Max–Planck–Institut für Biogeochemie
Max Planck Institute for Biogeochemistry

Max-Planck-Institut
für Biogeochemie

MPI für Biogeochemie · Postfach 10 01 64 · 07745 Jena, Germany

Editors
Biogeosciences

Verónika Ceballos-Núñez
Tel.: +49-(0)3641-57-6143
vceball@bgc-jena.mpg.de

2nd January 2018

Dear Dr. Akihiko Ito

We are very grateful for the time that you have dedicated to reviewing our manuscript, and for giving us the opportunity to submit a revised version. This new version is the result of several changes made with the recommendations and comments of the reviewers; particularly the 'Methods' section was rewritten in order to clarify the motivation behind this work and the approach that we used. Overall, we made clear that with this approach we are introducing an alternative way to diagnose models. We understand that some readers may be disappointed by the fact that the outcome of this work was not an improved model, but we hope that this can be overlooked because we provide a powerful tool that the researcher can use to make a more informed decision on which model performs better, with respect to a given question.

In the text below we quote the referee's comments in italics and provide our response below in blue:

Reviewer 1

General comments : The research article Ages and transit times as important diagnostics of model performance for predicting carbon dynamics in terrestrial vegetation models promotes using the distribution of C ages and C transit times of different tree organs for improving performance of vegetation model. In order to do so, the authors tested three different carbon allocation schemes into a simple vegetation model with the aim to discriminate the three allocation models in terms of C stock, C flux, radiocarbon, C ages and C transit times distribution. The paper concludes that C ages and C transit times distribution can indeed be used to evaluate the skill of the allocation schemes. Furthermore, the authors encourage the scientific community to use their approach for future model comparison or validation. In my opinion the authors did a good job in developing a powerful method to help reducing uncertainty in model output as well as facilitating model development.

We appreciate the motivating comments from Referee No. 1. This referee's impression from the manuscript captures what we intended to communicate with it.

I really appreciate that the R scripts are clear and easy to use by someone interested in applying this method to their own model. Im convinced that the ideas presented in this paper our of interest to the readership of biogeosciences, but the presentation itself needs to be improved. The methods and result sections lack essential information. Before resubmitting the manuscript, a senior researcher should carefully edit these sections such that they meet the minimal requirements for publication.

We agree that there were some explanations missing and that there were certain points of the methods section that could be improved. Thus, we made major modifications to this section, which can be found in the suplementary material of this response.

Furthermore, the following concerns should be addressed in the manuscript:
1) The study relies on theoretical simulations from an unrealistic vegetation model to match the requirements of the mathematical calculation of ages and transit time. I understand and agree with the need for this approach. This caveat should, however, be addressed in the discussion and the conclusion where it was argued that the method can be useful for evaluating more complex vegetation models. In my opinion this suggestion is overly simple and this statement is not needed for the study. I would recommend the authors to stick to the more theoretical conclusions. Based on my understanding of complex vegetation models, the numerical and conceptual work required to implement the proposed method into a vegetation model with full complexity is too large to be downplayed. If the authors disagree, I would expect at least a paragraph dedicated to how this functionality could be integrated into a model with phenology, environmental changes, forest management i.e. forest structure and genetics. This would really enhance the impact of the manuscript as it may convince modellers to apply your method in their models.

The explanation on how this approach could be applied to more complex models was shortly examined on page 17 line 17. In order to calculate mean ages and transit times of non-autonomous models (inputs and process rates change over time), we would need a different set of equations. Rassmusen et al. (2016) have provided equations for the mean age and mean transit time for non-autonomous models, which require knowledge on the entire history of inputs and cycling rates for the duration of the simulation. This information is not available for Harvard forest, and for this reason we did not include these computations in our analysis. However, for complex models in which process rates vary over time due to phenology, environmental change, forest management, etc., one would obtain nevertheless age and transit time distributions. These distributions would be time-dependent, i.e. one would obtain a different distribution for each time-step. Since this manuscript is an introduction to the main concept of age and transit time distributions in vegetation models, we believe it is better to keep it simple and show these distributions for the steady-state case only. This would allow the reader to grasp the main concept, and if needed, apply to the more general case in which process rates change over time.

2) The method section is concise, which in general I like, but the current description of the methods is too concise making it hard to follow. Each time the authors refer to a method described by another study they referring to it without a explanation or at least the reason why they choose this one. As a general guideline readers should be able to understand the method section without having to consult any other papers. What justifies the use of a two-step instead of one step optimization? Why do you need Bayesian optimization?

As mentioned before, we admit that the Methods section needed some improvements and we have worked on that for the new version of the manuscript. The Bayesian optimization gave us the possibility of exploring the parameter space, because it searches for alternative parameter sets that can result in a good fit of the model to the data. In order to have a good starting point for this search, we used the output of the classical optimization as an initial constraint. The parameter sets that we obtained were used in the uncertainty analysis, as explained in the P7L10.

Why do you test co-linearity between parameters (answered too late in the discussion)

We agree that the explanation of why we test collinearity should be included in the methods, specially because it is not a common practice in the field. The new version of the manuscript includes a rationale for the use of the collinearity analysis early in the Methods section.

3) Since the authors use a lot of R packages to manage complex statistical methods, I guess that they are familiar with statistics and coding. Surprisingly, simple tests such *as the t-test, anova, or metrisc like Root Mean Squared Error are not used to highlight their finding. The conclusions are not backed-up by any statistics, which is mandatory for publication in a journal like Biogeosciences.*

This is an interesting point because it is customary to support empirical findings with statistics. However, we are dealing with a completely different case here. In empirical studies one has data that is assumed to be random draws from an unknown distribution. Here, we have known distributions and no samples are drawn from them, only calculated mean values from the distributions. Unfortunately, in this case the usual t-test or even the Kolmogorov-Smirnov test for comparisons of samples from different distributions are not appropriate because we are not dealing with samples whose distribution is unknown; we are actually within the rare case in which the distribution is known and we not only have it's equation, but also the parameter values. We have to admit that we are not aware of any statistical test designed to compare two or more distributions without having to draw samples from them. But even if there would be a test, the results would be trivial. We know that the distributions are different because we have different parameter values from all of them. Therefore, we would always reject the null hypothesis that the distributions are different. Since this would be a trivial result, we refrain from adding any statistical comparison to our analysis.

