Dear editor of Earth System Dynamics

We would like to thank you and anonymous reviewers for your valuable comments and feedback on our manuscript "*Divergent predictions of carbon storage between two global land models: Attribution of the causes through traceability analysis"*. We have revised the manuscript according to reviewers' suggestions and comments. We provide a detailed response to reviewers' comments, point by point, in the "Authors' response" document (on next page of this letter). Specifically, we have revised results, discussion and summary sections, updated table 1, added supplementary material. All the major changes made are highlighted in red color. We believe that the core message of the study is better communicated now. We will be more than happy to improve our paper, if required.

Yours Sincerely,

Dr. Rashad Rafique

Pacific Northwest National Lab, USA

Phone: (608) 770-4236

E-mail: rashidbao@gmail.com

Anonymous Referee #1

Received and published: 28 September 2015

This manuscript examines the carbon cycle of two models, CLM-CASA' and CABLE, using a framework to trace carbon through the system. The traceability framework has been used and published previously – it is useful for comparative studies of the simulated carbon cycle in ecosystems. I appreciate the utility of the traceability analysis for identifying structural or parameter uncertainties in the models, and I think this paper provides a clear comparison between two widely used models. However I have two main concerns with the manuscript as it stands now. These are given below, followed by some specific comments/questions for the authors.

First, for a fair comparison between the models the same climate forcing should be used. Even if the mean values of temperature and precipitation are the same, it's more likely the extremes of these that have the largest impacts on the NPP and residence times. I am not convinced that some of the differences between the models is not due to the different climate data sets used.

Response:

We acknowledge reviewer's concern for not using similar forcing data as the extreme values may produce different results for different variables. This makes more impact if we focus on the temporal analysis of variables under consideration. In this study, we mainly focused on the global means of all variables after spin-up period to the steady states. We think the use of different forcing data, if their long term means are comparable, will not largely affect the results; however, we agree that the use of similar climate data produces the most ideal results. We have addressed the potential source of uncertainty due to difference in climate data in discussion section of the revised manuscript.

Second, the diagnosis that the differences in NPP are due to differences in Vcmax and SLA is speculative (at least as it appears in the text). External factors can affect the NPP such as temperature, radiation, or precipitation, and as I stated above I am not convinced these differences are not important. Also, internal model differences can affect NPP such as how canopy radiation and moisture stress are handled. Last, five biomes have either a lower SLA (ENF, EBF) or a lower Vcmax (DBF, C3 grass, Tundra) in CLM-CASA' despite that model having a higher NPP – so it is not clear to me the role of these parameters in determining the relative NPP.

Response:

Throughout the manuscript we have tried to convey the idea that in the two carbon cycle models there are major structural commonalities. Conceptually both models are similar: organic matter decomposition was a function of temperature, soil moisture, organic matter quality, and soil texture, therefore the differences in model output were not due to structural differences in the models, but due to differences in parameter values. However, it seems not to be the case for

NPP: as reviewer notes, CLM-CASA' NPP is always higher than CABLE NPP, which is not consistent with Vcmax being lower in CLM-CASA' for some biomes. The formulation of the photosynthesis modules differ between the two models: unlike CABLE, CLM-CASA' has moisture limitation (but it doesn't reduce NPP); CABLE uses combined model for C3 and C4 photosynthesis, whereas CLM-CASA' separates C3 and C4 photosynthetic pathways; the formulation of the RuBP-limited photosynthesis rate is different between the two models [section 3.2.4 c in Kowalczyk et al. (2006) and section 8 in Oleson et al. (2004)]. Autotrophic respiration was simulated differently between the two models: in CABLE wood and root respiration was proportional to their carbon pool sizes, and leaf respiration was proportional to the carboxylation rate; CLM-CASA' calculated whole plant autotrophic respiration as half of GPP. These differences also may have contributed to the differences in whole plant NPP.
We included these remarks in section 3.3 and the first paragraph of the discussion.

Specific Comments:
1. I think some clarification on the role of PFTs and biomes in the carbon pools would be helpful. I have a few specific questions: - I am curious how you handle multiple PFTs within a gridbox in the traceability framework. How many pools does each model have in total (what is n in the equations)? For example, does each PFT in CLM-CASA' have 3 plant carbon pools? Are the litter and carbon pools shared between PFTs? - How is the translation from PFTs to biomes done? A table showing the 16 PFTs of CLM-CASA' and which biome each fits into would be useful. - Please provide the full names of the pools in the caption of Figure 3.
Response:
We worked with the model output, so we assigned 1 pft per gridbox, i.e. the pft that occupied most of the area in a gridbox (CLM-CASA' simulates up to three pfts per gridcell). CLM-CASA' has 12 carbon pools (simulated at the pft level), and CABLE has 9 carbon pools.
We included the table below as a supplement to the manuscript and provided full names of the pools in the captions to Figure 3, for further clarity.

