

## ***Interactive comment on “Impacts of future agricultural change on ecosystem service indicators” by Sam S. Rabin et al.***

**Sandy Harrison**

s.p.harrison@reading.ac.uk

Received and published: 23 December 2019

Review of "Impacts of future agricultural change on ecosystem service indicators" by Sam Rabin et al. The paper presents and explores simulations with the coupled land use and vegetation model LandSyMM to quantify future land use change and resulting impacts on ecosystem service indicators. There is a lot of interesting and thought-provoking material here, and I am sure this paper will create a lot of interest. However, like many "future scenario" papers, there is a lack of consideration of plausibility or uncertainty. The authors do not help the reader to understand why these projections are better or more reliable than other estimates. The section on runoff and flood risk is not convincing, in part because the separation of responses takes no account of what is already known about impacts of CO<sub>2</sub> changes on runoff and in part because

C1

no case is made for using mean P95month as a measure of flood risk in a model with no water redistribution instead of relying on a projections using explicit hydrological modelling. The writing style is overblown (particularly in the Introduction) and often obscure (for example in the methods section and in the results section). The messages here could be conveyed in a clearer fashion with some pruning and rewriting, and this would considerably improve the readability of this paper. Shortening the existing text would leave room for a proper discussion section that would allow key issues to be explored. I hope the specific suggestions below can help the authors improve this paper and clarify their arguments, because the reliable estimation of future changes in ecosystem services is important for many purposes and people. Sandy P. Harrison

Specific Comments The Methods section is long, difficult to follow and at the same time does not give sufficient information to allow these experiments to be repeated. I think this needs rewriting, focusing on the information that is really needed to understand what is going on. I think it might be helpful to provide a paragraph at the beginning of this section to explain the logic of the order of presentation - I found some information I expected in one section in somewhere else completely, for example. Some of the information presented could be summarised in the form of a table and/or flowchart diagram, and this would certainly be helpful. Specific comments on individual parts of the Methods are given below. Section 2.1 LPJ-GUESS description. Given how important these simulations are for downstream results, it would be helpful to give a more detailed description of how the model simulates crops (i.e. what are the differences between the treatments of each crop type), how nitrogen limitation is handled, what information is used to specify nitrogen inputs to cropland etc. The information about how irrigation, water demand, water supply, and plant water stress are simulated may well be described in Alexander et al. (2018) but since these are crucial to the current simulations, the approach should be briefly described here. Even the description of how the model simulates natural vegetation types is given short shrift here, so that the claim that it handles CO<sub>2</sub> fertilisation is unsupported. It is also unclear from the present description how some of the service "proxies" are calculated by the model. For

C2

example, how does LPJ-GUESS simulate runoff? Please provide a better description of how the model works, so that it is easier to understand its strengths and limitations. Section 2.1 Ecosystem services. Most of Section 2.1 is given over to a description of ecosystem services. What I was expecting here was information about what model outputs were used as indicators of specific ecosystem services. However, much of the text describes why a particular service is important – which should have been information provided in the introduction and indeed partly is provided there. The description of the simulated index is brief and uninformative. What I think would be more helpful would be to reshape this in the form of a table, listing the service and the model output (or outputs). This would save some space which could usefully then be used to provide more details in the model description so that it is clear how these outputs are obtained. Section 2.2. Description of PLUM. Although a detailed explanation of the model is given in Alexander et al. (2018), it would be really nice to know a little more about it here. In particular, I am intrigued about the interface between the two models. What is the handshake, for example, between the four crop types in LPJ-GUESS and the seven crop types in PLUM? This is not explained here, nor is it explained in the description of the simulations. I do not understand how the crop demand optimisation works, and in particular whether this involves considering surpluses and surplus distribution (which should affect commodity prices) or whether it is assumed that there is always a surplus. Section 2.3. Given the complexity of the experimental design, the complicated linking of different models, and the multiple sources of inputs, I think it would be extremely helpful for the reader if you included some kind of flow chart here to guide us through. Section 2.3.3. The factor separation experiments are not well designed. Recycling 30 years of climate is not equivalent to a constant climate. As the results of the FireMIP experiments show, it is difficult to compare these constant climate experiments with constant other experiments when the constant other is based on a single year. Furthermore, the value of treating all climate variables as a single input seems a bit odd when thinking about productivity – it would be more interesting to diagnose what aspects of climate are crucial. In any case, a better factor-separation

C3

approach is needed. Alternatively, given that these results are "mostly not presented" (line189) you might leave this out.

