I do not fully understand how this technical detail is related to the sentences before. I think it would be good to elaborate here a bit more about this mapping (but only general, as all details are listed in Table 2).

  We clarified that the mapping of land categories is between IPCC 1996 guidelines and IPCC 2006 guidelines.

- Lines 282-287: I like that the authors mention and highlight the importance of choosing the correct method for uncertainty estimations. However, from the text I did not really understand what the difference between both methods is. I think it would be helpful to explain this better, and maybe also give an example for each case, such that it becomes clear what the difference is (and why it matters).

  The text has been revised as follows:

  "The second aspect is the ambiguity in the use of the standard error of the mean (SE) vs. the standard deviation of the population (SDp) to calculate the uncertainty of the carbon flux estimates. The SE is to be used to quantify the uncertainty of a variable that applies to the entire population from which the mean value of the variable has been unbiasedly inferred, e.g. the increment of the entire forest land when the increment value is derived from an unbiased forest inventory, so that the variability of the population does not determine uncertainty in the knowledge of the true value (only random errors in measurements do matter). In contrast, SDp is to be used to quantify thew uncertainty of a variable when the mean value of the variable (e.g. the average per hectare forest biomass carbon stock) is applied to only a portion of the population from which has been inferred (e.g. the deforested area), which means that the variability of the population contributes to the uncertainty in the knowledge of the true value; thus SDp always applies to every IPCC default value used in NGHGIs. Although such guidance is provided by IPCC (IPCC, 2019, volume 1, chapter 3), countries do not always properly use the standard error vs the standard deviation, which leads to underestimating uncertainties when the standard error is used instead of the standard deviation or overestimating uncertainties when the standard deviation is used instead of the standard error."

- Lines 342-345: Is the increasing number of available NAI countries due to the usage of NC/BUR, REDD+ and NDCs reports or is there another reason? I think it would be good to mention that here.
- We added that "These improvements reflect recent submissions (either NC/BUR, REDD+ or NDC) from countries that did not provide LULUCF data before."
- Lines 510-522: I think it would be nice to support the statements about the agreement between the different datasets by some statistical measures (e.g. correlation or trend analysis)

  Given the different time steps taken by the two datasets (every 5 or 10 years for FAOSTAT, up to one year in the NGHGIs), it is difficult to analyse the most relevant differences by statistics means. However, we added new information on the difference in trend that we consider more relevant, i.e. the period 2010-2020 for Forest land, where NGHGI DB indicate a constant sink whereas FAOSTAT indicates a decrease of the sink of 1.5 GtCO2/y. This is reflected in the new Supplementary figures 6-8 and in the following text:

  "Trends in carbon stock density (tC/ha) may help explain in part the differences of the trends in FL between NGHGIDB and FAO. In particular, after 2015 the NGHGI DB indicates globally a constant sink, whereas FAOSTAT suggests a decrease of the sink (Fig. 4b). These differences originate at country level - including for the major AI and NAI countries (see Supplementary figure 6) - and can be linked to the fact that FRA 2020 carbon stock density data, upon which the FAOSTAT estimates are based, are often constant after 2016 or 2017 (see Supplementary figure 7, e.g. Australia, Canada, Finland, USA, India, Indonesia and Mexico); this fact reflects a lack of data for the most recent years rather than a real decrease in sink capacity during the period 2015-2020. Indeed, the global forest area with constant carbon stock density for the period 2015-2020 is doubled that reported for the period 2000-2015 (see Supplementary figure 8). At the same time, the FL fluxes in FAOSTAT are estimated based on differences in carbon stock densities in 2020 and 2015 only. As shown in Supplementary figure 6, the use of 2020 FRA values with constant carbon density after 2015 (Supplementary figure 7) may lead to underestimating the sink strength in FAOSTAT in some case (for instance USA, EU27+UK, India) and overestimating it in other cases (for instance, Canada).

- Lines 641-642: Why might the datasets fail to capture the recent increase in deforestation?

