Biogeosciences Discuss., 11, C3865–C3873, 2014
www.biogeosciences-discuss.net/11/C3865/2014/

Biogeosciences
Discussions
Open Access

Interactive comment on "Structural analysis of three global land models on carbon cycle simulations using a traceability framework" *by* R. Rafique et al.

Anonymous Referee #2

Received and published: 28 July 2014

1 General comments

In this paper, the innovative traceability framework established and applied to CABLE in *Xia et al.* (2013) is applied here to CABLE, CLM3.5-CASA and CLM4.

Although the traceability framework is intriguing, it is unclear which conclusions, if any, could not have been gained through simple examination of the flowcharts presented in Figures 2-4, which summarize model structure, pool size and residence times, or from Figures 1 and 6, which display model output. The results and their discussion do not clearly outline the relevance of findings, or how these might be used to inform

model development, benchmarking or analysis, although this is presented as a central contribution of the paper. The traceability framework is also mentioned as a strategy to characterize the C cycle in models, but conclusions regarding model complexity are made on the basis of the number of pools, not on traceability results.

The main differences found relate to differing spatial distributions of NPP and ecosystem C storage between CABLE and CLM*. CABLE has much larger ecosystem C storage in arctic regions than CLM*, and CABLE indicates boreal regions as having the largest NPP, whereas CLM* have the largest NPP in tropical regions. Which of these models agrees more closely with MODIS NPP, or modeled residence time from HWSD and MODIS NPP? What drives these differing patterns? If the greatest differences observed are spatial, why is the traceability framework then applied at the mean global scale?

The traceability framework seems to have primarily been applied to determine the mean global NPP and mean global residence time of three model. How does this "help fully characterize the behavior of complex land models"? What insights does the traceability framework provide which other appraoches cannot? There are mentions throughout the paper of the benefits and importance of the traceability framework, such as: "This study provides useful information for data assimilation, benchmark analysis and future model development by evaluating the relative importance of model components and source of variations." However, these statements are empty unless the advantages of the traceability framework are well explained in the introduction, and unless the results, discussion and conclusions make use of the unique contributions of the traceability framework, and provide tools or insights which could assist with model development, benchmarking or data assimilation.

2 Specific comments

- Was ecosystem C storage similar or dissimilar among models? Why did certain models have more or less C storage than others?
  - Abstract lines 11-12: "the spatial distribution of total ecosystem C storage and residence time differ greatly among the three models"
  - Discussion p.9994 l.16-17: "The land models used in this study dispalyed considerable spatial variations in the global distribution of NPP, total C storage and C residence time."
  - Discussion p.9996 l.4-8:"the NPP is not highly variable among the models due to similar environmental conditions. As expected the resultant C storage was also very similar among the models. The variations amongst these models can also be largely explained by examining the C transfer rates from one pool to the others"
- Abstract lines 12:15 describe comparisons of models against "measured" C storage and residence time. It would be helpful to mention which measurements are used, and to refer to these as estimates, as MODIS NPP is not a true measurement. If there is a lack of space, the sentence beginning with "However," could be deleted entirely as it repeats details already mentioned in the previous sentence.
- The introduction aptly mentions that the strong dependence of C storage on NPP and residence time has been well-established in previous research. Residence time is then described as less studied, and challenging (p. 9981 l. 19:21). Some further explanation and justification of these statements on residence time is required. Why is residence time challenging? What aspect of the association between residence time and NPP has not previously been well established, but can now be described using the traceability framework.

- The third paragraph on p.9981 and first paragraph on p.9982 provide justification for using the traceability framwork on the basis that models are complex, and their uncertainties need to be attributed to their sources. It concludes with: "For example, Mishra et al. (2013) have identified the modeling uncertainties of soil C in permafrost regions but insufficiently attributed these variations to their sources. This shortcoming can only be addressed after gaining a thorough understanding of the model's fundamental structural differences and understanding its traceable components controlling the C dynamics." In this paper, is the traceability framework applied to attribute model uncertainties to their sources? If so, how is this achieved, and where are the results from this analysis? The traceability framework appears to describe differences in C dynamics according to differences in soil, vegetation and litter pools, not uncertainties. The comparison of modeled residence time to estimated residence time using modeled residence time from MODIS NPP and HWSD represents a further model output intercomparison, not an attribution of errors to their sources.
- A stated goal of the traceability analysis is to "characterize the complexity of C cycle in the models" (p. 9983 l. 2-3) but in the results sections, the number of pools alone is used to inform comparisons of model complexity. How did the traceability framework assist in this stage? Isn't it standard for model intercomparisons to assess differences in model structure to explain differences in model outputs?
- The description of models could safely be moved from out of the methodology section, since they were not built or modified for this paper. Furthermore, the information presented on these models differs widely, and makes it difficult for readers to examine differences between models. I would suggest rewriting and lengthening these descriptions so that they provide the central equations, environmental scalars, and processes through which C is transferred in order to inform later discussion. It is also crucial that obvious statements (e.g. that CLM3.5 is used for climate change simulations, or that CO_2 is released through respi-

ration), or any information provided directly in the flowcharts or introduction, be removed to reduce redundance. The descriptions at present are not informative for interpreting results.