CLM-CASA' PFT	CABLE-compatible biome
needleleaf evergreen temperate tree	evergreen needleleaf forest
needleleaf evergreen boreal tree	evergreen needleleaf forest
needleleaf deciduous boreal tree	deciduous needleleaf forest
broadleaf evergreen tropical tree	evergreen broadleaf forest
broadleaf evergreen temperate tree	evergreen broadleaf forest
broadleaf deciduous tropical tree	deciduous broadleaf forest
broadleaf deciduous temperate tree	deciduous broadleaf forest
broadleaf deciduous boreal tree	deciduous broadleaf forest
broadleaf evergreen shrub	shrubland
broadleaf deciduous temperate shrub	shrubland
broadleaf deciduous boreal shrub	shrubland

c3 arctic grass	tundra
c3 non-arctic grass	c3 grassland
c4 grass	c4 grassland

2. At the end of Section 2.1, the q10 values are given but not their context – do these apply to soil respiration?

Response:

The context of Q10 value has been added in section 2.1 of revised manuscript.

3. What determines the baseline residence time? Why do deciduous needleleaf and evergreen needleleaf forests have the highest baseline carbon residence times? And why is there such a large difference in the baseline residence time for tundra between the models?

Response:

Baseline residence time is determined by the matrices of A and C as well as vector B. Where, A represents the carbon transfer among pools, C represents the respirational losses and B represents the partitioning coefficients. The A, C and B components are determined by soil texture and vegetation lignin fraction.

The deciduous needleleaf and evergreen forests showed highest baseline residence time because they partitioned the largest fraction of NPP to woody biomass. For tundra the baseline residence times differed also likely due to the partitioning coefficients, because both models simulated similar environmental scalars of 0.1.

This has been clarified in discussion section of revised manuscript.

4. The residence time is higher in CABLE – and this is attributed to higher residence times in wood and a higher allocation of NPP to wood. Could another reason be that once carbon enters the passive SOM pool in CABLE, it does not interact with the slow pool? What is the effect of the more complicated interactions between soil pools on the residence times in CLM-CASA'?

Response:

The more complicated interactions between soil pools in CLM-CASA' slightly increase the residence time (but not significantly), because instead of leaving the system, carbon returns to another pool, thus staying in the system longer.

To demonstrate that, we can introduce similar interaction to the CABLE model structure (i.e. return of carbon from passive pool to slow pool, and from slow pool to fast pool), using the same two transfer coefficients as in CLM-CASA' (0.45 and 0.4). If we assume NPP is 600 g C/m2/year, and the parameters in CABLE are those listed on Figure 3 in the manuscript, the ecosystem carbon residence time will be 13.06 years. If we modify the model structure to include the interactions mentioned above, the ecosystem carbon residence time will be 13.68 years. This experiment illustrates that interaction between soil pools increase carbon residence time, but not significantly.

This has been clarified in the discussion section of revised manuscript.

5. At the beginning of Section 3.5, I think it should be reiterated that a lower environmental scalar limits decomposition and turnover and therefore increases the final ecosystem residence time of carbon.

Response:

The suggested information by reviewer has been incorporated in the beginning of section 3.5 of revised manuscript.

6. What is meant by the last sentence in the first paragraph of the Discussion: "Longer ecosystem residence time in CABLE was mainly attributed to higher environmental limitation of the organic matter decomposition." This seems to be in contrast to what is shown in Figure 6 – which shows a larger difference in the baseline residence time than in the environmental factors.

Response:

After taking consideration of reviewer's comment, the confusing sentence has been removed from the revised manuscript.

7. In general the abstract and Discussion/Summary sections communicate the main results of the study very well. I have one suggestion: the Discussion mostly addresses the 3 objectives given in the Introduction, but it could improve the paper to more explicitly address these objectives and the main conclusions pertaining to each.

Response:

We appreciate reviewer's constructive feedback. In order to make this study more consistent, we have clearly stated the objectives of the study. We have also highlighted the main take away messages of this study on abstract and discussion sections. We hope this will be helpful in conveying the overall message of this study.

