

ANSWER to Referee #3

(answers by the authors are in red)

Anonymous Referee #3, 20 Oct 2022

Review comments: Mapping land-use fluxes for 2001-2020 from global models to national inventories

I have been following this work for some years, and I am glad to see an article which describes the methods in more detail. Though, more detail leads to more questions! Overall, I think this is a well written article that makes the necessary data descriptions (with some potential modifications), and I do not see any major barriers to publication. The detailed method description does make me reflect on various historical assumptions (not made by the authors, but further locked in by the authors). In a sense, the authors mechanically go through a method to bridge the gap between different definitions, ok for ESSD I guess, but I think there could be a little more reflection on the appropriateness of past decisions and therefore potential pathways forward to avoid the ‘gap’ (the authors do some of this already in the discussion).

Thank you very much for the positive and constructive comments, and for the thoughtful and stimulating reflections and questions. In general, we agree with most of these reflections. Of course, as the reviewer also noted, there are no easy answers to all the questions. Below and in the revised manuscript, we try to reflect upon these inputs.

In a sense, the paper provides an ad-hoc “fix” to a problem that sort of should not exist. Of course, there may be good reasons to have different definitions, and therefore a method to bridge those definitions. Though, at some stage, one may start to question the definitions and whether they are appropriate for the times. Science evolves, as does policy. In both cases, BM and UNFCCC, are wedded to decisions made in the past. The authors lock this in with comments like “unlikely countries will change”. We tend to hold on to those decisions as hard as we can, even if the justification is rather weak!

Key questions that come to mind: Is it appropriate to use BMs in the carbon budget? How can the Loss of Additional Sink Capacity not be included in the carbon budget?

Why can countries continue to have such inconsistent definitions of managed land? Is managed land still a relevant proxy for anthropogenic? If BMs included indirect effects (such as in DGVMs) and UNFCCC had data based definitions of ‘managed’ land, the gap may diminish significantly? Ok, I don’t expect the authors to solve these issues, but their method and dataset puts them in a unique position to comment on these issues. The UNFCCC is based on IPCC guidelines, and IPCC guidelines are also informed by IPCC ARs, so if we have better science, or can do things differently and better, that should be communicated, not just assume BMs or UNFCCC will never evolve.

Thanks, we broadly agree (but of course our paper cannot address all these points)

On models: given the combination of BMs and DGVMs can closely approximate the NGHGI definition, implementing the indirect effects in BMs would provide little added value (but of course it can be explored in future studies).

On countries, we added: “Overall, in the short term (i.e., before 2030), it is difficult that countries will change their approach to reporting anthropogenic land-use fluxes from managed lands only, due to methodological reasons (most NGHGI are based on direct observations, which cannot fully separate human-induced and natural effects) and policy concerns (compliance risks)”. We also reinforced the text before this, with the aim to provide more robust arguments to what countries may or may not likely do in the future.

“We acknowledge the open debate on how to reconcile global models and NGHGI – see the instructive discussion between M. Meinshausen and S. Federici on adjusting models vs. adjusting country data (<https://doi.org/10.5194/essd-2022-245-CC1>). A pragmatic interim solution that we propose is to adjust the global models’ results if the analysis is partly or predominantly focused on country or regional levels, and considering adjusting the sum of country data to the models’ results if the analysis is focused on climate mitigation efforts at the global level relative to modeled emissions pathways. This approach, followed also by the UNEP Emissions Gap Report 2022 (see Box 2.1 therein) ensures that country estimates are consistent with those reported by countries themselves to the UNFCCC, and that global estimates are consistent with the carbon cycle, scenarios and climate science literature used in the IPCC Assessment Reports. Given the focus of this study on regional and country-level estimates and on disaggregating fluxes into different categories, we here adjusted BMs’ results to country data”.

A major comment on the method (description). I do not think the method has to be changed, but I think we need a full description of TRENDY. Is there a TRENDY paper that describes the different runs, consequences, etc? We need a list of S0, S1, S2, S3, etc, and a description of what they mean. A figure may be nice. It seems to be so many issues arise because of the way the carbon budget has defined things. What is the land sink in terms of assumptions (S2)? How do the BMs match to the Sx conventions? What are the DGVMs not used for net LUC? How does the LASC fit in? And what does all this mean for the ‘gap’? Ok, I can imagine a response from the authors would be that this should appear elsewhere, it should be in the global carbon budget paper, or in a TRENDY specific paper, etc. But, unless one has a 10-year experience with TRENDY and its protocol, understanding some choices made in the paper, and the consequences thereof, is difficult.