Results and Discussion section. There is a lot of detail here, but the selection of things to highlight seems somewhat arbitrary. This is particularly the case in the delineation of geographic areas (what, for example, is meant by South Asia?). I was, for example, somewhat surprised by the lack of commentary on changes in China. Given that these kinds of assessments are of largely political interest, I wonder whether there should be some refocussing here - away from biggest changes to most important regions? Some thought should also be given to tabulating results. I would strongly advise separating out the Results from the Discussion, creating a separate section. There are many issues affecting the results presented here, including the impact of methodological uncertainties, that really need to be discussed more fully in this paper. How sensitive are the results to specific inputs? what is the impact of mixing static and time-varying inputs? given that there are large differences between vegetation models in terms of their predictions, how reliable are the LPJ-GUESS productivity estimates? or perhaps, where are they situated with respect to other models? and how much does this matter to the final assessment? How serious is the mismatch between PLUM outputs and the scenarios? How much of an impact does this have on the projections? I am not suggesting that these issues invalidate the study, but I think it would be helpful to discuss the sources of uncertainty and I suggest that you add a Discussion section, where you can do this.

One additional issue that could usefully be included in the Discussion, but certainly needs to be treated somewhere, is the assumption that increased fertilisation will always produce an increase in production rather than a saturating relationship, shown by analyses of field data.

A second issue that could usefully be included is "CO<sub>2</sub> fertilisation" – given that this still appears to be controversial, that there is confusion about this is photosynthesis or WUE, that different models produce different strengths of fertilisation and so on.

C4

In comparing LandSyMM results with other models, it would be useful to include a discussion of the plausibility (or otherwise) of their/your assumptions. This would also deal with the questions: given that there are other simulation results, what does this paper add? and why should we believe the results are more plausible?

I would seriously consider taking out the section 3.2.2, but in any case it needs re-writing. Runoff. The impact of CO<sub>2</sub> on runoff is going to be strongly dependent on whether we are talking about semi-arid regions or not, and there is now considerable literature on this (which should be cited). I think a more logical way to organise this section would be around climate regions. The transition from global runoff increasing to "flood" and "drought" risk is abrupt and it would be helpful to actually explain regional patterns of runoff change first. The fact that LPJ-GUESS is not a proper hydrological model, i.e. it does not transfer water between grid cells, it does include groundwater recharge, it does not include surface storage etc. etc. is mentioned in passing here (line 345). But this is a key issue about what "runoff" means and what "flood risk" means. This has been alluded to earlier on by referring to meteorological flood/drought, but it potentially very mis-leading – not for the immediate readers of the paper but certainly for the "assessments" that will pick these results up and re-use them. The logic of focusing on biodiversity hotspots is different from the logic employed with other ecosystem services, in the sense that with the other services you allow for increases/decreases and for changes in geographic regions where increases/decreases can happen. Wouldn't this be a useful approach here too? Is it possible that there would be increases in biodiversity in some regions that are not currently considered hotspots?

Conclusion. If you split the results and discussion section into two, the you could consider including the conclusions in your discussion section. The current conclusions are not very startling (storylines with high socioeconomic challenges to climate change mitigation consistently have the most severe consequences for ecosystem services) or are simply a repeat of how important this information could be (which was already in the introduction).