  We added "It should be noted, however, that both datasets are not always very updated: FAOSTAT reflects FRA data collected up to 2017 (or earlier), while for NGHGIs it depends on the country: for Brazil data used here are up to 2016 - thus the increases in deforestation detected in the last years in Brazil (Silva Junior et al. 2020) are not included - for DRC are up to 2018, for Indonesia up to 2019."

- Lines 665ff: I think here it would be helpful to remind the reader again about the methods used to assess whether NGHGI DB or FAOSTAT is more complete (and to refer again to Supplementary Figure 2).

- We elaborated the text as suggested, adding clarifications: "Results show (Table 8, see also Supplementary Figure 2) that for 73 NAI countries, the NGHGI DB appears more complete/less uncertain than FAOSTAT on carbon fluxes. This occurs especially when the NGHGI reports non-biomass pools (not included in FAOSTAT), and when FAOSTAT estimates zero carbon fluxes on FL-FL (because of a single value of carbon stock reported over time in many country FRA reports). In the latter case, we assume that one or more estimated values of carbon flux in a NGHGI represents more information than a single value of carbon stock reported over time in a FRA report. It should be noted that here we speak of "completeness/uncertainty" because, according to the IPCC 2006 guidelines, the lack of completeness is a source of uncertainty."

- Lines 680-687: I did not fully understand how the percentages indicated here were obtained. Is this directly connected to the assessment of completeness for NGHGI DB and FAOSTAT? Or is it estimated based on a different approach? It might be worth to mention this more explicitly.

  Thank you for the comments. We added information and clarified the caption of figure 7, which now reads as:

  "Disaggregation of the differences in net global LULUCF CO_2 fluxes (average 2000-2020) between the NGHGI DB presented in this study and FAOSTAT (upper right column), and between the NGHGI DB and the UNFCCC GHG Data Interface (UNFCCC GHGDI) (bottom right column). For the comparison between NGHGI DB and FAOSTAT, for AI countries we assume the NGHGI DB to be more complete in terms of land categories and carbon pools than FAOSTAT. For NAI countries, we distinguish cases when: (i) the NGHGI DB is more complete on non-forest land uses, or more complete/less uncertain on Forest Land (FL) and Deforestation (DEF) (reflecting the analysis in Table 8) than FAOSTAT; (ii) it is unclear which source is more complete; (iii) FAOSTAT is either more plausible or more complete than the NGHGI DB. For the comparison between the NGHGI DB and the UNFCCCC GHGDI, we consider the NGHGI DB more complete because it uses sources either not considered by the UNFCCC GHGDI (recent REDD+ and NDC submissions), or simply more updated (NC/BURs with the IPCC 2006 format). Whiskers indicate the estimated global uncertainty (95% confidence interval) on the net LULUCF flux for the NGHGI DB (see Methods) and FAOSTAT (Tubiello et al. 2021). See text for details."

Furthemore, we changed the colours of the columns in the figure and changed the text accordingly, which should make it easier to follow the reasoning:

"In the first case, 59% of the total difference (1.6 Gt CO2/yr, striped parts in the upper-right column in Figure 7) may be explained by a more complete/less uncertain reporting by the underlying NGHGIs included in our dataset, in both AI and NAI countries. For another 26% of the gap (0.7 Gt CO2/yr, dotted gray part in the upper-right column in Figure 7), it is difficult to identify a clear reason, and often both datasets appear rather poor. The rest 15% (0.4 Gt CO2/yr, yellow parts in the upper-right column in Figure 7) can be explained by more plausible sinks or more complete/less uncertain reporting in FAOSTAT than in the NGHGI DB for NAI countries, including on organic soils. Based on the information above, the hypothetical combination of the best databases for each country (i.e. NGHGI DB or FAOSTAT) would yield a global net LULUCF sink in the range between -0.5 Gt CO2/yr and -1.3 Gt CO2/yr (i.e. the dotted gray area in the upper right column, Figure 7). "

I am looking forward to seeing this study accepted and published in ESSD.

