Revision of Ms. Gmd-2014-254 N entitled "Improving the ISBACC land surface model simulation of water and carbon fluxes and stocks over the Amazon forest" by E. Joetzjer et al.

We are grateful to both reviewers for their helpful and constructive comments. Please, find hereafter our point-by-point response in italic.

Anonymous Referee #1

This paper details the improvements made to the ISBACC through the introduction of: 1) a new soil water stress function, which alters the modeled photosynthesis and 2) a new autotrophic respiration scheme. The research tests some interesting adaptations to the model to work towards improving the ability to simulations of respiration and the impacts of seasonal variations in water flux in tropical forests; work which is greatly needed. However some work is needed on the manuscripts format and writing style.

Major comments:

1) I find this manuscript confusing to read in places. Informal language is used frequently throughout the manuscript and the order or appropriateness of words often seems wrong.

We had the manuscript read by a native english scientist who corrected parts of it.

2) I find the structure of the manuscript odd. Firstly some of your results are introduced in the methods as justification, rather than a clear explanation of what questions and hypothesis you will test.

We are not sure what part of the manuscript the reviewer is referring to. Is it the introduction to section 3? If yes, we indeed wanted to explain why we chose to modify the model in such a way. We believe that bad model results is indeed a motive for improvement. We did add however a sentence reminding the scientific goal already stated in the introduction:

"Indeed, large biases in the simulated latent heat and respiration fluxes are not acceptable when modelling a region where precipitation recycling is important and where changes in the carbon fluxes could have profound effects on the global climate."

Secondly the text seems disjointed in the results and discussion section; I find myself repeatedly trying to find the explanation for a stated result given and either find no explanation or an explanation given in a separate section. In a combined results and discussion section I would expect results to be stated and followed by an explanation, which is put in the context of relevant literature. Perhaps this section can be restructured and made clearer for the reader.

We chose not to change the structure of section 4, but we did add explicit references to the explanations given elsewhere in the text ("as is shown in section...")

3) I do not understand why ORCHIDEE is used in this manuscript; no clear explanation is made of how the comparison with this model adds to the conclusions of the manuscript.

The idea behind the use of several models was to try to estimate the uncertainty due to the land-surface model choices in the simulations of the Amazon future. For that purpose, we first had to have 2 models that behaved reasonably well, hence the work on improving ISBA. It had already been done with ORCHIDEE. Now, this exercise showed us that in a forced setting, as it is the case here, both models react fairly similarly despite very different hypothesis behind the representation of fundamental processes like photosynthesis, allocation, or phenology.

4) Section 4.1: I would argue that soil texture is not the only issue controlling SWC distribution with depth. SWC with depth will be strongly influenced by your root water uptake too and therefore dependent on: a) differences in root biomass between the model versions, b) differences in the vertical distribution of roots and c) differences in the LAI and gs. I feel that these issues need to be addressed and shown, to test what is truly driving the differences in SWC. This is briefly touched on in the section 4.2, but is not comprehensively dealt with to assess what really drives the changes in SWC distributions between models.

We agree of course that soil texture is not the only factor controlling vertical soil moisture distribution. But in this particular case, we believe it is the most important because a) an b) water uptake in ISBA does not depend on root biomass. Roots have a predefined depth and vertical distribution (based on Jackson et al 1996 work) kept constant. So root depth and profile is identical between the simulations and can't play a role in the differences between simulations. Now looking at the bias common to both simulations (too fast wetting of the soil), it could be explained by too low water uptake by roots during these periods. But increasing water uptake by roots during these periods would increase too much evapotranspiration. But we agree that this section is a bit too short. We rewrote it and included the uncertainty in the root water uptake.

5) The key changes caused by the addition of the respiration and the new water stress function are changes to both the LE and H fluxes and the carbon stocks/allocation. There is however little effect on the NEE, Reco and GPP caused by PS+R. Presumably there are many trade-offs between the model formulations that allow such drastic shifts in C allocation and storage, and shifts in H and LE whilst maintaining the same C flux balance. These trade-offs are very important for the interpretation of the results, but not presented clearly or well explained in the manuscript.