C5

#### Minor comments

Line 15-16. The statement about future population changes is expressed rather badly and is difficult to grasp, please rephrase. Line 25. Is this really a feedback *sensu stricto*? Lines 47-48. The processes operate on the plant functional types rather than among them. Can you rephrase this to describe the model more clearly. Line 51. When you say C3 cereals sown in winter and spring, presumably these are considered as two PFTs, so it would be clearer to say " C3 cereals sown in winter, C3 cereals sown in spring ...." Lines 68-75. It is impossible to judge whether these measures provide reasonable proxies for water availability, freshwater ecosystem condition, or flood risk because there is no information on how runoff is generated in LPJ-GUESS. is runoff simply the difference between P and ET in a gridcell? or is there transfer of surface water between gridcells? is there a contribution from groundwater? Line 75. If you mean that hydrologic drought is not the same as meteorologic or socioeconomic drought, why not simply say so? This sentence is unnecessary, and begs the question: what is e.g. socioeconomic drought. Line 76-81. How does LPJ-GUESS calculate total nitrogen loss? Do you separate out nitrogen loss from natural ecosystems and agricultural systems? Line 82. The linking of climate change and human health here led me to believe that you were going to look at ecosystem services that mitigated the impact of climate change on human health. Apart from the mention that BVOCs affect ozone which in turn can have impacts on health, you don't really go into this in any depth. For example, you don't mention e.g. mineral dust and the role that vegetation plays on mitigating dust emission pace China. Perhaps changing the emphasis here to plant emissions (which have multiple effects, including on climate and on health) would be a better way to introduce this section. Line 94-102. I can see why the focus on the hotspots is attractive, but in this modelling framework is would also be possible to make a more general assessment of biodiversity loss and this would also be valuable. Line 109. The plant name *Miscanthus* should be in italics.

Line 115. Is 500 years really sufficient to bring the carbon pools into equilibrium? or

C6

is the phrase a realistic starting point mean to imply that they are not necessarily in equilibrium?

Line 125-126. This first sentence should be moved into the description of the model, as confirmation that PLUM works reasonably well. It is not relevant to a description of the modelling protocol.

Line 127. Given that PLUM has 7 crop types and LPJ-GUESS only four, how do you input PLUM land use into LPJ-GUESS?

Line 129. Please can you bring this flowchart and the table into the main text?

Lines 133-135. Please indicate that the details for these sensitivity tests are given in a following section and reference the section here.

Line 139. Surely this should be: Viovy, N. 2018. CRUNCEP Version 7 - Atmospheric Forcing Data for the Community Land Model. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. <http://rda.ucar.edu/datasets/ds314.3/>.

Line 140. Either spell out what these problems are or refer to a paper that does. Maybe Tang et al. (2017)?

Line 147-149. I am having difficulty with this description. You use time-varying allocations of cropland area per gridcell but a static data set of what these crops were. How did you apply this? Simply assuming that the area might change but the crop remains the same? How much uncertainty does this introduce in your results? Unless you address this in a Discussion section (as suggested above) you need to say something here.

Line 148. Is this still in prep.?

Line 154-155. This sentence is a bit unclear. The manure N is held constant in the calibration run but varying in the other runs?

C7

Line 179-180. Did you estimate these values or are they provided?

Line 184-185. Recycling 30-years of climate is not "constant climate"

Line 210. "first two or so ...." please state what period it actually decreases over.

Figure 2. This is unreadable at the size reproduced here and, with the grey background, the paler colours are not sufficiently visible. You need to find a way of making the changes more visible. Maybe splitting this into two figures would help. (Note the same comments apply to Figure 4, 5).

Line 222. I confess that I find this agricultural expansion in Alaska a bit implausible, even for a high-end scenario, given the topographic constraints and the issue of permafrost. I would be intrigued to know when the permafrost disappears in this scenario. And how infrastructure (or lack of it) would impact this expansion?

Line 228. Since South Asia is not a widely-recognised geographical term, it would be helpful to define where exactly you mean here. Are you including southern China here?

Line 234. What climate change produces more favourable growing conditions in South Asia?

Line 236. Even larger ... even larger than what?

Line 250 et seq. I find this discussion of other model results here confusing. I think you want to separate this from the presentation of all the results from your experiments and move this type of comparison into a separate discussion session. This would allow you to discuss the plausibility of the other assumptions compared to the assumptions encapsulated in your simulations.

Line 267. I do not understand what you mean by "friction" here.

Line 269-274. Please take out this speculation about the impacts of including forest products in LandSyMM based on work that has not yet been done.

C8

Line 275-283. And so what? You appear to be saying that you have different results from one study because they used unrealistic inputs, and that you have different results from another study because they made a set of different assumptions. In the first case, perhaps you could assume that a reader might guess that your results are "better", although you never actually use the "unrealistic" word. In the second case, however, you might give use a hint about the assumptions made by IMAGE would produce more/less realistic results and why.

