**Response Letter**

Dear Editor and reviewers,

Thank you for your precious and constructive comments and changes for improving our manuscript. As you can see from changes in the “Comparison\_of\_the\_MainText.docx”, “Comparison\_of\_the\_Figures.docx”, and the responses highlighted in blue below, we have carefully revised the main text (CLM\_PiTs\_v9\_09102015\_Biogeosciences.doxc) and the figures (CLM\_PiTS\_v9\_figs\_09102015\_Biogeosciences.docx). We do appreciate your wonderful ideas and insights, which made us rethink the whole picture and improve the quality of our paper.

Reviewer #1 (Comments to Author):

In this paper, the authors attempt to reproduce observations of canopy behavior during a shading experiment following introduction of a pulse of high-concentration 13CO2 and subsequent shading treatments. The results are that CLM performs better in pretreatment conditions when tuned to site-specific values, and CLM is unable to capture the observed track of the 13CO2 pulse through the canopy and roots. The finding that tuning improves model performance at a site is hardly new or novel. This has been done many times before (e.g. Collello et al., 1998; Prihodko et al., 2008; Rosolem et al., 2012, and many more).

Response: As emphasized in those references the reviewer listed, we agree that improving the model parameters is one of the most important and effective ways for model development. But as you can see, our effort on optimizing CLM parameters for the pre-treatment period is not the only story for this study. We found that even with the updated model parameters there were still some differences between the simulated and observed variables especially during the treatment period. So, by isolating the impact of model parameters, we addressed that those errors should be mainly attributed to structural model uncertainties, which can illuminate further model developments. We have modified the text to better explain the purpose of parameter optimization.

The following of 13C through the system is the more interesting component of the paper, and, in my opinion, a lost opportunity. Instead of describing what the observations mean with respect to the behavior of the natural canopy, the authors simply gave a clinical description of how the model differed from the observations. “We ran a model, here’s what happened.” Come on. I can think of a conceptual model whereby the 13C pulse is first taken up by the leaves, then takes time to work through the system. You can see some of this in Figure 6, particularly 6(a) and 6(d), but the behavior in the phloem and bulk roots is more subtle and complex. But a description of processes and mechanisms at work in the real canopy are never addressed, and the reason for model departure from observations is glossed over, the authors merely saying that the allocation scheme “needs attention” and a more labile storage pool should be added. Isn’t this the time to do it? I would be very interested to see a paper that demonstrates the mismatch between modeled and observed 13C, posits some reasons for the mis- match, addresses them, and runs the model again. That would be a very interesting paper.

Response: This is a nice point. We wish we could do all the model evaluation, development and calibration in one paper. However, we think this manuscript adds value to the community by introducing the capability to perform a detailed model-experiment comparison using a global land-surface model. This study adds new capability to CLM in particular: (1) the explicit use of the 13CO2 observational information by reading new stream file rather than assuming fraction of CO2 as in the default CLM version, (2) improvement of the site-level simulation structure and leverage of the capability for single-point implementation of the Point Version of CLM (PTCLM), (3) calibration of the selected model parameters for the pretreatment period with parameterization optimization techniques, and (4) accomplishment of the functional test capability for the photosynthesis subroutines and multi-scale evaluation using both leaf and canopy scale data. But as you can see from the title and content of this paper, we have mainly focused on the evaluations (not just the carbon isotopes but other important carbon and hydrological variables) of the CLM model with the unique PiTS manipulations for the first phase of our CLM-PiTS work. To achieve this purpose, we actually have conducted a lot of novel work. We think those efforts should be unique at least for the land-surface modeling community, and worth reporting as a separate paper. We have improved the discussion section to place our results more in the context of relevant canopy processes. A successful LSM model development effort will require that we understand these processes across broad temporal and spatial scales. A forthcoming manuscript will use not only the short-term datasets of this PiTS 1 experiment but the subsequent PiTS 3 experiment. Both experiments (and other published work) will be used to develop and test a new C partitioning regime in CLM – one that allows short term C storage pools that can reflect abrupt shifts in GPP, seasonal changes in ratios of C allocation to different pools, and inclusion of additional C sinks such as fruit or fungal symbionts.