Thanks. We have re-written the methods (“Global models”) to address all points raised by the reviewer. However, we refer to two original papers (Sitch et al., 2015, and Obermeier et al., 2021) for further details - since we do not use any other simulation than S2 we refrain from explaining the other simulations here. We have instead clarified that we only use the S2 simulation.

Another major comment, which I did not notice to the end, is that the data is just a table? I was expecting something more comprehensive than this. Basically, the paper is providing a summary of key results, not the data used in the paper and for the analysis?

Thanks. We considerably expanded the dataset in the online repository (<https://zenodo.org/record/7541525#.Y8WF8ezMJEI>), including for each country a time series 2000-2020 for: each BM, with results for each land-use category; the ensemble mean of the DGVMs, with the sink in non-intact / managed forest (i.e. the adjustment applied to BMs); the NGHGI data for each land use category. Furthermore, in the same repository we made available at two different resolutions the forest map that we used and the detailed protocol to process DGVM data.

I have some more specific comments, building on my points above, in order they appear in the article:

1. General, lines 67+. What is the dataset? It is a dataset which reconciles the different between two datasets? Or is it an additional disaggregation of the TRENDY dataset? Perhaps something like this is needed: “Here we provide additional disaggregation to existing models runs from DGVMs to allow a reconciliation between...”.

Thank you. We rephrased that sentence but added the term “disaggregation” elsewhere in the text.

I have not looked at the actual dataset, but I am sort of curious of what it actually is. Is it a better version of the TRENDY dataset?

Yes it is based on Trendy 11, just the disaggregation is different (intact forest, non-intact forest, other). Now it is online and includes also all results from BMs and NGHGs

It should be noted that we updated the results to make it consistent with the Global Carbon Budget 2022 (at the time of the original submission, only GCB 2021 data were available). While results changed a bit (i.e., now we have a better match between BMs and NGHGs at the level of LULUCF, but a less good match for forest land and deforestation), the message of the paper does not change.

2. General, dataset: Ok, I have now looked at the dataset, and it is Table S1? Is that it? I was expecting country level estimates at the same level of detail as in the figures. I can't really do much with this dataset, it is only a table with summary statistics?

Thanks. We considerably expanded the online dataset, see above

3. Line 181+: Sland is S2? This assumes the land areas in PI? If the land areas were allowed to evolve over time, as they did in reality, this leads to the LASC but also includes LUC? But the LUC estimate, based on BMs, does not include any indirect effects. The total land sink is S2+BM+LASC? Alternatively, this is S3? LUC according to DGVMs would be S3-S2? To my general comment above, it really needs a figure to explain this, and put a magnitude on some of these effects. According to Friedlingstein et al 2022, “The resulting loss of additional sink capacity amounts to 0.9GtC/yr”, like this is ~10% of the total emissions? This is not trivial! DGVMs have the ability to consider evolving land areas and indirect effects on LUC, but those results are ignored? In the context of comparing to NGHGs, this seems puzzling! DGVMs are uniquely positioned to bridge the gaps, and this is what the authors do, and so what are the implications if BMs are not the starting point and instead DGVMs are?

In response to this and other comments by this and the other reviewers on the use of the TRENDY DGVMs and the loss of additional sink capacity (LASC), we have consolidated our discussion of the LASC and moved this part to the end of the methods section. We have also condensed the explanation, since the key element of our translation between global models and NGHGI - the managed forest fluxes - is not substantially influenced by the loss of additional sink capacity. We thank the reviewer for pointing out to us that we have not clarified sufficiently that this issue is only of marginal importance to our approach and results.

Line 181+: Building on the previous comment. What is the justification for starting with the BMs? One could start from the LUC estimates from DGVMs? That would be more consistent? It would be useful to explain this more, and potentially show how

the BMs differ to the DGVMs for LUC? I presume through the Sx runs, it is possible to make a self-consistent definition of the total net land sink, and then disaggregate into LUC and land sink components that are most useful. If the DGVMs are used for LUC, does the gap between UNFCCC and DGVM LUC differ as much as if BMs were used?