Line 290. "intermediate carbon fertilisation ..." Not phrased felicitously, since it implies that C-fertilisation itself has multiple levels of working (off, half-on, on). Please rephrase.

Line 299. Sorry, I might have missed this – what do you do about the conversion of secondary vegetation to pasture in terms of carbon. Are we looking at gross or net here? Line 302-303. I am not sure why you are picking out one model from this study. I think you should give the range of estimates. I don't know whether the quoted value for IPSL-CM5A-LR is low-end or high-end. Figure 3. Please don't abbreviate emission on this Figure. Line 305. "probably"? You could establish this by looking at what the difference in loss of non-agricultural land between these simulations and yours is if you take out the pasture expansion and the expansion of cropland in Alaska. Line 308 et seq. So the difference is caused by the differences in the scenarios, right? But later on you imply its because the models don't include nitrogen-limitation. I think you need to make it clear what you think is giving rise to these differences, scenarios or model set-up. It would, of course, make it interesting to run your experiments with the older scenarios - and this would be helpful in terms of uncertainty analysis. Line 309-310. Your comparisons with other simulations are unbalanced - having described the results from the Brovkin et al study in some detail you here say that the results are low compared to the Nishina et al. (2015) study even when comparing just to the models in that study with nitrogen limitation. But no details. How low? what was the range simulated by the N-enabled models in Nishina et al.? Do you have any idea why you

C9

get a different result? Lines 311-318. So, different models produce estimates less than LandSyMM as well as above. What do we learn from this? You imply this is because of differences in scenarios (while hedging your bets in terms of climate forcing), but what is needed here is a back-of-the-envelope calculation of whether the differences in cropland and pasture area would produce a comparable estimate in LandSyMM. If you wanted to be ultra-realistic, you could use the areas where they show biggest changes in area. Line 319. Photosynthesis scaling parameters ..... what scaling parameters? Line 338. Why did you not run it in coupled mode then? Line 341. Please can you explain what was done in the Asadih and Krakauer (2017) analyses. Were these full hydrological models? Since this was an ensemble, presumably there is a range of estimates for at-risk of flood and at-risk of drought? Please give these ranges in the text. How much of a difference does using monthly versus daily values make to the estimates of area affected? Line 349. How many classes are there? You need to spell out that you are talking about increases/decreases in flood risk and drought risk areas). Line 379. Please explain why high CO<sub>2</sub> suppresses BVOC formation and provide references here. Line 382-383. This sentence is confusing because it seems to say that boreal forests are causing declining monoterpene emissions – whereas I think the idea is that decreases in boreal forest area coupled with less effective BVOC production in the surviving boreal forest area are responsible. Please rewrite. Line 393. Please italicise *Miscanthus* Line 396 et seq. This paragraph states that predicting the effects of changing BVOCs is difficult because the model framework doesn't include atmospheric chemistry. I am not sure what "a surface-level discussion of possible effects" means here. You can predict potential changes in BVOC emissions, and so perhaps this is the point to stop at. There is no need to go further and speculate about what the impact of these changes might be on atmospheric chemistry, climate and/or health. Line 403. The fact that this one region is not defined as a hotspot, and that it has a big impact on the results, makes a good case for extending this analysis to consider changes in biodiversity everywhere. Line 410-415. Please rewrite this section to first make the comparison and then explain possible sources of differences. Line

C10

415. Are you saying that species-area curves are an inappropriate tool for estimating extinction rates rather than species numbers? Lines 419-420. I hadn't realised that climate changes and CO2 could have an effect on models too! Please rephrase this. Line 423. "We may see a similar effect ...." Please clarify: do you or don't you? Line 429-430. Considering that the results from LandSyMM have been compared to a range of other model simulations in this paper, the first sentence really doesn't make sense. Maybe this is more comprehensive in terms of the range of scenarios and the range of outputs, but what else is different here?

---

Interactive comment on Earth Syst. Dynam. Discuss., <https://doi.org/10.5194/esd-2019-44>, 2019.