Detailed comments on the tables and figures:
Table 1: the column called final is not clear! Why this column differs from Best2 column ? I guess it is because of the constraint optimization on C stock but it is not described in the caption. Do not show the median and the quantiles of Best2 if you will not use these values in the next step.

We agree that the names that we used for the parameter sets were confusing. We now describe them in a better way in table 1 as well as in the Results section. 'Final' is the parameter set that was most frequently chosen by the Bayesian optimization method and was used for all of the simulations, unless otherwise noted. Furthermore, the above mentioned 'quantiles' are the quantiles of the distribution of the values of each parameter, as a result of the parameter space exploration performed with the Bayesian optimization.

Table 2:I am not sure this one is necessary ...

We moved it to the appendix

Figure 3: Two figures instead of one can be better to avoid the compressing of the leaf C stock.

This is a good point, we followed the recommendation of the reviewer here.

Figure 5,7: Put some statistics to highlight your results. You argue that mean ages and transit time are different across allocation model. Support it with a probability! Figure 8: Again show that mean ages are significantly different.

Please see our response above regarding the applicability of statistical tests in this case.

Reviewer 2

Despite an encouraging title and a promising abstract, I find the study led by Ceballos-Núñez & al. quite disappointing and I doubt that the modelling approach proposed here can be used by a larger community to constrain carbon dynamics in terrestrial vegetation model.

It is unfortunate that the reviewer was disappointed with this manuscript, but we are convinced that this issue can be easily resolved, because there seems to be certain misunderstandings that will be unveiled in the following points.

The "vegetation" model described in this study is in fact a very simple box model, calcu-

lating fluxes between different carbon storage compartments, as well as the carbon stock of each compartment, comparing three model structures (i.e. increasing the number of carbon pools). The model was forced by a constant input of carbon (GPP=1400 gC m-2 year-1) and run on a yearly time step, with no change in environmental forcing (climate, CO2,etc.). The results are shown during the transient spin-up (e.g. fig4) or at steady state (most figures). In my opinion, this approach (i.e. yearly time step, constant GPP, no external forcing) is absolutely not appropriate for "predicting carbon dynamic" as claimed in the title. The actual dynamic of the carbon cycle, i.e. the increasing terrestrial carbon sink, is happening because of transient changes in external environmental conditions affecting the terrestrial carbon dynamics. Therefore it is impossible to draw any conclusion from this study, with respect to the actual dynamic of the system.

It is important to clarify that autonomous systems, as those modeled here, are still dynamic, and an useful tool to assess processes that occur within the vegetation, which in this case is the distribution of carbon among different compartments. If we were interested in predicting the effect of a specific disturbance such as time-varying atmospheric CO2 or temperatures, we could still predict time-varying ages and transit times distributions. However, this is not the objective of our study; we are rather interested in presenting **the concept** of ages and transit times as useful diagnostics of model performance and as a tool to explain mixed ages of non-structural carbohydrates previously reported in field studies. For the case of transient simulations, which the reviewer advocates here, formulas do exist to calculate mean ages and transit-times (see Rasmussen et al. 2016), and if we would have knowledge on the time evolution of process rates at the Harvard Forest for the simulation period, we could have calculated the time evolution of age and transit time distributions. But as an introductory paper on the main concept, we do not consider appropriate to include the additional complexity inherit of the time-evolving formulas. For this reason, we decided to use the autonomous case to introduce our concept.

I also disagree on the way the authors claim they evaluated their model using observation. [i.e. 'We found a good fit of the three model structures to the available data' [abstract]]. First, I understand that the data used for model "evaluation" are the same that the one used for model optimisation as these are the only data mentioned in the manuscript. Is the model calibrated against Harvard Forest data (i.e. results in Table 1) and then compared to the same data for evaluation (Figures 3 and 4)? Or did I miss something? Second, there is no information on how the simulations were done for evaluation.

This is a good point with regard to the "evaluation of the models", since in case that we actually wanted to evaluate them we should have used another data set. However, we actually never mentioned that we evaluated the models. In the figure 3 we simply showed that the model simulations fitted the data points, but this is only to give an idea that the predictions of C stocks are in accordance to a particular forest. We understand that this might be a source of confusion, but it is important to highlight that this work is a theoretical exercise, and the fit of the models to the data is only to have a rough example that can be related to a 'real' forest. We made our intentions clearer in the methods and in the results, as can be seen in the material that we included as supplement of this response.

Figure 4 clearly shows that the model is in transient conditions from 1950 to 2010, with wood carbon stocks increasing and being comparable to the observations in 2010. That would make some sense if the model actually started in 1950, with external forcing (climate, CO2, land use, etc.) changing from year to year. My understanding is that this is not the case here. The model is simply spinning up, slowly reaching steady state. The agreement in 2010 is hence completely artificial. Cstock wood does not seem the have reached equilibrium, running another 100 years and it would be well above the observations. Unless I missed these two elements, there is strictly no evaluation in this paper.

Again, we are not interested here in finding the best model that reproduces the entire history of C accumulation for the Harvard Forest site as modified by changes in atmospheric CO2 concentrations and climate change, but rather to find a set of realistic parameters that at least can reproduce the trend in carbon accumulation for some of the measured pools in this site. The data shows that this forest is not in equilibrium yet, and our transient simulation approaches these dynamics well. We consider this is enough to obtain a set of parameters that allows us to show some examples of the main concepts we want to introduce: age and transit time distributions of carbon. Please keep in mind that this is a conceptual paper, and we make no claims regarding the accuracy of the predictions for the specific site. We are rather interested in introducing a new set of model diagnostics that can be very useful for more specific simulations.