We have re-written the methods section on global models to clarify why BMs and not DGVMs are used to quantify the land-use emissions in our study and in the Global Carbon Budgets, and why estimates from the two types of models are not combined. We acknowledge that the methods for all budget terms in the Global Carbon Budgets evolve over the years to reflect the state of the art and we are happy to revisit the reviewer's suggestion once BMs and DGVMs can be usefully combined. Since the purpose of our paper is to provide a pragmatic approach to translate between the two most common assessments of land-use emissions (NGHGI and the estimates from BMs), we believe that a discussion about alternative approaches via DGVMs would not add to the clarity of our study's goal.

4. Line 205: This text sounds like the country-based data are the national submissions? Perhaps "We used the most up-to-date and complete compilation of country-level LULUCF estimates (Grassi et al), ... This dataset builds on ... [and mention submissions]". Since "submissions" is used before Grassi et al, it reads as though Grassi et al is the UNFCCC country emissions. Some rewording will avoid confusion.

Thank you, we did follow the suggested text.

5. Line 275: The "but exclude land-use change". Technically, the process is included in the DGVMs, but the area that underwent LUC is excluded (basically, help at PI). Something like "but the land area is held at PI values as a proxy to exclude LUC"? If correct?

The sentence has been deleted

6. Line 275+: The steps "results were first...and then" could be better explained. The "first" step, with Hansen et al, is to ensure consistent definitions of forests? The "then" or "second" step, is to mask managed land by assuming intact forests? The main point is to explain why you do the "first step", as you don't say why (I don't think). There is also an issue of how the PI area maps to the current intact area? Are these all subsets (PI > Han > Intact)?

The paragraph has been slightly rewritten, and the full protocol has been made available in the online repository.

Importantly, we clarified that the Hansen mask "ensures that the current forest area, and not the pre-industrial one used by S2 runs, is applied in this study"

7. Line 280+. Somewhere, perhaps in a separate paragraph, can you define what "intact" forests are, and why they are a good proxy for "managed land". Not obvious, even if the statistics look rather good in the figures.

Thanks. We added "Intact forest is defined as areas without detected signs of human activity via remote sensing (Potapov et al., 2017) (Figure 2a), which a previous study (Grassi et al. 2021) found being a relatively good proxy for "unmanaged" forest in country reports."

8. Line 316: Just clarifying. The managed land mask is frozen at 2013 values, it has no trend? Any thoughts on the potential implications? The 2% is based on the total area,

since LULUCF is a change, a 2% change to forest area is rather significant? Basically, it is the changes that are relevant? And 2% is a big change?

Yes, there is no trend in the intact forest area we used. This simplified the analysis. If this approach had impact on the estimate of land-use changes from BM, then 2% in forest area would be relevant. But since it affects only SLAND, the impact is small: it could be roughly assumed that our 2001 SLAND sink is underestimated by 1% and the one in 2020 is overestimated by 1% – well below the uncertainty from DGVMs. We have clarified it in the text.

9. Line 321: The LASC sits on the LUC side? Or you saying that the intact mask has an area less than the PI forest area, so any LASC are masked out? I don't really understand this paragraph, but it is obviously worth mentioning. It probably needs an explanation for people that are not experts on TRENDY protocols and LASC!

In response to this and previous comments (see above) we have consolidated and re-written in a clearer way the explanation of the implications of the LASC.

10. Line 340: What is deforestation here? Is this the process of forest to non-forest, or is this a positive (gross) flux defined in BMs?

Yes, it is forest to non-forest transition, and it is also the positive (gross) flux defined in BMs. We tries to clarify in all occasions that we refer to “fluxes”

11. Line 343+: It sticks out that H&N is quite different to the others. Can you provide some more explanation on this?

We now used the GCB 2022 values, and now H&N is much closer to the other BMs. The main reasons is that previously H&N allocated shifting agriculture differently than the other BMs. The reasons why we changed from GCB 2021 data to GCB 2022 is that the latter data were not available at the time of our submission (July 2022). Even if the new results would not change the overall message of our study, we think that this change makes sense. Since our manuscript will eventually likely be published in 2023, ignoring the latest data from the GCB 2022 would sound a bit strange.

12. Line 383+: Again, this LASC seems to be something the reader has to have a good handle on to interpret some of these results. You say it is taken care of, which I can trust, but in the methods I think the reader has to have a much better of understanding of LASC, etc, to be able to pass the comments here.

See comment 9.

13. Line 400+: Why is the method worse in Canada, China, India, etc. You touch on this later, but perhaps just add some pointers that you dig into these differences later. As a reader, I am interested in in why things differ, not just stating they differ.