Kowalczyk, E.A. et al., 2006. The CSIRO Atmosphere Biosphere Land Exchange (CABLE) model for use in climate models and as an offline model. CSIRO Marine and Atmospheric Research.

Oleson, K.W. et al., 2004. Technical description of the community land model (CLM). NCAR Tech. Note NCAR/TN-461+ STR: 173.

Comments from Reviewer 2

The manuscript submitted by Rafique et al entitled "Divergent predictions of carbon storage between two global land models: attribution of the causes through traceability analysis" is an interesting work on the behavior of land models. The authors used an analytical approach (traceability framework) to decompose model predictions of ecosystem carbon (C) storage into a set of common parameters. The authors also made good attempt in writing the manuscript, however, there are few concerns (mentioned below) those can help in improving the manuscript. I believe that after considering below mentioned concerns, this manuscript can be suitable for publication in Earth Syst. Dynam. Discuss journal.

Major Comments

(i). The traceability framework used in this study can be discussed in more detail to make it more compelling. This can be done through the better articulation in the objectives section. In the Results section, the text mainly focusing on detailing the differences in models which is good but should not be stretch too much and rather authors should focus on the key differences and the importance of those differences in modeled NPP and carbon storage.

Response:

This study mainly focused on the application of traceability framework. Therefore, the detailed explanation of the traceability was avoided. We have cited the following reference, which has developed and explained the framework. However, in order to accommodate reviewer's concern we have also revised the objective section of the manuscript.

Xia, J. Y., Y. Q. Luo, Y. P. Wang., E. S. Weng and O. Hararuk:. A semi-analytical solution to accelerate spin-up of a coupled carbon and nitrogen land model to steady state, Geosci. Model Dev, 5, 1259-1271.

The minor details from the results section have been removed. The current version of manuscript shows the key differences between two models, CLM-CASA' and CABLE, in estimating carbon storage capacity.

(ii). The discussion section is reasonably organized and describes a summary of the differences in the models. However, this section needs to focus on the model performance, and why this approach is most useful than previously studied. Also, what are the model uncertainties?

Response:

We appreciate reviewer's concern about the models performance. However, the objective of this study was to implement the traceability framework for the relative comparison of the two models in estimating their carbon storage capacities. Evaluation of models performance against any benchmark was not the aim. We acknowledge reviewer's idea in evaluating these models and verifying the applicability of traceability framework. We would consider this idea in future research.

(iii). The summary section of should focus on key findings and take away message rather than repeating most of the results.

Response:

The summary section has been revised to highlight the key findings. The repeating results, if found unnecessary, have been removed from the summary section.

Specific Comments

(i). Abstract: The abstract is well written, however, it can be further improved by more focusing on the take away message. Also, highlight what we learned from this study.

Response:

Takeaway:

The following takeaway message has been highlighted the manuscript.

"Overall, the traceability analysis showed that the major causes of different carbon storage estimation were found to be parameters setting related to carbon input and baseline residence times between the two models."

What we learned:

The biggest challenge to model diagnostics is model intractability. The more processes incorporated; the more difficult it becomes to understand model behavior. As a result, uncertainty in predictions among global land models cannot be easily attributed to their sources. The framework used in this study analytically decomposes complex land models (CLM-CASA' and CABLE) into traceable components, which are helpful in attributing the models variations to their respective sources. This study showed that the models predictions were mainly controlled by the parameters differences.

These messages have been mentioned in the manuscript.

(ii). Introduction: Authors mention that the future CO2 concentrations depend on the balance of C uptake and C loss from ecosystems. Why is "in simulations" used? The future of CO2 concentration depends on how the terrestrial carbon cycle will (actually) respond to various external factors, not on how we simulate it. Further, the sentence needs work...Many studies have evaluated and compared the carbon cycle components of ESMs...Also you focus on ESMs here, but the analysis presented in the paper is using land models (not ESMs). This distinction is not clear.

Response:

The word "simulation" has been removed from the revised manuscript. Also, the world "earth system model" has been replaced with "land models" in the revised manuscript.

(iii). Materials and Methods: Again, here referring to ESMs, when analysis is focused on land models. If CLM-CASA' and CABLE are forced with different climate drivers? If so, this needs to be made clear. Somewhat glossed over here.

Response:
The word "ESM" has been replaced with "land models" in the revised manuscript.
The detailed analysis of climate forcing, used in this study, is given in the section 3.4 of manuscript. This analysis shows that the two climate forcing data sets were not substantially different at global level. However, at biomes level, few biomes show some differences.