No real discussion was given for why we are interested in simulating carbon isotopes through the ecosystem. Is it simply to gain a better understanding of biogeophysical processes? Could we expect to see better simulation of net carbon flux and/or the Bowen ratio with better understanding of 13C? Are there implications for ecosystem response to changing climate?

Response: Our interest in accurately simulating carbon isotopes through the ecosystem is not simply, or even primarily, to gain better understanding of biogeochemical processes. Rather we are interested in using the isotope tracer to identify areas of improvement specifically in the model representation of carbon allocation, and thus to improve the representation of C storage and CO2 exchange with the atmosphere, with the implications for simulation of climate-carbon feedbacks and future climate when these improvements are incorporated into ESMs. We have added text in the Introduction outlining this interest, the why, of our interest in simulating carbon isotopes in the ecosystem.

My initial inclination is to recommend rejection for this paper, but I think there is an opportunity here. Take out, or at least minimize the sections on parameter tuning. The community has already done this. A more detailed focus on what is going on with the isotopes as they move through the real system is needed, as is discussion of model success/failure in reproducing the observations and what it means. Finally, the authors should hypothesize some ways to modify CLM, and implement them. This would result in some actual hypothesis testing, as opposed to a paper that reads “We ran a model: here’s what happened.” My formal recommendation is acceptance with major revisions.

Response: As suggested, we reduced the parameter optimization related part and placed it in the context of the ultimate goal of evaluating the model structure. We also added some sentences on the importance of isotope study, and presented potential model developments that could reduce the uncertainties and biases of the CLM in isotope simulations.

Specific Comments:

The unit testing was mentioned as being very important, but not described. If the Wang (2014) paper is all the reader needs to know, cite it and move on. If more detailed description is needed, share it with the reader.

Response: Quite as suggested by the reviewer, we have had the section 2.2.3 for this purpose.

Increasing the Ball-Berry slope and intercept parameters to extreme values made little or no difference in the one plot where they were shown (Figure 4c). Obviously, then, this was not the reason for model error. Why not just say that modifying the BB parameters made no difference and move on? Also, after demonstrating that the BB parameters were NOT important, the authors state in the conclusions that they ARE. This is a contradiction.

Response: The mp and bp parameters were proved to be very important for the pretreatment optimization against the observed transpiration but not for the treatment period. So, the lines 498-523 of section 4.2 were designed to discuss the possible reasons behind the failure of the mp and bp optimization during the treatment period. As suggested, we made slight changes on this expression in the Conclusions part.

Figure 6: there is no explanation given for del 13C, the y-axis on all plots. The scale amplitude differs by an order of magnitude between the panels; the reader needs to be told what is going on here. I’m assuming that the standard treatment is used, where the sample 13C/12C ratio is compared to a standard; is it PDB? Not all readers are isotopists, so some description and context would be helpful. The change of the del value from negative to positive might confuse some readers, so more explanation is warranted.

Response: We added a brief description of the calculation of del 13C in the caption of Figure 6.

Where did the carbon isotope treatment come from? I’m familiar with Suits et al. (2005) and van der Veld et al. (2014). Does the CLM methodology follow these or something else?

Response: Please refer to the section 25 of CLM 4.5 tech note for detailed description.

The del of the respiration is extremely dependent on the spinup, and changing del 13C through the industrial era. How was this treated?

Response: The Suess effect is modeled using changing global atmospheric 13C and 12C from 1850-present as input. Model spinup was performed using pre-industrial 13C values and default model parameters for soil carbon turnover. The text has been modified to clarify the simulation procedure.

In section 4.2 the authors say “. . .modeled soil CO2 efflux was too high on the first day of labeling and too small afterwards.” Actually, Figure 5b shows this to be false. In actuality, the del 13C was too high on the first day, and too small afterwards (Figure 6d).

Response: Changed as suggested.