In this part we describe the results for LULUCF, later on we provide more information for each category, trying to add more explanation than before. We have rewritten the text to reflect the changes from GCB 2021 to GCB 2022. Some of the reasons of the differences are difficult and may deserve future studies.

14. Line 534: After reading the size of the LASC, can you be more specific on the “is likely an overestimate”. The LASC seems to be not insignificant! This is a constant theme throughout the paper. LASC seems like a rather significant issue, but it is randomly brought up (it is as if a diligent author noticed “oh, don't forget LASC, we should mention that here”). Does the article need a more systematic discussion of LASC?

See comment 9.

15. Line 550+: A “short-term and pragmatic fix”. The discussion is really how to marginally modify the status quo, what can BMs tweak and what can UNFCCC tweak? I sort of see more fundamental questions rising. Are BMs, which ignore indirect effects, too outdated now? They were good when Houghton first did it, but hey, we can do better now? Line 558 mentions a fairly fundamental issue of land management and demographic models, but then Line 559 says “a greater disaggregation” of BMs. What is the point of disaggregating BMs if demographic models are needed? One could even think in terms of observational constraints. UNFCCC inventories built on forest inventories is close to an observable, but BMs don’t include indirect effects? DGVMs may be more similar to a UNFCCC inventory in terms of processes included (indirect effects), give or take the area issue and annual variability, so could DGVMs be better overall for the LUC (vs BMs)? Or are DGVMs too uncertain? Basically, could a “short-term and pragmatic fix” be to drop the BMs? (sorry to those running BMs). BMs seems like a simple climate model that ignores the carbon cycle? BMs are not really inputs into climate models, that is more land transitions? To rephrase your comment, “what is the long-term and non-pragmatic fix”?

See our previous answers that we now clarified in the manuscript why currently BMs are still advantageous to be used and DGVMs only provide ancillary information on the land-use emissions term. We have also rephrased the text to clarify that it is not a “fix” to the BMs that we meant with the “pragmatic fix”, but that we were referring to the translation between NGHGI and BMs.

16. Line 574+: A good point to bring out is how much of the gap is due to differences in areas and how much is due to indirect effects? Quantifying that would be a very valuable exercise, and gets closer to original notion of separating anthropogenic effects? Historically, the UNFCCC approach of managed land is a proxy for anthropogenic. Do we have better science now that we can use a different proxy? Does the data in the paper give a method to better define anthropogenic?

Thanks. We tried to reinforced the text in this section. Overall, we do not see easy alternative to the managed land proxy in the short-medium term. And nor did the authors of the 2019 Revised IPCC Guidelines.

On the area of managed land, we added new text and numbers in the method section:

“At the global level, NGHGIs indicate about 3.7 and 0.7 billion ha of managed and unmanaged forest, respectively. In comparison, the IPCC Special Report on Climate Change and Land (IPCC, 2019b) indicates 2.9 and 0.9 billion ha of “forest managed for timber and other uses” and “forest with minimal human use”, respectively. In terms of global ice-free land surface (ca. 13 billion ha), about 75-80% of land is considered to be under some form of human management (Erb et al. 2017, IPCC 2019), with the rest being unmanaged forested and unforested ecosystems (ca. 2 billion ha) or other land (barren, rock). By contrast, the BMs consider a much smaller area of managed forest than NGHGIs (e.g., 1.4 and 1.3 billion ha by BLUE and H&N, respectively). Finally, the areas used in this study - based on the combination of non-intact and intact forest plus country-specific information (for Russia, Canada and Brazil) - are about 3.3 and 0.8 billion ha for managed and unmanaged forest, respectively (Figure 2b, Supplementary Table 1). Australia is the country with the greatest difference between the area of managed forest used in this study (0.04 billion

ha) and the NGHGI (0.13 billion ha, although the NGHGI assumes a large part of this area to be in carbon equilibrium)”

17. Line 618: “it is unlikely that countries will change”. Bleh. What time frame, short to medium term? Countries routinely make changes, revisions to LULUCF can be rather significant too. I disagree with this starting point. If we (the scientists) put forward good reasons to change, in IPCC reports, maybe they will change. That doesn’t mean they change to BMs or DGVMs, but they change. The managed land issue was a poor proxy for anthropogenic, with many issues. If the science has evolved to do better, just say that, and don’t prejudge whether countries are rigid or not. Same goes for the science community too... The budget once had H&N, then included BLUE, and now OSCAR also. It used to have a residual sink, now it has DGVMs. We have to think in terms of evolution, not holding onto old ideas.