Finally the co-authors conclude that Differences in model structures had a small impact on predicting C stocks in ecosystem compartments, but overall they resulted in very different predictions of age and transit time distributions. I will argue here that considering the fact that each of their model parameters was constrained using the same carbon stocks, this is a result one should expect as a direct outcome of their methodology and no conclusion on the reality of the processes can be drawn from it.

We agree that sentences such as the one cited can be interpreted literally as “the model structures had a small impact on predicting C stocks differences”. However, the C stocks were not listed as one of the metrics that we used. Thus, what we meant was that although all the models had similar predictions in C stocks, they had important differences with regards to other metrics. To avoid this confusion, we made our point clearer in the abstract, results and discussion. It is anyway noteworthy that the results of the sensitivity analysis show that these data is an insufficient constraint, since different combinations can result in the prediction of similar C stocks, hence the equifinality section.

Reviewer 3

The manuscript by Ceballos-Núñez et al. describes a study on transit times of C through a temperate forest, with the help of a framework with a number of pools and transfer time coefficients between these pools. The number of non-structural pools is varied to investigate the importance of these for representing the transfer. Whereas the concept used in the manuscript is generally well explained, more information should be provided on the details. I have some reservations to the chosen methods, but this may also be partly caused by misunderstanding due to the too concise description. Some parts of the study need to be clarified, and slight alterations to the methodology may be required. This manuscript can become an interesting contribution to the literature on modelling carbon allocation in (temperate forest) ecosystems, but I would recommend the authors to clarify the methodology, and also to rethink some of the concepts that they apply in their setup. I hope that the comments below can support the authors in revising their work.

We thank referee 3 for an accurate summary of our manuscript. We agree that in the concise Methods section we left out some important details that could have increased the clarity of our work. Therefore, we made major modifications to the Methods section, which we have included as a supplement to this response.

Major remarks
A general issue with the methodology is the limited amount of observations that is used to constrain the transfer matrix B. Firstly, these data should be described more carefully, e.g., Figures 3 and 4 seem to suggest that wood or biomass numbers for multiple years are available, which does not appear in the Methods section. Secondly, and more importantly, it seems to me that the available data make it impossible to separate the non-structural pool(s) in the models, which leaves the system underdetermined. The authors also seem to refer to this in their methods (p 7/l 5). It is unclear how a Monte- Carlo setup for sampling the parameter space could help to constrain these, and in fact the authors seem to suggest that it does not. I would suggest to either explain more carefully how the Monte-Carlo setup was applied, which parameters were estimated with it, and which constraints were used, or to skip the Monte-Carlo setup and replace it by a more simple parameter

estimation based on literature values, which seems more appropriate for a system that cannot be constrained by observations.

We agree with the reviewer in that there was limited data to constrain values for the transfer matrix **B**, and that we did not carefully explained what data was used for the parameter estimation procedure. We included a description of the data that we used and the way in which we calculated the C stocks from it, which addresses the concern on the use of the wood biomass numbers. However, the issue of limited availability of observations is an issue that remains and for which we can do little about. We did in fact used reported values from the literature to fix some of the parameter values (P6L11) or to set ranges for the parameter optimization routines. Nevertheless, we had non-identifiable issues and we think it is important to report them in our analysis in a transparent way. In this sense, the Bayesian optimization we performed, helped us to honestly report uncertainty ranges for possible parameter values, but it is in no way a method to fix the non-identifiable problems we found. We think that by reporting uncertainty ranges for the parameters and the predictions we can better deal with the uncertainty related to the lack of observations and values from the literature.

Although the setup of the three steady-state models is generally well described in the methods, there is one aspect that I have not been able to resolve with regard to the photoassimilates pool. The text seems to suggest that this is not a real pool but just partitions the input vector u to the other pools and to autotrophic respiration, leaving it in essence a flux (or sum of fluxes) rather than a pool. However, in Fig. 2, C_P *receives its own box, and the presented results (e.g. Fig. 8) indicate that the photoassimilates pool also has an age. If this is indeed the case in the model, there would be one nonstructural pool (namely,* C_P*) in Storage 0, and two respectively three non-structural pools in Storage 1 and Storage 2. Maybe the authors can clarify whether this is the case or not. If it is, I would recommend to remove it as a pool, to reduce the size of the vector x and thereby reduce the number of parameters that need to be estimated.*

This is a compartment that we decided to keep from the original model proposed by Richardson et. at. (2013). This is in fact not a storage compartment, it is one of the two types of foliage compartments. Since the foliage is divided into photoassimilates and structural foliage, the former is strictly a compartment, and has its own state variable. However, this compartment does retain the carbon for a short time, giving the impression to be a component that partitions the input flux. Since this confusion might have been originated from the model schemes in figure 2, we changed the name of the variable C_f to C_{Strf} and we included a short statement about these two compartments in the legend of that figure. Also, to clarify, u is not a vector, it is a scalar, the partitioning vector is β.

The discussion section describes the limitations of the method to some extent, but it should also be used to discuss the validity of the assumptions that are made in the methods. E.g., how important is the annual timestep, and would the results look very different when resolving the annual dynamics? How would varying input u or varying transfer matrix B (as a function of meteorology) change the conclusions? How would a representation of changes in forest structure or regeneration affect the conclusions?

Thanks for the suggestion. Indeed, resolving annual dynamics and making the system time-depend would not have any impact on the main conclusions we derived from this study: 1) age and transit time distributions strongly depend on different carbon allocation schemes imposed by the model structure, and 2) observed mixes of carbon age in non-structural carbohydrate pools can be easily explained by the existence of age distributions in vegetation compartments. Explicitly resolving intra and inter-annual dynamics simply result in time-varying age and transit time distributions, but does not invalidate the existence of these distributions. We added a few sentences to the discussion to address this comment.