Reviewer #2 (Comments to Author):

General Comments:

This paper combines a site level manipulation experiment (PiTs) of a young loblolly pine plot in Oak Ridge, Tennessee as a tool to test the model structure of CLM4. The authors use observations from two levels of shading to evaluate ecosystem response, and inject highly enriched 13CO2 into the atmosphere to evaluate model allocation of carbon biomass. The authors find that through calibration of a subset of parameters that the biomass pools, respiration and transpiration are captured fairly well. Based upon differences between del13C concentration in biomass pools and del13C in the soil respiration the authors determine that the existing fixed annual allocation methodology within CLM is insufficient to represent the observed patterns in carbon allocation. They recommend the inclusion of a non-structural carbohydrate pool to help delay the allocation of carbon to plant tissue. In addition they recommend a dynamic allocation sub-model that is influenced by environmental conditions rather than the existing fixed annual allocation approach. This reviewer felt it was unclear how the model was initialized from the spinup-phase to the sapling state in 2003. Was a disturbance imposed and/or the biomass characteristics of the sapling simply prescribed into the model? In addition, it was also left unclear which sets and how much of observed data were used in the calibration phase and which were withheld (e.g. the light response data). Although the authors demonstrate that the existing formulation is deficient in simulating allocation timing, it was less clear whether this formulation was insufficient in simulating multi-year to decadal allocation. I would have welcomed more discussion of model skill in terms of the model simulation of biomass variables from 2003-2010 against observations (Figure 3) to help contribute to this discussion. The data is already there in Figure 3, it just needs to be evaluated, and perhaps shown more clearly than in a log plot. Finally, C13 labeling, as demonstrated here, is used as a means to an end in order to evaluate simulated allocation and timing. To that effect, C13, isotopes and isotopic fractionation are not the focus of the paper, however, the author gives the impression that the model itself is a passive placeholder for C13. In reality, the model is imposing its own mechanistic representation of C13/C12 photosynthetic fractionation upon the canopy air and is actively partitioning against the heavier C13 isotope during the photosynthetic process. The degree of fractionation is dependent upon stomatal conductance, assimilation rate and nitrogen limitation, which in turn are dependent upon environmental variables. Although this fractionation mechanism may be relatively minor as compared to the overwhelming signal of the injection of enriched 13CO2, this reviewer would have benefitted from a more careful discussion of the simulated behavior in biomass pool del13C behavior in Figure 6. How much did the modeled fractionation processes contribute to the model-data mismatch both in terms of bias and trend, if at all?

In general, this paper represents a valuable fusion between manipulated site level experiment to test and improve CLM skill, and recommends acceptance with revisions based upon the suggestions mentioned above.

Response: Thank the reviewer for the detailed and positive comments on our work. Detailed responses can be found below.

Specific Comments:

How was the model initalized after the spinup? Was a harvest or planting initiated at 2003 or were the sapling biomass variables prescribed into the model? Only in the results section does it become clear that the simulation was started from near bare ground in 2003 and then run forward. This should be described more carefully in the methods.

Response: All aboveground biomass was removed in 2003 to simulate a harvest, and a planting occurred in 2003 using a prescribed biomass. We added a clearer description of initialization in Section 2.2.1.

Although there were many measurements taken at the site, it was unclear what were actually used for the calibration, how many data points, and at what temporal resolution.

Response: The variables used for pretreatment calibration are described in the Section 2.2.2. We also added the temporal information associated with these observational variables.

Page 6979, Line 19: Should read 13CO2.

Response: Changed as suggested.

Page 6981 Line 26: I would say standard ‘parameter’ CLM version, instead of just standard CLM because this is confusing whether you mean parameter or structure.

Response: Changed to “Compared to the CLM results with standard parameters”.

Page 6985: Lines 3-6: An increase in soil respiration suggests increase in root respiration? How do you not know it isn’t increased soil carbon respiration?

Response: We were mainly addressing the little sensitivity of CLM simulated root and soil carbon respiration to the soil water condition. No magnitude comparison of these two respiration were involved. The observations may also be driven by increases in heterotrophic respiration, and we modified the text to reflect this.

Page 6986: lines 5-10, The fact that the model overestimates the del13C could be a function of allocation turnover time or the fractionation process itself. Did you demonstrate what the pre-treatment fractionation process is. . .? Could have taken pretreatment del13C leaf samples and gotten a baseline for the fractionation process itself. It could be useful to account for the influence of fractionation on model output, and help distinguish issues with the fractionation vs. pool turnover and allocation.