We added an explanation of the trade-offs in section 4.3: "So, there is a trade-off in the model between f0 and Ammax, that can be expected from the photosynthesis module. A lower maximum assimilation rate (Ammax) tends to reduce the carbon assimilation (see eq A7 in Calvet and Soussana, 2001). On the other hand, with a higher f0, intracellular CO2 increases (see equation 4), which favors carbon assimilation."

In section 4.4 we didn't want to repeat why Ra and Rh partly compensate each-other, but we tried to improve the clarity of the paragraph that explains it.

Minor comments:

P1299: “The fractions of newly formed assimilates or reserves allocated to these pools are parameterized as a function of soil water content, temperature, light, and soil nitrogen availability” I am sorry I don’t really understand how these parameterisations are made.

We wanted to just give a sketch of ORCHIDEE carbon allocation scheme. But it is indeed too short to be understandable. As it is not essential to our purpose we dropped this sentence.

P1299: L13-15: What is an above ground metabolic and structural pool, is this carbon pool and if so how is this differentiated from an above ground biomass pool?

There was a mistake in the sentence that we corrected and we now specify consistently that we are talking about biomass pools.

P1301: DS is suddenly introduced here, but there is no explanation of this acronym.

Indeed, there was a mistake in the text, D in place Ds (corrected). Note that Ds is also defined in the table 2.

P1306: L 20: I don’t really understand what you mean by “model diversity” and what we can illustrate with it.

Ok, model diversity is exaggerated considering that we only have to models. However it's interesting to compare ISBAcc to Orchidee considering that ISABcc has never been evaluated on tropical forest while Orchidee has been largely evaluated for this PFT. We did reword the manuscript.

P1307:

L3: What wet bias? You reference no figure or give no quantification! Also the lower panel actually shows that soil moisture contents seem to be relatively similar between the models. If you were able to put error bars on the observations to represent spatial heterogeneity (which can be very large) would you expect either model to be outside of these error bars?

We agree. That is why in the text we wrote “the slight wet bias”.

L8: “...allows the model to simulate a relatively wet top-1m horizon as observed.” Again you reference no figure or give no quantitative way for the reader to assess this.

Added (Fig. 3 mid panels).

P1308: L6: Acronym SD used and not explained.

Standard deviation (corrected).

L7: “The CTL runs show a systematic overestimation of H that is strongly reduced in both PS and PS+R versions.” This sentence is a bit repetitive of previous sentences.

We are describing the bias (bottom) part of the figure. But it was not clear so we rewrote the sentence.

L10-15: No explanation is provided as to why LE and H biases vary between sites here.

We added a sentence about the low energy closure of the Manaus measurements

L15-24: What is the explanation for the model improvement?

Explained before "The partitioning of the energy budget is better represented with the simulation using $A_{m,max}$ and f_0 parameters derived from the in situ observations (PS version, Table 3)."

L19: Again you reference no figure and give no quantification to back up your result.

Indeed, Figure 5, reference added. The quantification is given in the Taylor diagram.

L21-24: I disagree with this statement. I would say that the likelihood of two models being wrong at the same location is not small, particularly in tropical forests where, as your paper is suggests, there are many mechanism and processes we understand very little about and are nowhere near being able to model.

This is only a suggestion that is partially backed up by the weaker energy closure at the Manaus and Jaru sites. We added a sentence about this.

P1309:
L1: what is the "CTL experiment"?

Sorry, CTL simulation, corrected.

L17: I don't understand what you mean by "The scores". L18: RSD acronym used without explanation.

Statistical scores, RSD = relative standard deviation. Corrected in the manuscript.

L25: Can you give an explanation as to why you think the biases vary from site to site. You say about eddyflux errors below but do not actually directly say that it is linked to this issue. Also are there any modelling biases which you would expect to vary by site?

We rewrote this part. The NEE biases vary from site to site mainly because NEE is a small flux resulting from the difference of 2 large ones (Reco - GPP). It is important to evaluate modeled NEE to observations since it is the flux actually measured but not much can be said from small differences between sites.