The results indicate that the slow non-structural carbon pool has lifetimes that exceed those *of the slowest other pools in the models (notably the wood). I would expect the authors to comment to this, as it is at least counterintuitive to have the non-structural carbon of higher age than the structural carbon. This may also be because of the un- derdetermination in the setup (see my comments above). Could the authors comment on that?*

We appreciate this detailed remark. Actually this is only an artifact of cropping the ranges of the distributions in the plots. Since this is clearly causing a confusion, we extended the ranges, so that it is possible to see that the tail of the age distribution for the wood compartment is longer.

The figures in the appendix need to be used in the paper. Is there a difference between Fig. 6 and Fig. A5?

Thank you for this remark; we added the missing references. The difference between figures 6 and A5 is the parameter set with which we run the models to obtain them; for the simulations presented in figure 6, each model was run with a different parameter set, while for figure A5 the same parameter set was used for the three models.

Minor remarks
- The abstract describes the model in very general terms before explaining the application to Harvard forest. Whereas it is fine to present the model in general terms first, I think it could help to mention early on that you are talking about forest ecosystems. You may consider adding "forest" to the title as well.

Although the examples presented in our manuscript correspond to a forest, the approach that we are suggesting is applicable to other systems. The use of age and transit time distributions as diagnostics can be implemented to models portraying other systems, as long as they follow the assumptions explained in the formulas. For this reason we kept the description of the model general enough so it can be applied to other systems.

- p 3/l 12: "random paths" - I do not think that "random" is appropriate in this context; the C can follow multiple paths, but it does not move randomly through a tree. Please consider replacing.

We understand that the term 'random paths' can be misinterpreted in the context of tree physiology. This is why we edited that line to read as follows: At each time step, a particle may randomly stay where it is, or flow to the next compartment with a rate (speed) given by the transfer coefficients (also know as cycling rates).

- p 6/l 14: There is some repetition in this sentence, and the crucial information on which data sets are used appears to be missing. Please check.

We are grateful for the identification of this issue. We indeed had the same title for different links. We reduced the redundancy and explained more what kind of data we used.

- p 6/l 21: Please explain what the constraint is that is determined from "the NSC calculations from Wood". Do you have a number for the 2nd NSC pool that can be used?

We explained better that we calculated the C stocks of NSC by using equations that derive it from the C stocks in wood.

- p 7/l 6: What is the parameter set with the highest frequency?

The parameter with the highest frequency is the one selected more often by the Bayesian optimization method during the parameter exploration. We clarified this in the text.

- p 9/l 8: replace "tree" with "three" (I suppose) - p 9/l 11: replace "stronger" with "larger"

We appreciate this remark, we made the corrections.

- Figure 3: One of the foliage lines appears to be invisible (or hidden behind one of the

other lines). Please check.

There was an overlap between the lines, so we fixed it by dividing that plot into two separate ones, one for wood and the other for foliage.

- It seems that Fig. 4 is not referred to in the text, whereas it is crucial for the paper. Please add a reference to it. - Figure 4: What causes the large spike in the root C stocks? Is this an initialization problem?

In fact we referenced the figure, but only with a number, which is probably the reason why it was overlooked. We have corrected that. Also, it is probably true that the cause of the large spike in the root C stocks might be due to having the wrong initial values for that compartment.

- Figure 7: The median for storage 1 appears to be missing from panel A.

The line of the median was hidden behind the mean line of another model, we corrected this by lowering the range of the x-axis.

- p 17/l 13: How do you determine the 20-year longer growth? The curves should follow an asymptotic behaviour towards the steady state.

The age and transit time distributions were obtained under the assumption that the models were in steady state. Thus, we can say that the mean ages are predictions of the mean ages of the vegetation once it reaches steady state. Since in each model the mean ages are different, we can consider that their difference is the remaining time that it takes them to reach steady state.

- Figure A4: Some lines appear to be missing.

In this case it is more complicated to fix the overlap because an outcome of running the three models with the same parameter set was that they all had similar mean and medians in their distributions.

Reviewer 4

The article develops an interesting approach to distribute structural and labile carbon into age distributions with the resulting transit time of carbon in the vegetation.They tested three different allocation models and used measured carbon values from Harvard Forest to parameterize each model and atmospheric isotopic carbon to compare against the isotopic signatures calculated from each structural component.

We are glad that the referee found the article interesting, but we have to clarify that we did not ‘distributed the carbon into age distributions‘; we calculated the age and transit time distribution of carbon particles in a vegetation system and its compartments.

Age and transit time distributions were different for the different allocation models, showing older age distribution in the model with more storage compartments, as well as with the wood components.
1. At the end of the introduction it is stated that the models are diagnosed according to four metrics, but only metrics 1 and 2 (carbon fluxes and bomb spike) are based on observations, and even those are not necessarily for each component, so the diagnosis is really just an analysis of how the different model results compare or what they imply rather than an actual metric based on observed values. There needs to be better distinction between what is actually observed vs what is modeled throughout the paper.

We interpret this and the previous comment as a sign that we need to make our Methods section clearer. Please find enclosed the new version. The only measurements that we used for parameter optimization were published aboveground biomass, LMA and LAI. These measurements were only used in the optimization procedures, to find suitable parameter values. The model comparison was mainly qualitative, using simulations, e.g. the method used for the radiocarbon is clearly stated in the first line of the section 3.2.2. In the rest we discuss that we were comparing the model simulations with respect to their expected dynamics, depending whether they were fast or slow cycling components.

2. So, one justification for this research is that age and transit times can be measured with isotopic tracers, but that was not done in this study. There should be some examples in the literature of these types of measurements or some attempt to use them to verify the results of this study. It would be nice if there were some way to prove that adding more compartments improves the age distribution and transit times. However, compared to the total atmospheric isotopic signatures, the multi-compartment schemes actually do worse, which the authors attribute to perhaps the lack of phenology.