(iv). Results: The statement "In general, biomes with higher carbon storage capacity of models, showed moderate NPP and higher ecosystem residence times" does not seem to accurately describe the relationship between Uss and τE. This only seems to describe ENF. Please check it again. The sentence "Three biomes, evergreen broad leaf forest, C4G …." is unnecessary and too wordy. "similar diverse"??? This needs to be fixed. Majority of the text is describing figures only. Please shorten the text (as commented above in the General Comments) to highlight the main points and their importance.

Response:
We agree with reviewer on the statement "In general, biomes with higher carbon storage capacity of models, showed moderate NPP and higher ecosystem residence times". After careful observation, we found a complex pattern (shown in Figure 1). Therefore, we have removed the statement from revised manuscript.
The too wordy sentence "Three biomes, evergreen broad leaf forest, C4G …." has been also removed from the revised manuscript.

The sentence starting with "similar diverse ..." has been fixed in revised manuscript.

Overall, text has been shortened, as much as possible, to highlight the main results of this study.

(v). Summary. See response in above General Comment section.

Response:
Summary section has been revised in manuscript (see earlier response).

(vi). Figures: Figures can be improved by mentioning in the caption about the black circle and the open square symbols. These things have not been mentioned in the figure 1. Same apply for the Figure 2. Further, in Figure 4, the time period of the weather data should be mentioned.

Response:
Figures' captions have been updated in revised manuscript, where necessary. The weather data description has been added in the method section of manuscript.

Anonymous Referee #3

Received and published: 10 November 2015

General comments

The authors have applied the recently developed traceability framework for benchmarking terrestrial carbon cycle models (Luo et al (2012), Xia et al (2013)) to two such models, to compare their simulated ecosystem carbon storage capacity and to explain the differences they find. The study demonstrates the power of the traceability framework approach in elucidating the mechanisms underlying differences in behaviour between models. The outcome of this work is suggestive of how useful a larger study with a greater number of terrestrial carbon cycle models could prove to be. The paper is generally well written and clear.

Specific comments

The only concern I have with this work is the difference in forcing used to drive CABLE and CLM CASA'. The authors do address this, and demonstrate that the forcing is largely comparable, although for some biomes the differences in precipitation and air temperature were significant. A repeat simulation by one model with the forcing of the other (or the same for both) would be useful to indicate how important this difference in forcing might be to the overall result though I suspect it would not change the results significantly. However should a similar study be undertaken with a greater number of land surface / terrestrial carbon cycle models I would hope common forcing to be a feature. Also, it would be good to see a little more on the soil carbon stores which are barely mentioned in the paper.

Technical corrections

Page 1580 line 14: was a function OF the
Response:
We have corrected this in the revised manuscript.

Page 1580 line16-17 – Lines 17-19 constitute a more detailed version of lines 16-17 so presumably lines 16-17 should be deleted.
Response:
The lines 16-17 have been removed from the revised manuscript.

Page 1581 line 1: should be (Sitch et al. 2015)
Response:
We have corrected this in the revised manuscript.

Page 1581 line 7: The "The" at the start of the sentence is a bit unnecessary.
Response:
We have corrected this in the revised manuscript.

Page 1581 line 19-20: “for the period 1850-2100” makes more sense

Response:

We have corrected this in the revised manuscript.

Page 1582 line 18: Luo et al 2003 is not referenced – the Luo et al 2001 reference listed is actually from 2003 though so I think it is that which needs changing.

Response:

Luo et al., (2003) and Luo et al., (2001) are two different studies. They have been checked and clarified in the manuscript.

Page 1582 line 18: Should this be Luo et al (2003) given the above?

Response:

Luo et al., (2003) and Luo et al., (2001) are two different studies. They have been checked and cited properly in the manuscript.

Page 1583 line 22: Missing comma after CLM.

Response:

We have corrected this in the revised manuscript.

Page 1584 line 9: Could perhaps do with another sentence giving a bit more detail as to how the aggregation was done.

Response:

We worked with the model output, so we assigned 1 pft per gridbox, i.e. the pft that occupied most of the area in a gridbox (CLM-CASA’ simulates up to three pfts per gridcell). CLM-CASA’ has 12 carbon pools (simulated at the pft level), and CABLE has 9 carbon pools.

We included the table below as a supplement to the manuscript and provided full names of the pools in the captions to Figure 3.