P1310:

L3: Data can, and should be filtered for U*. Did you do this or test this affect?

The fluxtower data used in this analysis are indeed filtered depending on u. The references of the data are given in table 1. As many papers advice to take the u* filtered data we didn't test the effect of the filtration.*

L12: You state 330 TC ha-1 but you do not discuss any of the errors on this and similar numbers in Fig 8. There are no errors on any of your observed values. Observed data should not be used extensively without considering its error, particularly as some of the errors stated in Malhi et al., 2009 are substantial.

You are right. We added rough error estimates (from Malhi et al paper) in the text.

P1311: L3-4: Can you reference a figure or quantify the underestimation of Rh and the Rh C stock?

Yes Fig. 9 (reference added in the text)

P1312: L13-16: This has not been discussed in the text directly and it seems odd to introduce this in the conclusion.

Ok, to make the manuscript easier to read we moved this paragraph at the end of the results and reworded the conclusion.

Figure 3: I do not see any blue PS line on here. I am assuming it is under the PS+R line. If so maybe show a single line and note that the responses are identical. Also on this and all other figures can you add a letter in the panels (e.g. a., b.) as in the text it would be much easier to follow. I also do not really understand how you did your re-scaling or what the bottom panel is showing. Is it averaged or max SWC over a 10m-2 area or a 2m-2 area?

Ok, we modified the figure and the legend (where we better explained the rescaling : "We linearly rescaled the soil moisture content of the 10 m pit (Bruno et al., 2006) to the values of the 2 m one (da Rocha et al., 2004) by multiplying the 10m SWC by the ratio of field capacities between the 2m and the 10m pit)"

Figure 4: RN acronym used without explanation in legend, but then R Net used in the Figure.

Indeed, corrected.

Figure 5: Standard deviation should still have units associated with it and if not there should be an indication that the data has been normalised in the Figure legend. Also I would suggest labeling what the lines in the Taylor plots indicate, as there is a lot of information in these plots. Finally you do not mention the period over which the comparison is made in the figure.

Ok, see legend.

Figure 6: Please put the whole figure legend in and not a reference to another figure.

Ok.

Anonymous Referee #2

This paper presents results based on improvements made to various aspects of the ISBAcc model. The developments are grouped into improvements relating to photosynthesis and soil water stress (PS) and additional improvements to the respiration of various biomass pools (PS+R). With the PS +R version, the biases in latent and sensible heat flux, GPP, and ecosystem respiration are generally reduced. The model performs comparably to the ORCHIDEE model. To me the most notable improvements are the carbon stocks and division between heterotrophic and autotrophic respiration. The paper shows important progress in simulating fluxes in the Amazon and should be published in GMD, I do have some suggestions for improving the manuscript and advise minor revisions.

The paper is well organized and the results are clearly explained. However, there should be more attention to the uncertainty in the observations and a link between the model results and site-specific processes. The authors state several times that there is large uncertainty in the flux measurements and particularly in the partitioning between GPP and RECO. I am in favor of using these measurements for model evaluation but I think uncertainty bars should be added to the figures. If not, a more quantitative discussion of the results within the context of observational uncertainty would help the reader judge the improvements in the model.

My second major comment is the link between the site-specific results and processes at the sites. It is stated a few times that the differences in model biases between sites are possibly due to errors in the forcing data or observations (for ex. Around Line 20 on page 1308). I'm not sure I agree with this statement, especially based solely on the fact that the two models perform similarly. It seems more likely that these differences are because of fundamental processes that are different between the sites that neither model captures. I have some specific examples and suggestions below.

We thank the reviewer for the suggestion to look at the overall energy closure of each site. This helped us quantify what we suspected from modelling only. We also added a few references to link our results to the specificities of each site. Our initial goal however was to evaluate a land surface model over that is used within a global climate model. Having this goal in mind we needed to test if the model could reproduce fluxes and pools over a fairly wide range of sites within the Amazon forest (and the CTL version could not). Our goal was not to use the model to gain insight in the functioning of each site, which, we agree is probably more interesting scientifically.