In the Discussion we compared published mean ages of NSC with the ones predicted with the models. We agree that it would be more interesting to have measured C age distributions to compare with the simulations, but they are unfortunately unattainable with the current empirical methods. Also, one important clarification is that radiocarbon values for the different pools cannot be compared directly to atmospheric radiocarbon values as this comment seem to imply. They can only be compared to radiocarbon values measured on the tissue of the different pools.

3. The models used here are purely parameterized models without any processes at all, so how would this approach actually apply to models that were process-based? Bayesian optimization of parameters has been done before, but it is impressive that this approach was taken here. Any speculation about how this age and transit distribution approach could be applied and tested in some of the major ecosystem models?

The models we use are simple, which helps us to introduce the concepts of age and transit time of carbon in vegetation compartments. Independent on the complexity of the model being used, and the time steps of specific simulations, we can always expect mixes of C ages of different compartments, and this results in age and transit time distributions. The concept is independent on the complexity of the model, but the specific shape of the distributions would in fact depend on the specificity of the model. So, there is no consequence on whether the concept can or cannot be applied for more complex models, it only would affect the particular shape of the distributions.

4. Non structural carbon would seem to be the same thing as labile carbon, so maybe indicate that first time the term is used.

Thank you for this suggestion. We indicated that in the introduction.

5. The allocation models here really refer to models with different quality storage pools. The allocation itself seems to be simply constant rates but I would consider allocation model to imply different methods of determining allocation, i.e. literature-based rates, cost-benefit analysis, etc. I would prefer more clarity on how the constant rates differ between the pools, and how they are determined.

We agree that there are other methods to estimate the allocation rates, and that we did not explained the reasons why we chose to estimate the parameters of our models using the Bayesian optimization method. In the new version of the Methods section we mention that we chose this method because it gave us the possibility of exploring the parameter space and with that find parameter sets to do an uncertainty analysis of the model. The goal of the parameter estimation procedure was to obtain parameter sets for each model, which could allow them to simulate similar C stocks to those obtained from the Harvard Forest Archives. The upper and lower boundaries used to constrain the parameter values during the estimation were obtained from the literature, for a similar forest. As we explained in the point 3, we decided to use simple models (with fixed coefficients) to illustrate the concept of 'age and transit time distributions'; this is why the environmental forcing was not part of the estimation of the rates. However, it is important to clarify that even if we would have used other methods to obtain the rates, we would have still obtained different ages and transit time distributions for each model'.

6. P. 7, and of first paragraph: How do the functions used to calculate age and transit time distributions relate to the formulas in the introduction? Last Paragraph: What is meant by lower-diagonal with respect to the figure?

The formulas mentioned in the introduction were implemented as functions in the 'R' package 'SoilR', in order to calculate age and transit time distributions. With "lower-diagonal" we were referring to the values under the diagonal. We corrected that sentence.

7. P. 9, Figure 3: Are the observations data the dots and vertical lines? Seems like a large range for the error bars why so large, and what are they based on?

We agree that we did not explained properly the way in which we used the data obtained from the archives of Harvard Forest. We added a short explanation on how we used the measurements of aboveground biomass, LAI and LMA to calculate the C stocks that we show in figure 3. The large deviations from the C in wood is probably due to the large variation in the aboveground biomass measurements.

Table 2: What is the criteria for positive or negative correlations (i.e. how large and R2 value?)

The assignment of positive or negative correlations was performed based on the sign of the values observed in the pairwise plots. Only R^2 values < -0.1 and > 0.1 were assumed to account for correlations.

8. Figure 4 is not mentioned in the text. More detail on the source of the isotopic car- bon data would be useful. How are the "bomb spikes" used to determine radiocarbon values?

We had actually referenced the figure, but only with a number, which is probably the reason why it was overlooked. We corrected that mistake. The calculation of radiocarbon in the different compartments relies on a standard atmospheric radiocarbon curve (Hua et al. 2013), which is used to introduce radiocarbon in all vegetation compartments at rates specified by the vector u and the transfer matrix $\mathbf{B}$. This functionality is part of the 'SoilR' packaged used here. We added a short description in the methods about the incorporation of radiocarbon in the simulations and provided a citation for additional details of its implementation in the SoilR package.

9. P. 11: Not sure whatis meant by "Notice that distributions with the highest peak (occurred) closer to 0 years, and with younger mean and median ages had the youngest C".

That sentence was meant to explain how to interpret the shape of the distributions and, based on that, show that the model *Storage: 0* had younger C than *Storage: 1* and *2*. However this explanation might not be needed, so we removed that sentence.

10. P. 16, second paragraph: What parameters are the ones that compose the beta vector and B matrix?

The elements that compose the β vector and the $\mathbf{B}$ matrix are described in the equation (1). The components of the matrix $\mathbf{B}$ are represented by the symbols next to the arrows in figure 2, which are the parameters that control the C transfer through the compartments.

Not all the appendix figures (A4 to A9) are mentioned in the text.

We appreciate this remark, we have added the missing references.

At bottom of page, what are implications of the different age values i.e. how do the 1.71/2.14 and 45 yr relate to the 7.2?

This is the part that we mentioned in the response to the second comment. We were comparing the simulated mean ages of the NSC to the ones obtained from literature. The published values had a wide range. Thus, although the mean ages of the storage compartments with fast cycling were lower than the mean published value, they were still within the range.

We hope that we addressed the comments of the four reviewrs adequately and that this new version is suitable for publication.

Sincerely,

Verónika Ceballos-Núñez

Ages and transit times as important diagnostics of model performance for predicting carbon dynamics in terrestrial vegetation models

Verónika Ceballos-Núñez[1], Andrew D. Richardson[2, 3, 4], and Carlos A. Sierra[1]

[1]Max Planck Institute for Biogeochemistry, Hans-Knöll-Str. 10, 07745 Jena, Germany
[2]Department of Organismic and Evolutionary Biology, Harvard University, 02138 Cambridge, MA, USA
[3]School of Informatics, Computing and Cyber Systems, Northern Arizona University, Flagstaff, AZ 86011, USA.
[4]Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, AZ 86011, USA.