Response: The pre-treatment del 13C measurements were used for model calibration. Parameters controlling both stomatal conductance and allocation were included. The leaf del 13C was modeled quite well (Figure 6a). The model may display a bias in bulk root del 13C (Figure 6c) because of errors in allocation timing or a failure to represent post-photosynthetic fractionation. Further discussion has been added to section 4.1.

Page 6987 Line 11 The leaf level light response data was not used in calibration, but I am not completely clear what was included.

Response: In terms of the calibration variables, please see the response above.

Page 6987 Line 17 You found that the seedling biomass was important for the simulation, so did you use that as a calibration tool? You only show in table 1, the parameters that were being optimized, but if state variables are also optimized this needs to be included in table 1.

Response: A typical seedling biomass (1 g C m-2) was used to initialize the model for each simulation, including those used for calibration. These were fixed values rather than variable parameters. While uncertainty in the initial state may be important in determining calibrated parameters, we did not have suitable information and so used the fixed values. Section 2.2.1 was modified to better describe the initialization process.

Page 6989, line 23: “The simulation does not show any difference in respiration between LS and HS plots prior to the shading”. But that’s the same for all variables, shown in Figures 4,5 and 6. Why would it not be this way? Aren’t these simulated exactly the same way between plots, or did you calibrate individually? It was my impression that the only thing that would make a simulation difference would be imposing different meteorological conditions at the point of shading. Nothing was different before this between the state of the LS and HS plots as far as the model was concerned. This is not the model’s fault, and you would have to optimize or initialize the model based on subplot characteristics to get different pre-treatment simulations.

Response: This is correct. HS and LS plots are identical in the pre-treatment simulations and were not calibrated separately. In the actual experiment, the HS trees had slightly higher biomass and leaf area, leading to differences between HS and LS in pretreatment observations. We did not have quantitative information about the biomass components of each plot to conduct separate calibrations for HS and LS separately. Soil carbon may also vary among the plots and was not measured too. We felt that this section added little to the discussion and removed it.

Page 6990 Lines 12-25: “It is clear from this study that additional work is needed to improve allocations schemes in CLM.” I would tend to agree with this statement if we are concerned with diurnal scale variation (as show in this paper, and even then it doesn’t seems quite so much an allocation problem, but a timing, or staging problem, i.e. assimilated carbon is too quickly allocated to biomass and respired), perhaps seasonal allocation variation (not specifically addressed in this paper). Has this been demonstrated for decadal time scale allocation and growth? Isn’t a better indicator of model skill how well the biomass variables were estimated from start (2003) to treatment (2010) (Figure 3a). The log scale makes it hard to see, but it would appear that the calibrated simulation provides biomass variables consistent with observations, and isn’t this the time scale more relevant for climate, perhaps the simple allocation scheme isn’t bad for longer time frame..?

Response: Good point. However, we do not have annual observations of the biomass components from 2003-2010. We only know these trees grew up from one-year-old seedlings (1 g C m-2) planted in 2003, and the estimated leaf, stem and roots carbon in 2010 based on the loblolly pine trees that have similar age (Fig. 3a). Because the PiTS 1 experiment only lasted for several weeks, strictly speaking, what we’ve seen here for CLM errors in the timing and magnitude of C allocation should be mainly limited to the daily scale or short term. The Conclusions part was slightly changed to clarify this. We would expect to get more insights on the CLM capability in biomass simulation at longer time scale by doing detailed evaluations of the simulated biomass components again latest gridded or multisite observations (Thurner et al., 2014; Liu et al., 2015). Investigations with these datasets are actually underway and details of which will be reported elsewhere.

Page 6992, Line 15. . ...‘[the model]’ is not able to reproduce the observed patterns of allocation as revealed by the 13C labeling experiment.” I thought this was too strong a statement: certainly the labeling experiment demonstrated deficiencies in model timing of the allocation, but ultimately this statement is based upon the growth patterns based upon DBH in Figure 5, which in itself was a proxy for biomass based on allometric assumptions and highly variable based on water status. It does seem like the model overestimates stem growth on this short time scale, but again is this consistent with longer term allocation as shown in Figure 3a? This might make for an interesting comparison if this short time scale overestimation in growth translates to 7 year allocation. The stem biomass seems to be captured quite well.