Lastly, I recommend some proofreading and editing. There are several instances where the wording is not precise or sentences are unclear.

Specific comments

Page 1297 (Section 2.1): What method was used to calculate GPP and RECO from the NEE? And why was this not done for the GFG site?

The fluxnet data are available online, the partition between GPP and RECO is done following Reichstein 2005. For the GFG site the PI (Damien Bonal) did not release the GPP data.

Page 1298 (end of section 2.1): This would be a good place to mention potential problems with energy closure at the Fluxnet sites. How well do these observations close the budget (ie what is LE +H/Rn for each site)?

We did calculate the overall energy closure according to Wilson 2002. Thanks for the suggestion because it shows a priori that the Manaus and, to a lesser extent, Jaru sites have observation data that are less coherent with each other than the other sites.

Page 1299 (Section 2.2): I think there are some typos around Lines 20-23. Are the 3 carbon pools active, passive, and slow? Or is the 3rd pool both slow and passive? Also do you mean 'labile' instead of 'liability'? This should be reworded.

Indeed, corrected.

Page 1300 (Section 3): The text in this section explains the different model versions in a clear way, but I found it difficult to follow Table 3. For one thing it's not clear what the "tolerant" and "linear" experiments are. Also the columns seem to switch halfway down, from depicting differences between CTL and PS to differences between gm and f0 calculations. Other suggestions: The sentence about Table 2 should be earlier in the paragraph, and it should be explicitly stated that Table 3 refers to the parameters used in the PS experiments.

Indeed, the table was modified during the edition process. We changed it to improve clarity.

Page 1302 (Section 3.2, Near Line 15): Perhaps to help with Table 3, the equations for soil water impact on f0 and gm can be moved here. Also it's not clear how these equations changed between the CTL and PS experiments.

We prefer to keep them in the Table (although reorganized) because this paper does not focus on the water stress functions (see Joetzjer et al. 2014). And there is very little moisture stress in these 5 sites during the time frame studied.

Page 1303 (Section 3.3) In the description of B4 pool: is this pool for the sapwood of the roots? If so there needs to be an apostrophe after roots to clarify (roots' sapwood).

Indeed, modified (also modified for B6 root's heartwood)

Page 1304, Equation 7: Double check this equation. Should the LAI term be part of the exponent?

Indeed, thanks, modified.

Page 1305 (Section 3.3.2): Typo in heading name (change 'trunc' to 'trunk').

Indeed (modified).

Also what are the values for βwood, E0 (mention that the values are given in Table 4)?

The values are now in the text.

Is the β the same in Equations 11 and 13?

Yes, also added in the text.

Page 1306 (Section 3.3.3): The SLA is mentioned here but it's never stated where in the model the SLA is used.

Indeed, we did add a paragraph (section 3.3.4) to be more clear Note that we chose not to describe in detail the parametrization of the SLA. We only summarize the concept and give appropriate references.

Page 1307, Line 5: What do you mean by "successfully evaluated", can you briefly state the results of that evaluation (ie: are the model results similar between K67, Caxiuana, and K83)?

We add a reference to the figure from (Joetzjer et al. 2014) showing the evaluation of the Soil Water Content at K67 and Caxiuana in the text. We showed in this study that monthly Soil Water Content in the top 3m was correctly simulated by the modified version of ISBAcc (with the linear WSF) in terms of quantity (biais < 5%) and seasonal variability (correlations =0,9 at Caxiuana, but only 0,6 at K67 because the meteorological forcing did not cover the whole period (see details in the paper).

Page 1308, First paragraph: Taylor diagrams are now commonly used but it still would be helpful to orientate the reader as you begin to discuss Figure 5 (for example: Lines of constant correlation extend from the origin, and standard deviation relative to the observations is denoted by the blue radial lines, etc). This is especially true because you have several different statistics displayed in the figure, and sometimes it is unclear whether you are referring to the Taylor diagram or to the bias plots.