- Why was the SASU method used for only two models? Could this have affected the output?
- Why do the models use different meteorological forcing? If different forcings are used then there must also be some comparison provided of the forcings. There are otherwise too many confounding factors which complicate any true model intercomparison.
- "These simulations were customized to not account the effect of any disturbance effect" is unclear, and needs to be rephrased.
- Why were three different resolutions used for model runs? What does this assist with? How does the traceability framework make use of multi-scale data? Providing justification for the method used, and explanation of its implementation, should be established in the methodology section.
- The first section of the diagnosis provides motivation or context, which belong in the introduction. Instead, this subsection should fully describe the implementation of the diagnosis.
- The fraction of C entering each pool is described as a hidden variable; however, can this not be easily determined from model output? Why is a traceability framework required?
- Similarly, why is a traceability framework required in order to determine accumulation in pools?

- Does Figure 6 represent mean values from 2000, the year for which MODIS NPP was used? Please also provide diagrams of the "measured" data (modeled from MODIS NPP and HWSD), as well as figures indicating the difference between spatial outputs from these models and the "measured" data.
- How did you combine HWSD with MODIS to get C residence time? Do you consider the resulting estimates to represent measurements of C residence time, or would these better be considered as 'estimates'? Has this appraoch previously been applied? How accurate are the HWSD and MODIS NPP considered?
- All context and justification belong in the introduction, and all methods belong in the methodology. Portions of the second paragraph of the results belongs in the methodology section. Restructure the methodology and results sections so that they correspond to one another.
- Why did you do a comparison of spatial, global-scale model estimates from process-based models (CABLE, CLM*) with MODIS NPP (?) using barplots of mean global values? Over which period of time were these values averaged? Were these results generated using the traceability framework, or are they simple mean values from model output?
- Why are spatial patterns of NPP and ecosystem C storage more similar in CLM* than CABLE?
- The description of model allocation schemes should be one of the central components of the model description subsections, presently in the methodology. This only belongs in the discussion if it is used in the interpretation of results.
- Are the allocation schemes in CABLE and CLM 3.5 constant? If so, what purpose did the traceability framework fulfill? It seems a major part of this project involves summarizing what is already known about these three models.

- Subsection 3.3 should be summarized in a table, and the results should be described in relation to existing findings rather than merely repeated.
- CABLE seems to represent fine roots as having the same residence time as wood, which is contrary to many other models, and contrary to what is known. CABLE is described as having the most realistic spatial distributions of NPP, and as having the best performance in terms of C storage capacity (concusion: p. 9999 l.8-12). On what basis do you decide that CABLE performs best in terms of C storage capacity, when its roots and wood have the same residence time? This would be one of many central points that should be brought up in the discussion.
- It seems the final result of this paper is summarized in Figure 8, which shows a scatter plot of NPP and residence time for the three models. However, only one value per model is shown in each scatter plot. Is this the global mean NPP vs global mean residence time? What insights into model performance does this provide? The scatter of the points does not appear to have a strong shape or direction, and these plots are not well interpreted in the discussion.
- Why are mean global values used, when the largest differences in models exist at the spatial scale? Can the traceability framework be applied over selected regions to gain insights into why
- There are many unfounded statements in the last section of the discussion, conclusions, and abstract that describe what the traceability framework will do, or has done, but these appear unaccomplished in the text. Can the residence time, and associations with NPP, be acquired from model output without a traceability framework? Which advantages does a traceability framework provide?
- A central result from the traceability analysis of CABLE by *Xia et al.* (2013) is that environmental scalars control C dynamics. It is strange that the analysis presented in this paper was then not informed by this finding, but instead repeats this

recommendation. “One way to elaborate the analysis is to further examine the effects of environmental scalars and environmental forcing data which strongly influence the C residence time and transfer from one pool to another.” Why were these environmental scalars not presented? If meteorological differences drive differences in model behaviour, then why were different environmental met forcings used to drive models?

- The main conclusion from this type of analysis cannot be simply that there are differences in models because of differences in C transfer and pools. After completing this analysis, the results should be fully describe why certain models have greater/smaller NPP, residence times, storage etc., and describe these differences according to specific aspects of model structure.
- Provide justification for this statement: “The results of this research will provide valuable information for the future study of model development, data assimilation and benchmark analysis.” How can model development, data assimilation or benchmark analysis be improved using the results and discussion you have presented?

3 Technical corrections

- The clarity of the ideas presented in this manuscript is undermined by incorrect use of the English language. Grammatical and spelling errors need to be fixed, and sentences need to be better structured.
- The paper as a whole needs to be restructured to allow readers to quickly grasp what is being presented (e.g. methods in methodology section, results in the results section).

4 References

Xia, J., Luo, Y., Wang, Y.-P., and Hararuk, O.: Traceable components of terrestrial carbon storage capacity in biogeochemical models, Global Change Biology, 19, 2104–2116, 2013.

Interactive comment on Biogeosciences Discuss., 11, 9979, 2014.