CLM-CASA' PFT	CABLE-compatible biome
needleleaf evergreen temperate tree	evergreen needleleaf forest
needleleaf evergreen boreal tree	evergreen needleleaf forest
needleleaf deciduous boreal tree	deciduous needleleaf forest
broadleaf evergreen tropical tree	evergreen broadleaf forest
broadleaf evergreen temperate tree	evergreen broadleaf forest
broadleaf deciduous tropical tree	deciduous broadleaf forest
broadleaf deciduous temperate tree	deciduous broadleaf forest
broadleaf deciduous boreal tree	deciduous broadleaf forest
broadleaf evergreen shrub	shrubland
broadleaf deciduous temperate shrub	shrubland
broadleaf deciduous boreal shrub	shrubland
c3 arctic grass	tundra
c3 non-arctic grass	c3 grassland
c4 grass	c4 grassland

Page 1584 line 12: Q10 is used without any prior explanation as to what it is so a few words stating what it is would clarify

Response:

The role of Q10 has been clarified in the revised manuscript.

Page 1584 line 19: using THE following equation

Response:

We have corrected this in the revised manuscript.

Page 1584 line 22: ...vector of length n "representing the carbon pool sizes"?

Response:

Yes and this has been clarified in the revised manuscript.

Page 1585 line 1: "exit rates of carbon left in pool" is not all that clear – presumably it means the rate of loss of carbon from each pool via decay or respiration

Response:

The "exit rates of carbon left in pool" has been replaced with "carbon losses through respiration" in the revised manuscript.

Page 1585 line 12: here the baseline ecosystem residence time is a function of A, C and B, but on page 1582 line 24 you state that the baseline carbon residence times are usually preset in a model according to vegetation characteristics and soil types. Confused me at first, but on re-reading I think it is the case that A, B and C are all usually function of model parameters and therefore it follows that the baseline ecosystem residence time is also? If so this could be spelled out a little more.

Response:

We appreciate reviewer's deep concern on this point. We acknowledge that A, B and C are functions of model parameters, which are usually preset in models, based on vegetation and soil characteristics. This has been clarified in the introduction section of revised manuscript.

Page 1586 line11: Xia et al 2013 not 2012

Response:

Our apology, the reference by Xia et al 2012 has been added in the revised manuscript.

Page 1587 first paragraph. A table summarising the residence time, NPP and ecosystem carbon storage capacity for each pool for both models might be a useful

Response:

We appreciate reviewer's concern on the presentation of results. However, after considerable thinking about this suggestion, we believe a diagram representation is preferred over tables. The figures 3 and 6 represent the residence times, NPP and carbon storage capacities for the two models.

Page 1591 lines 1-2: ...and evergreen broadleaf forest AND SHRUB? in CABLE, and evergreen broadleaf forest, C4 GRASSES, SHRUB in CLM-CASA',

Response:

This has been clarified in revised manuscript with the following sentence:
"The C4G, evergreen broadleaf forest and shrubs in CABLE and C4G, shrubs and evergreen broadleaf forest CLM-CASA', showed the highest temperature scalar values amongst all other biomes, respectively".

Page 1591 lines 4: tundra in both CABLE and CLM-CASA' ?

Response:

The information about "tundra in both CABLE and CLM-CASA'" has been updated in the revised manuscript.

Page 1591 lines 8: ...(0.87) in CABLE and EBF (0.98?) in CLM-CASA' ?

Response:

The correct values have been verified and updated in the revised manuscript.

Page 1591 lines 9-10: Overall, the lowest water scalar was DNF in CLM CASA' and the lowest temperature scalar was Tundra in CABLE?

Response:

The suggested information "Overall, the lowest water scalar was observed in the deciduous needleleaf forest in CLM-CASA' and the lowest temperature scalar was observed in Tundra in CABLE" has been added in the revised manuscript.

Page 1591 line 11: ...for most biomes

Response:

This is corrected in the revised manuscript.

Page 1595 line23: , 2003 not 2001

Response:

Luo et al., (2003) and Luo et al., (2001) are two different studies. They have been clarified and properly cited in the manuscript.

Page 1598 final two columns might be a little clearer if CABLE was expressed as a fraction of CLM-CASA' or the other way round, rather than one minus the other.

Response:

The table 1 in revised manuscript has been updated with the fractional differences of parameters between CABLE and CLM-CASA'.

Page 1599: presumably the units of the grey contour lines are kgC ?

Response:

Yes, thank you. The unit has been added to the caption of figure 1.

Page 1604: line 4: T and W need subscripting.

Response:
The subscripts have been fixed in the revised manuscript.