Ok added (text and legend)

Page 1308, Second Paragraph: The improvements in the model appear to be substantial. Do these occur year-round or are the improvements focused during a particular season?

Improvements are not season dependant. It is said in the 1st paragraph of section 4.2 but we did a reminder in the last paragraph of the section.

Also, why is the bias still so high at Jaru? Here is one place where a link between the model results and characteristics of the sites would be helpful. Also can you be more specific with your final sentence in the paragraph – is there evidence in the literature for which processes might be missing? See for example Baker et al. 2013: Surface ecophysiological behavoir across vegetation and moisture gradients in tropical South America, Agricultural and Forest Meteorology (attached as a supplement), and da Rocha et al. 2009, (already cited in this study).

Thank you for the papers. As pointed out by Baker et al, 2013, Jaru is wetter than the Santarem sites, with a longer dry season and a rather different radiation seasonal cycle, being located at 10 degree S. But in our case, we also have the Guyaflux site, that is also very different. It is even wetter *than Jaru, with an equaly pronounced seasonal cycle and is located at less than 20 km from the ocean. So we could expect a large bias too, which is not the case.*

So we agree with the reviewer that site specificities are important and we added a few lines and references in section 4.2 but in our case we believe that the coherence of the dataset itself plays an important role.

Note also that a high bias in H expressed as a percentage is not that important since H is relatively low in the Amazon. So a 100% bias on H at Jaru is actually not that great amount of energy.

Page 1309, Line 1: I would not say the GPP is correctly simulated by the CTL experiment, although the annual magnitude appears to be roughly correct.

We added "annual magnitude" in the text.

Page 1309, Line 12: What is meant by "the model behaves as expected"?

It means that when the soil moisture is not limiting and the radiation increase, the model simulates an increase of GPP. The seasonal cycle of GPP is actually very similar to what Baker et al simulate at this site with SiB (fig4).

Page 1309, Lines 13-15: This implies that the latent heat flux from the model is mostly due to flux through the stomata, while the observed LE has other important sources. Is the modelled LE mostly from transpiration? What are the other sources of LE in ISBACC ?

In ISBAcc the evapotranspiration term is the sum of the evaporation from the soil, the evapotranspiration from the water intercepted in the canopy and the transpiration from the plants. Over the Amazon, the transpiration represents about 70% of the evapotranspiration.

Page 1309, Line 26: What are the "observation uncertainties"? The discussion that follows regarding the measurement uncertainty is useful but also highlights why it would be good to quantitatively include these uncertainties in the analysis.

Page 1311: Is there mortality in the model?

The turnovers rates can be assimilated to a background mortality in ISBAcc.

Table 4: Is 1/SLA constant in both the CTL and PS+R experiment? Also what is the Ts and Tp in the CTL column referring to?

As now explain in the text (see paragraph SLA) in the CTL version the SLA was calculated. As indicated in the table Ts and Tp are respectively the surface and the soil temperature

Figure 2. The figure legend refers to Calvet et al. 1999, but the corresponding text at the bottom of

Page 1300 refers to Calvet et al. 1998. Which is correct?

Indeed, the correct reference is Calvet et al. 1998. We modified the figure.

Figures 4 and 6: The display of diurnal cycles for each month is very useful, but for the seasonal cycle it might be easier to judge the model if the 3 years of data are averaged together.

We agree, but we also wanted to show inter-annual variability, and having 3 graphs would have been cumbersome. We preferred to show the seasonal change in diurnal cycle (left panels) and the interannual variability of the seasonal cycle (right panel).

You could also indicate the standard deviation of observed fluxes to give some estimate of the uncertainty in the measurements.

As standard deviation is highly dependant of the time step considered we chose to don't add them. Indeed, if it's calculated with hourly data the diurnal cycle hide the errors, over hourly data they show the atmospheric conditions (e.g. rain/no rain) while over monthly data, we see the seasonal cycle and the interannual variability.

Also, can you denote on the figure which months are the dry season?

Ok, added.