REVIEWER 2 (Richard Houghton)

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

This analysis by Grassi et al. compiles a new, up-to-date, gap-filled data set of annual emissions of CO2 from land-use change for individual countries. The data set is based on country reports to the UNFCCC, including annual GHG inventories from Annex 1 countries and several types of communications from non-Annex 1 countries. The data set provides a valuable resource; it will be important for the imminent global stock take.

The study compares the new data set with two other commonly cited analyses of LULUCF emissions (UNFCCC and FAO) and provides reasonable explanations for differences. It also explains why the data that were compiled by the Washington Post were somewhat different.

The paper is well written and clear. The data set will be used by many scientists and policy makers. The authors are well qualified scientists involved with the IPCC, FAO, and UNFCCC .

A consequence of the comparisons is that the paper identifies future needs, such as: Are current rates of deforestation in the tropics increasing (Feng et al., 2022) or decreasing (FAO, HYDE)? How important are emissions from non-forests (e.g., peats)? Are forests getting denser (biomass increasing)? Or the reverse? Altogether, this is a valuable and unique contribution to the carbon and climate change communities.

Thank you very much for the nice words and the useful comments.

We added in the conclusion the following text, inspired by the reviewer's comment: "Furthermore, our analysis reinforces the urgency for the global models used in the Global Carbon Budget (Friedlingstein et al. 2022) to address research questions such as: Is managed land a net sink or a net source globally? Have rates of deforestation in the tropics been increasing or decreasing in the last two decades ? How important are emissions from non-forest lands?"

REVIEWER 3 (William F. Lamb)

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

This is an excellent dataset and description that will have a high impact in the field. Would the authors consider extending it back to 1990? As they know, this is the first year of UNFCCC reporting, and is relevant for tracking both historical policy outcomes and future pledges.

Thank you for the suggestion. While for Annex I countries the entire time series 1990-2020 is easily accessible from UNFCCC web site (https://di.unfccc.int/flex_annex1), data for Non Annex I countries are indeed rather scattered.

Despite we keep the focus of the study on the 2000-2020 period, in the online dataset and in the Suppl. Fig 5, we now present data from 1990, aggregated for AI and NAI countries. Due to lack of LULUCF information from many NAI countries, data for the pre-2000 period should be considered more uncertain than for the post-2000 period.

REVIEWER 4 (Maria Sanz)

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

In the context of large expected contributions of the FOLU sector to achieving the net zero emissions by 2050. Several studies pointed out the discrepancies between top down studies and bottom up aggregation of data from countries reporting to the UNFCCC, as well as differences between top down modelling. Efforts, such as this paper, to point out and understand these discrepancies are critical at this point in time when countries are starting the process to collectively assess progress to achieve the Paris Goal through the first Global Stock Take to be concluded by the end of 2023. The paper goes beyond and proposes a compiled tame series for 192 countries built upon multiple data from five sources (GHG national inventories, Nat Comms, BUTs, REDD+ submissions, NDCs) that makes available as a basis for further analysis.

The construction and choices made are well documented and transparent, and the original data is identifiable in the tables of the database. This allows for the reconstruction and work on alternative choices or criteria for the construction, highlighting the transparency of the work done.

The comparation with the estimations made by FAOSTAT highlight the importance of incorporation emerging more accurate data from improved reporting by countries that not always is provided to all data compliers or statistics at global level. Which leads to sometimes simplifications that can lead to misunderstandings or misuse of data (for example in modelling exercises).

Overall constitutes a first key step to provide a common basis for further global analyses in the coming months. And that can contribute to the Global Stock Take in relation to a very difficult and potentially controversial sector.

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

Some comments that may be useful for the authors in successive improvements of the data base:

- On the completeness tables in the supplementary information, it will be desirable to identify from which source if there are several each of the columns comes from.