Correspondence to: Carlos A. Sierra (csierra@bgc-jena.mpg.de)

Abstract. The global carbon cycle is strongly controlled by the source/sink strength of vegetation as well as the capacity of terrestrial ecosystems to retain this carbon. These dynamics, as well as processes such as the mixing of old and newly fixed carbon, have been studied using ecosystem models, but different assumptions regarding the carbon allocation strategies and other model structures may result in highly divergent model predictions. We ~~modeled three systems of vegetation compartments and assessed their performance by calculating the age of the carbon in vegetation system and within each compartment, and the overall transit time of C in the system. We used these diagnostics to assess~~ assessed the influence of three different carbon allocation schemes on the ~~rates of~~ C cycling in vegetation. First, we described each model with a set of ordinary differential equations. Second, we used published measurements of ecosystem C compartments from the Harvard Forest Environmental Measurement Site to find ~~the best set of~~ suitable parameters for the different model structures. ~~Second~~And third, we calculated C stocks, release fluxes, radiocarbon values based on the bomb spike, ages, and transit times. We ~~found a good fit of the three model structures~~ obtained model simulations in accordance to the available data, but the time series of C in foliage and wood need to be complemented with other ecosystem compartments in order to reduce the high parameter collinearity that we observed, and reduce model equifinality. ~~Differences in model structures had a small impact on predicting~~ Although the simulated C stocks in ecosystem compartments ~~, but overall they~~ were similar, the different model structures resulted in very different predictions of age and transit time distributions. In particular, the inclusion of two storage compartments resulted in the prediction of a system mean age that was 10-20 years older than in the models with one or no storage compartments. The age of carbon in the wood compartment of this model was also distributed towards older ages, whereas fast cycling compartments had an age distribution that did not exceed 5 years. As expected, models with C distributed towards older ages also had longer transit times. These results suggest that ages and transit times, which can be indirectly measured using isotope tracers, serve as important diagnostics of model structure and could largely help to reduce uncertainties in model predictions. Furthermore, by considering age and transit times of C in vegetation compartments as distributions, not only their mean values, we obtain additional insights on the temporal dynamics of carbon use, storage, and allocation to plant parts, which not only depends on the rate at which this C is transferred in and out of the compartments, but also on the stochastic nature of the process itself.

1 Introduction

The global carbon cycle is strongly controlled by the source/sink strength of terrestrial ecosystems. Vegetation in particular, is one of the major controls of global C sources and sinks with respect to the atmosphere (Canadell et al., 2007); it has the capacity to be either a strong C sink or a source, depending on the amount of C fixed by the canopy and the time that C takes to transit through its components back to the atmosphere (Luo et al., 2003). Strong sinks therefore, not only fix carbon at a fast rate, but have also the capacity to store this carbon for long periods of time (Körner, 2017).

The C storage capacity of an ecosystem is determined by the collective behavior of vegetation compartments such as foliage, wood, and roots, which may also act as C sources and sinks among each other (Xia et al., 2013; Luo et al., 2017). The capacity of a vegetation compartment to oscillate between C source and sink has important implications for ecosystems ~~to respond~~ in their response to perturbations and environmental change, i.e their resilience. Carbon fixed during photosynthesis is transported from the leaves (sources) to other parts of the plant (sinks). One of these sinks is the labile or non-structural carbon (NSC) (Hartmann and Trumbore, 2016; Trumbore et al., 2015; Martínez-Vilalta et al., 2016), which may turn into a C source during critical events, such as the start of the growing season (after periods of limited photosynthesis) (Richardson et al., 2013), and the recovery from disturbances such as drought (Hartmann et al., 2013), cold temperatures (Hoch and Körner, 2003), pollution (Grulke et al., 2001), or nutrient stress (Ericsson et al., 1996). Despite the importance of the source-sink capacity of NSC reserves, many questions remain unsolved. For example, are NSCs completely depleted when needed, and replenished afterwards? Is the C that has remained stored for many years still available for the plant? (Richardson et al., 2013).

It is indeed possible that the carbon stored in vegetation compartments, including NSCs, have been fixed at different times, resulting in a mix of ages (Muhr et al., 2016). Studies across wood rings in temperate forest trees revealed that the mean age of NSCs in stemwood can be up to several decades old (Richardson et al., 2013; Trumbore et al., 2015). Trumbore et al. (2015) explained these old ages with a simple model consisting of one NSC compartment with inward mixing of younger and older C. Alternatively, Richardson et al. (2013) proposed a model with two separate storage compartments (NSC)-with old and young C, respectively- that exchange material among each other. It is therefore uncertain how this mixing of NSCs of different ages occurs: In the form of one single compartment in which all ages are mixed, or in different compartments with separate ages?

Previous studies have focused mostly on determining ages of NSCs using radiocarbon-derived mean residence times, but this approach has limitations. One limitation is the ambiguity in the term “mean residence time”, which has been defined in different ways across studies; in some cases it implies the mean age of C in an ecosystem or ecosystem compartment and in other cases it implies the time it takes C molecules to leave the system of compartments (Sierra et al., 2016). Another limitation is the use of mean values instead of complete frequency or density distributions to assess the spread of C ages in vegetation compartments.

~~Certainly the~~ The study of C age distribution in vegetation requires challenging empirical methods, but can also be approached using ecosystem C cycle models. However, not all of the models perform equally well because the assumptions behind their structures may result in highly divergent predictions (Lacointe, 2000; Friedlingstein et al., 2006; Friend et al., 2014; Schiestl-Aalto et al., 2015). The performance of such models has been diagnosed by comparing their predicted C storage capacity and residence times (Friend et al., 2014; Yizhao et al., 2015), but if the above mentioned ambiguities are resolved, ages of carbon in vegetation and in respiration fluxes can serve as excellent additional diagnostics of ecosystem models, and can give important insights about carbon metabolism in vegetation under stress conditions, such as in the case of drought stress.