Response: Please refer to the above response for this comment. In addition, we have changed the expression of that sentence.

Figure 1: Labeling of relative humidity should be consistent. Hr and RH should be same thing. Choose one.

Response: Changed to Hr as suggested.

Figure 3a: Log scale, although convenient in order to apply all biomass simulations on one plot, makes it very difficult to assess the simulated vs. observed biomass variables. Do they agree within error. . ..hard to say.

Response: Based on this suggestion, we modified the Figure 3a and added the observed biomass values in the caption of Figure 3a.

References:

Liu, Y. Y., van Dijk, A. I., de Jeu, R. A., Canadell, J. G., McCabe, M. F., Evans, J. P., & Wang, G. (2015). Recent reversal in loss of global terrestrial biomass. Nature Climate Change.

Thurner, M., Beer, C., Santoro, M., Carvalhais, N., Wutzler, T., Schepaschenko, D., Shvidenko, A., Kompter, E., Ahrens, B., Levick, S. R. and Schmullius, C. (2014), Carbon stock and density of northern boreal and temperate forests. Global Ecology and Biogeography, 23: 297–310. doi: 10.1111/geb.12125

Reviewer #3 (Comments to Author):

General comments

Mao and co-authors present an interesting study evaluating photosynthesis and C allocation parameters in CLM4.0 with data from a 13CO2 labeling experiment in a young Loblolly Pine plantation. Although their efforts to evaluate process representation based on short-term experiments are novel and interesting, I’m surprised there’s less introspection on the implications of their findings. It’s not surprising that adjusting parameter values produces better results but as the author’s stress CLM is a global model used for climate change projections. As such what are the implications of tuning up model parameters for Loblolly Pine trees? Should the Ball-Berry parameters in CLM be changed for all plant functional types based on these findings? Would the modified parameters fit within observations constraints of data in databases like TRY or Glop- net? Are there larger structural uncertainties or biases in CLM that this study exposes? If so, how can they be corrected? As presented this work illustrates how models can be tuned with data, but misses a potential opportunity to draw broader conclusions or gain much insight. Like any good study, this paper raises more questions than it provides answers, but the answers provided here are not very compelling. The authors very clearly state their objective and focus at the end of section 2.2. This is a paper that more narrowly focuses on building tools and capabilities in PTCLM to facilitate model-data comparisons from experimental manipulations and site-level observations. This is an important, valuable contribution. The aim here isn’t necessarily to evaluate and improve CLM, but to build and document valuable tools that facilitate site-level comparisons. Perhaps one path forward would be to more narrowly cast the paper? As presented the paper seems to communicate that the authors were able to tune some parameters that modestly improved growth and photosynthesis parameterization for a particular tree species (Fig. 3), but neglected the heavy lifting of improving HUGE biases in soil physical & hydrology, plant C allocation, and photosynthesis/stomatal conductance (Figs 4-6). The authors adequately highlight these shortcomings and suggest solutions, although it doesn’t really appear they’re interested in addressing them in the future. One reading of this paper would conclude that by applying a novel (but incomplete) experimental design allows for fine tuning of parameters for a particular plant species that aren’t really that bad in the global parameterization of CLM and highlights some huge biases and structural issues with a model that the authors can’t address, or aren’t really interested in fixing.

Response: Thank the reviewer for the detailed and positive comments on our work. As suggested, we reduced the parameter optimization related part and placed it in the context of the ultimate goal of evaluating the model structure. Since some of the concerns raised here are similar to those from the other reviewers, we merged them and made corresponding changes as can be seen in the revised version. Other detailed responses can be found below.

Specific comments

Why shading? What does this kind of manipulation tell you about model response to disturbance or environmental change? The 13CO2 seems novel and valuable but the value of shading seems less clear, especially w/o a control during the experimental period of focus unless it’s to illustrate biases in low light (nighttime?) stomatal conductance.