  The sources are the same as the one indicated as “selected” in Suppl table 1 (online). We now added this column also in Supp Table 2 (online), to increase clarity.

- On the estimated uncertainty (page 11 line 270 -285) a good explanation is provided, but could the authors elaborate more on other possible factors such as AD sources (statistics/NFI versus Remote Sensing products for example).

The 2 problems identified refers to errors in the application of IPCC good practice. Regarding the use of 2 alternative approaches to collect data i.e. sampling versus census, IPCC considers equally consistent with good practice both approaches. Sampling is usually the most applied method given that is less resource intensive; while census (including wall-to-wall) providing more data may result more accurate, although this is not necessarily true since complexity may more likely. determine sources of bias. Anyhow we did not make an analysis of methods used and of the associated quality of area data, which is often determined by how the methodological approach is applied.

- Understanding that the information on other land uses than forest, it will be good in the future to reflect in a more disaggregated manner by land uses and use conversions this estimates. Please confirm that when referring to Forest lands it refers to forest lands remaining forest lands, and clarify how the lands converted to forest are dealt with in the paper, as well as the HWP data.

As explained in Table 2, in our study (consistently with the IPCC / UNFCCC reporting tables) we use the term "forest land" to refer to the sum of "Forest land remaining forest land" and "Land converted to forest land". In the same table, HWP is also allocated to "forest land". To increase clarity, this information is added also in the caption of Fig. 4.

- In the supplementary table 2, can the authors explain differences between the LULUCF column and the others.

LULUCF is sum of other categories. Now this information is added in Supp table 2 (online), as suggested

- An expanded table per countries with land use and pools where data available and how many years, and if the methods in different years are compliant with which IPCC GL will be useful in future updates of the data base.

For the whole LULUCF, originally not gap-filled data ('how many years') are already included in Suppl Table 4 (online) . We now added in Suppl Tab. 1 (online) new information of the IPCC guidelines used (based on the information we found). On pools, we added a new table in the main text (tab. 5) with key statistics on carbon pools (number of countries reporting, average CO_2 fluxes) for the main land use categories and sub-categories in the NGHGIs of AI countries. As suggested, in future updates we will try to add more information on carbon pools of NAI countries.

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

REVIEWER 5 (Philippe Ciais)

Thank you very much for the 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 further 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. 2021. 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). 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. “

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 comment is valid, however 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 uncertainty values in absolute terms (as 95% confidence interval, following IPCC guidelines).

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

2. Most relevant changes made in the manuscript page

Beyond the changes already highlighted above, the note the following changes:

- We updated the numbers using UNFCCC country data up to July 2022, including all new 2022 GHG inventories from Annex I countries and updating the data from several non-Annex I countries. Overall, global numbers did not significantly changed.
- We removed the supplementary table 1-10 because they are already available online (https://zenodo.org/record/7034483#.Yw4XtexBxm8). This makes the Supplementary section easier to read.
- We added the following new table in the main text:
    - Table 5. Statistics on carbon pools (number of countries reporting, average CO2 fluxes) for the main land use categories and sub-categories in the NGHGIs of AI countries.
- We added the following new Supplementary figures:
    - Suppl. figure 4. Estimates of uncertainty for global LULUCF net emissions from 2000 to 2020
    - Suppl. figure 5. Trends of emissions from LULUCF from 1990 to 2020 in the NGHGI DB
    - Suppl. figure 6. Trends of emissions from deforestation in the NGHGI DB and in FAOSTAT
    - Suppl. figure 7. Trends of the values of C stock density (tC/ ha) from country reports to FRA 2020
    - Suppl. figure 8. Trends of the forest area for which countries report to FRA a non-zero change in C stock.
- We added new data on harvested wood products (HWP) and natural disturbances in the main text.
- We added a comparison with global LULUCF estimates collected in the recent UNFCCC Synthesis report for the technical assessment component of the first global stocktake (https://unfccc.int/documents/461466)