1.1 Definitions: ages and transit times

CO_2 molecules are fixed continuously by photosynthesis during the growing season; so, C particles enter the vegetation (from here on called "system") at different times during the year. After fixation, the photosynthetic products transit through the vegetation compartments until they eventually leave the system, either as CO_2 back to the atmosphere or as litter and exudates to the soil. This means, that at a given time (t) each particle ~~of~~ in the system has a different *age* , which is the time that it has remained ~~in~~ within the system since its fixation from the atmosphere. The time that each particle spends transiting through the system, from arrival until exit, is called *transit time* (Bolin and Rodhe, 1973).

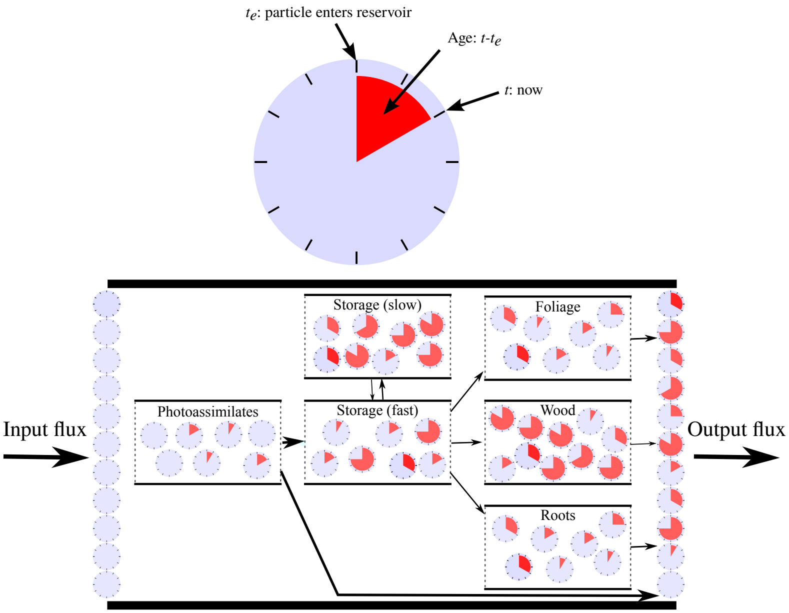


Figure 1. Graphical representation of the concepts of age and transit time distributions in a vegetation model. Carbon particles are represented here as clocks that measure the time ~~since~~ they have been in the system. **System age** can be defined as the age of all particles in the system at a given time, while **transit time** as the age of particles in the output flux. Adapted from ~~(Sierra et al., 2016)~~Sierra et al. (2016)

~~Although C particles follow random paths within a system, the speed at which C flows into or out of the system and from one compartment to another is given by certain rates (or transfer coefficients~~ At each time step, a particle may stay where it is given certain probability, or flow to the next compartment with a rate or probability given by the transfer coefficients (also know as cycling rates). This means that the age of carbon in a system ~~is determined by~~ results from stochastic and deterministic processes, which can be illustrated in Figure 1; given two organic molecules in a compartment, one that has remained in the system for longer time than the other, they both have the same chance to either leave the system, or move on to another compartment. Thus, if by chance older molecules remain for longer times, then the system's age gets older. But the pace at which these molecules are transiting is moderated by the cycling rates. This is why slow cycling compartments have older C.

Lets ~~consider~~ describe a system of well-mixed C ~~,~~ (distributed in multiple compartments~~. We can describe it with the following~~) with the system of ordinary differential equations ~~1:~~

$$
\begin{aligned}
&\dot{\boldsymbol{x}}(t) = \mathbf{B} \times \boldsymbol{x(t)} + \boldsymbol{\beta} \cdot u, \\
&\boldsymbol{x}(0) = \boldsymbol{x}_0,
\end{aligned}
\tag{1}
$$

where $\dot{\boldsymbol{x}}(t)$ is how much the quantity of carbon in vegetation compartment x changes with respect to time, $\mathbf{B}$ is a matrix of carbon transfer coefficients between the plant compartments, $\boldsymbol{x}(t)$ is a vector of states for vegetation (state variables), $\boldsymbol{\beta}$ is a vector containing the partitioning coefficients of photosynthetic input, and u is a scalar that represents that input. This linear system does not include environmental variables, or any other variables that depend on time, thus, it is an *autonomous linear system* with multiple interconnected compartments.

Given that each particle in the system has its own age and transit time, the age and transit time of the whole system can be considered as random variables. Additionally, the age and transit time of particles in a system's compartment is exponentially distributed. Then, the age and transit time distributions of the entire system would be the sum of those exponential distributions, i.e. a phase-type distribution (PT) (Metzler and Sierra, 2017).

The calculation of how many C particles have a certain age, or *the age density distribution of a system* ($f_A(y)$), is determined by the probability of entering the system through a given compartment and the rates at which C is transferred from one compartment to another until it leaves the system. Consistent with the symbols from the previous equation

$$
f_A(y) = \boldsymbol{z}^T \cdot e^{y \cdot \mathbf{B}} \cdot \frac{\boldsymbol{x}^*}{||\boldsymbol{x}^*||}. \tag{2}
$$

$f_A(y)$) is a function of (i) how fast the carbon is leaving the system: the row vector of release rates, which is the column-wise sum of the elements of the ~~matrix~~ $\mathbf{B}$ matrix ($\boldsymbol{z}^T = -\mathbf{1}^T\mathbf{B}$), (ii) the transition probability matrix ($e^{y\mathbf{B}}$), and (iii) the relative amount of C stock at steady state with respect to the total ($\frac{\boldsymbol{x}^*}{||\boldsymbol{x}^*||}$). Notice that we use here the symbol $||\cdot||$ to represent the vector norm, which is the sum of all entries of the vector.