Response: The shading treatment was designed to abruptly reduce GPP, which will affect the internal plant C flux and C partitioning to various sinks. It was used to create a strong disturbance in observable C partitioning, which could then be used to explore model sensitivity. Measured parameters such as stem C and sap flux displayed large shifts over the three-week treatment, and here we indicate poor model simulation of this abrupt disturbance.

It’s confusing understand what features of different version of CLM were used in this study. For example, CLM4.5 (Oleson et al. 2013) documents the 13C features and PT-CLM configuration used here? However, the photosynthesis, C allocation, and hydrology came from CLM4.0? It seems ironic that the authors would note the significant efforts to improve the CLM (P 6974 L 1-6), yet used a version of the model that does not reflect those changes (CLM4.0) in particular changes to canopy photosynthesis (Bonan et al. 2011, 2012) and known issues with soil hydrology in the models (e.g., Swenson & Lawrence 2014). Can this be somewhat clarified for readers not familiar with different versions and configurations of the model?

Response: CLM 4.0 was used for this study. Oleson et al. (2013) is referenced solely for the PTCLM configuration, which does not affect the model structure. Modifications of soil scheme in Swenson & Lawrence 2014 were mainly focused on the improvement of CLM over semiarid regions. Although we noted biases in the soil hydrology, there is little influence on carbon fluxes and allocation patterns since this site is not water-limited. The canopy photosynthesis in CLM4.0 does a reasonably good job against our evaluation metrics, including leaf-level light response data, and the CLM 4.5 will be evaluated in future manuscripts. As suggested, some further discussion and clarification on model version were added.

The experimental design that this study is trying to replicate seems surprising to me. As I understand it there are measurements for pre-treatment, two shading experiments, and post-treatment, but during the shading experiment there’s no untreated control group?! Repeating the experiment is well outside the scope of this manuscript, but a more compelling study would have been to parameterize the model for the pretreatment period and then see if it can even replicate control conditions during and after the experimental shading. This is especially troubling since the 13CO2 pulse came at the start of the shading experiment, and for which there are no control (i.e. unshaded) data?

Response: This is a good point and something we discussed. We considered using a shade treatment and an un-shaded control treatment, but were worried about the potential differences in environmental conditions caused by the shade cloth, independent of changes in light (i.e., relative humidity, wind). Thus we decided to use a heavy shade as the treatment, and the light shade as the control. We wanted to make sure all trees received the same initial CO2 signal, thus we applied the label immediately prior to shading, then could use the two shade levels to impact the partitioning of that 13CO2 label. We have since performed a new study, with isolated individual trees, using shaded or un-shaded controls as suggested by the reviewer, but also included comprehensive measurements of differences in environmental conditions between each tree (leaf and air T, relative humidity, above and below the canopy, and estimates of long-wave radiation based on sky or shade cloth temperatures. Both experiments (and other published work) will be used to develop and test a new C partitioning regime in CLM – one that allows short term C storage pools that can reflect abrupt shifts in GPP, seasonal changes in ratios of C allocation to different pools, and inclusion of additional C sinks such as fruit or fungal symbionts.

P 6980 L 10, I appreciate the honesty about the approach, but why was this approach chosen, instead of one that provides estimates of parameter uncertainty?

Response: This is a computational limitation, other methods that estimate parametric uncertainty (e.g. MCMC) would require a larger number of simulations and are harder to parallelize. Uncertainty quantification is an active area of research in land surface modeling. Because parametric uncertainty is not a focus of this study, this sentence was removed.

I appreciate the work that went into optimizing parameters for a young loblolly pine stand, but it’s hard to see and never discussed how far off the standard configuration of the model was from observations (Fig. 3a)? Are there error estimates on the observations? If so, can they be displayed (or are they already)?

Response: We modified the Figure 3a and added the observed biomass values in the caption of Figure 3a.

I understand the focus of this paper is on biogeochemistry and C fluxes but how well does PTCLM handle water and energy fluxes? How do changes in parameter values influence transpiration rates, latent and sensible heat fluxes? Has anyone looked at data from ‘nearby Walker Branch and Chestnut Ridge eddy covariance sites’, or other flux towers in loblolly pine plantations? At the very least there should be some introspection on how the suggested parameter changes influence other parts of the model, not just C fluxes.