Thanks. We changed the text as “Overall, in the short term (i.e., before 2030), it is unlikely that countries will change their approach to reporting anthropogenic land-use fluxes from managed lands only, due to methodological reasons (most NGHGIs are based on direct observations, which cannot fully separate human-induced and natural effects) and policy concerns (compliance risks). In addition, any changes would first need to be included in new IPCC guidelines and approved by the UNFCCC, a process that usually takes many years”.

We understand the arguments that, if better solutions exist, science should propose them and insist for their application. But right now we do not see viable and pragmatic alternatives for NGHGIs, other than the incremental improvements we suggested in our discussion. In that regard, models are more flexible than NGHGIs, models are not constrained by IPCC Guidelines, and we therefore suggest that the adjustments be made to models not to the NGHGIs.

18. Line 651: I don’t like this sentence. It implies that the science community has done it wrong, and when corrected, land is a net sink overall. One could also write: “When countries include the natural land sink together with source from LUC, they can report a sink”. What are you trying to say, perhaps: “When results from DGVMs and BMs are redefined to include an expanded definition of forests and indirect effects as a part of the anthropogenic sink, our analysis confirms that sink estimated in NGHGIs”. It is difficult to get the language right, but I think you want to avoid implicitly assuming NGHGIs are superior.

Thanks. We shortened the conclusion and deleted that sentence. But trying to take into account the suggestion elsewhere

19. Conclusions: An important point to emphasize somewhere here is the distinction between climate modelling and inventories. The NGHGIs cannot be fed into a climate model (or integrated assessment model), without those models making potentially significant modifications. Sure, let countries get excited about the sink in their NGHGI, but this has to be balanced with a smaller carbon budget, potentially significantly for some countries, and the ability to maintain a sink under climate change, may not be trivial. A major point here is the NGHGI approach is really a poisoned chalice (in my opinion), countries and some scientists have yet to realise that!

Thanks. We extended the discussion on the risks for countries associated with large managed area. On the smaller C budget, we see your point and agree, but also feel that it is a bit outside the scope of this paper (it has more broadly been discussed in Grassi et al. 2021). Nice analogy the ‘poisoned chalice’ !

20. Figure 1: Can you make a similar figure, but for S0, S1, S2, S3, LASC, BMs, NGHGs? Not sure if possible, but consider it a challenge!

Since we have re-written the text to focus on S2 (see previous comments) we do not see the need to provide more detail on other TRENDY simulations or the LASC.

21. Figure 3: I guess the average as a median would just be the middle model? In any case, the “gap” to reconcile is really a heavy function of the BM. If OSCAR was correct the gap would be much smaller, H&N much larger. What are these BMs doing differently?

We now use the GCB 2022 data (see above).

On the mean/median, we use the mean and not the median exactly because of the issue mentioned by the reviewer.

The main difference between BMs is probably the input data (we have now added this text in the methods), as discussed in Friedlingstein et al. 2022 (and now added in our text):

- BLUE is based on LUH2
- H&N is based on FAO data
- OSCAR uses both LUH2 and FAO data to constrain the model runs

And when they are compared to NGHGs, it really depends on what you look at. The gap for OSCAR is smallest for forest, but largest for deforestation. This might again hint to a different separation of fluxes from shifting cultivation in OSCAR.

22. Figure 5: I guess a is SLAND, disaggregated? For b, at what stage do you filter out models. Is there reason to believe the model at -2 and constant? Is there a trend in SLAND, and is that trend the same / different as the trend in non-intact? How does the LASC fit into this figure?

Yes, it is SLAND, we now added it. We take all the 16 DGVMs used in GCB 2022, without filtering.

We do not perform a trend analysis since our translation between global models and NGHGI does not require statistical measures on the trend.

On the LASC, please see the previous comments.

23. Figure 6: Why is H&N doing so well?

We now use the GCB 2022 data (see above), and H&N changed

24. Figure 10: Potentially useful figure. It would be even better if the LASC could be included? That would be a nice visual of the size of this effect and its potential importance. In the same vein, perhaps including LUC from the DGVMs?

Thank you for the suggestion. Since, as explained in previous responses, we now clarified that the LASC does not play a major role for the key element of the NGHGI-

BM mapping (the managed forest fluxes) we do not think adding the LASC to Fig. 10 will be helpful to the reader.