

Comment on essd-2022-245

Anonymous Referee #3

Referee comment on "Harmonising the land-use flux estimates of global models and national inventories for 2000–2020" by Giacomo Grassi et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2022-245-RC3>, 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).

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.

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.

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?

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

- 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...". 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?
- 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?
- 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 NGHGIIs, 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?
- 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?
- 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.

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

- 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) but 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 pre-judge 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.
- 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 NGHGI". It is difficult to get the language right, but I think you want to avoid implicitly assuming NGHGI are superior.
- Conclusions: An important point to emphasize somewhere here is the distinction between climate modelling and inventories. The NGHGI 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 realize that!
- Figure 1: Can you make a similar figure, but for S0, S1, S2, S3, LASC, BMs, NGHGI? Not sure if possible, but consider it a challenge!
- 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?
- 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?
- Figure 6: Why is H&N doing so well?
- 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?