Response: In addition to the C fluxes, we also evaluated the CLM simulated soil water content and transpiration against the PiTS observations. Although nearby, the Walker Branch and Chestnut Ridge sites are different vegetation types and ages, and it wouldn’t be appropriate to include them in this evaluation.

Looking at soil respiration (Fig. 5) seems like a spurious analysis given the data presented. In models like this rates of respiration are largely determined by soil C pool size but as these data are never presented, it’s hard to assess if the model is producing plausible results (as implied) with realistic initial states? Moreover trying to justify potential experimental differences in soil respiration seems speculative and distracts from the focus of this paper (photosynthesis and plant C allocation).

Response: Here we focus mainly on short-term variations in soil respiration, which are impacted by environmental conditions as well as the initial states. Although we don’t have enough information to judge model performance in simulating soil C pools, the short-term variability is clearly not being simulated correctly. We think this is relevant in that this could be reflecting structural limitations in the model including those related to root allocation and respiration.

The authors seem to have begun this study acknowledging the C allocation scheme is and isotopic fractionation in CLM is very simple (section 2.2). Thus, findings that it does poorly against observations hardly seem noteworthy. Even still, I’m surprised the authors don’t go deeper in their discussion of these results (P 6990 L 15-27). Section 4.3 reads it was taken from the DOE-MODEX website and/or a proposal the authors just submitted. No one is arguing about the value of bringing models and experiments closer together, however, this section is completely void of specific modeling needs that the results here highlight. It seems like the authors learned they need to collect some data more effectively, but it’s not clear how the results from PiTS-1 inform the model development directions (and measurements) that should be prioritized. This seems like an excellent opportunity to reflect on specific knowledge gained from both model and experimental work and how that insight will be applied moving forward.

Response: As suggested, we significantly modified the Discussion part to make our point clearer.

Technical Corrections

The sentence on P 6980 L 6-7 seems incomplete. A genetic algorithm that does what? [“To reduce this possibility, we used a genetic algorithm (Runarsson and Yao, 2000)”]. Please provide a bit of background for readers not familiar with this approach.

Response: We edited this sentence for clarity.

P 6982 L 12-21. I don’t completely follow this discussion. Did the authors change how CLM estimates limitations on carboxylation rates? If so, where is this described? If not, this text seems out of placed and either should be moved to the discussion or removed (with preference for the later).

Response: As suggested, we deleted several sentences beginning and after “The persistently….”, and deleted the references associated with these sentences.

P 6986 L 23 Can statistics be provided for the statement that parameterized results are better than the standard configuration? This statement seems true, but the stan- dard results don’t look that bad (for using a generic evergreen pft parameterization to simulate growth of young loblolly pine stand).

Response: From our modified Figure 3a, the evident improvements with the optimized parameters can be seen clearly.

Also in section 4.1 it seems like another important parameter change is the fraction of NPP that builds stems in the optimized parameterization. This also makes sense since loblolly packs on a large amount of wood for a relatively low LAI, making it a valuable timber / plantation tree. This is never discussed.

Response: Thanks for pointing this out. We have added some text to the discussion.

Figures are small, complicated, and hard to read. In my experience this gets even worse when papers are formatted to journal styles. Can text in figures be made larger, and cluttering information (e.g. formula be put into the caption). Some of the color choices for lines are either nauseating or unreadable (green and cyan), and insets in Fig. 6 are too tiny to be useful. Please take care to generate illustrative figures that help communicate & clarify the story being told here. Where possible figures should communicate observational uncertainty. It’s shown on Fig 4b, but not elsewhere.

Response: As the reviewer suggested, we have made significant changes of most figures in the updated version.

In Fig. 6, why do simulated 13C concentrations seem to spike before the labeling experiment actually happened (day 0?). This an error in how the figure is drawn, how the label was applied in the simulations, or a misunderstanding on my part?

Response: The labelling happened at midday on day 0. Because CLM allocates carbon to structural carbon pools and autotrophic respiration on every half-hour time step, the spike results primarily from root growth respiration during and immediately following the labeling on day 0.

Throughout there seem to be formatting errors with the subscripts on 13CO2.

Response: Changed as suggested.