The mean age is given by the expected value ($\mathbb{E}[A]$)

$$
\mathbb{E}[A] = \frac{||\mathbf{B}^{-1} \times \boldsymbol{x}^*||}{||\boldsymbol{x}^*||}. \tag{3}
$$

Likewise, the transit time density distribution ($f_{FTT}(t)$) is also a function of $\boldsymbol{z}^T$ and the transition probability matrix ($e^{t\mathbf{B}}$), as well as the vector of input distributions ($\boldsymbol{\beta}$).

$$f_{FTT}(t) = \boldsymbol{z}^T \cdot e^{t\cdot\mathbf{B}} \cdot \boldsymbol{\beta}. \tag{4}$$

The mean transit time is defined as ($\mathbb{E}[FTT]$):

$$\mathbb{E}[FTT] = ||\mathbf{B}^{-1} \times \boldsymbol{\beta}|| = \frac{||\boldsymbol{x}^*||}{||\boldsymbol{u}||} \tag{5}$$

In this case, the definition of mean transit time coincides with the commonly used *stock over flux* approach (turnover time), but note that the definitions presented here can only be applied to autonomous systems at steady state. ~~There are other~~ For non-autonomous systems, i.e. models in which inputs and process rates change over time, formulas for the estimation of mean age and transit time ~~distributions of non-autonomous models (e.g., Rasmussen et al., 2016)~~can be found in Rasmussen et al. (2016).

From these equations it is evident that age and transit time calculations mainly depend on the schemes of C partitioning ($\boldsymbol{\beta}$) and cycling ($\mathbf{B}$) within a vegetation model. Therefore, if we want to understand processes such as the mixing of old and newly fixed NSC using ecosystem models, it is critical to model proper carbon allocation (CA) strategies. Unfortunately, it is still uncertain what assumptions and simplifications should be done: How many carbon compartments are necessary to describe carbon cycling in vegetation? How are these compartments interconnected and how fast are they transferring C among each other? These are important questions that need to be addressed to improve our understanding of vegetation dynamics, and predict consequences of environmental change on vegetation.

In this contribution, we address the question: how different C allocation schemes affect the ages and transit times of carbon in vegetation models? In particular, we are interested in understanding whether different carbon allocation strategies would lead to different patterns of mixing of ages for the NSC compartment. For this work, we implemented 3 carbon allocation schemes based on ~~(Richardson et al., 2013)~~Richardson et al. (2013); the models have either no storage, 1 storage compartment, or 2 storage compartments (fast and slow C cycling). ~~First~~Our approach is mostly theoretical, and we are mainly interested in introducing the concepts of age and transit time distributions as useful model diagnostics as well as as an approach to explain mixes of C age in NSC compartments. For this purpose, we used published measurements ~~of~~ on ecosystem C compartments from the Harvard Forest Environmental Measurement Site to find ~~the best set of parameters~~ suitable parameter values for the different model structures. ~~And second, we~~ We also diagnosed the performance of these models using as metrics 1) C release fluxes (respiration and other carbon losses such as litterfall), 2) the dynamics of radiocarbon (based on the bomb spike) for individual compartments, 3) the transit time distribution of the system, and 4) the age distribution of C in the system and in each compartment.

2 Methods

~~We~~

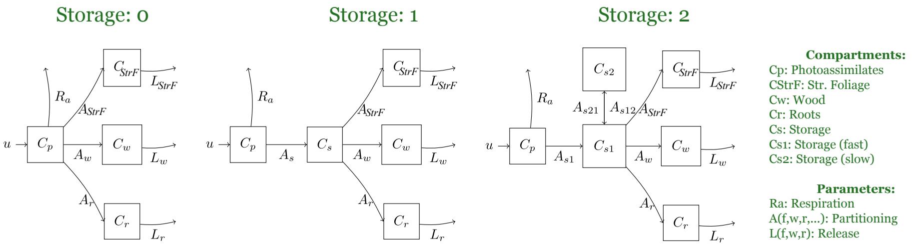


Figure 2. Three carbon allocation strategies in vegetation models. These strategies differ in the number of storage compartments, Storage: 0, Storage: 1, and Storage: 2. Adapted from (Richardson et al., 2013). The parameters are the rates at which the carbon cycles into and out of the compartments; thus, the fluxes are proportional to the C stocks and the rates. Notice that the foliage is divided into two compartments: Photoassimilates and Structural Foliage.

2.1 Model Implementation

Each model was written as a set of ordinary differential equations (based on Equation 1) within the environment of the R package `SoilR` (Sierra et al., 2012). All models met the requirements of equation 1; they are autonomous (their dynamics do not depend on variables that change with time) linear systems with multiple interconnected compartments. The initial carbon stocks, and some of the parameter values needed to solve those equations were obtained from the literature, from a deciduous-evergreen model with similar carbon allocation schemes (Fox et al., 2009). Other parameter values were obtained from an optimization procedure (see below).

As means to assess whether the carbon allocation strategies had an impact on the mixing of C age in vegetation compartments, we implemented three models whose carbon allocation strategies varied depending on the number of storage compartments (0, 1, or 2) (Figure 2), following the hypotheses proposed by Richardson et al. (2013). ~~The core structure of the models is~~ Given that we aimed at a theoretical comparison of the above mentioned strategies, we eliminated other potential sources of variation that may act as confounding factors by assuming that all C transfers between the compartments depended on constant rates. We also assumed that there was a constant photosynthetic input -gross primary production (GPP): u-of 1400 $gCm^{-2}year^{-1}$ (Urbanski et al., 2007)~~, which~~ . Environmental variability, which operate mostly on an hourly-daily time scales, was not not considered here because we ran the models at an annual time scale; i.e., without diurnal cycles or phenology.

Fixed photosynthetic input enters the system through the Photoassimilates compartment~~. Part of the C in this compartment~~ , and part of the carbon is released back to the atmosphere at each time step, in a flux proportional to the size of Photoassimilates and the constant rate R_a ~~; the environmental variables (which operate in an hourly-daily time basis)were not not considered because we ran the models at an annual time scale; i.e., without phenology~~(Figure 2). In the model without storage compartment, the C stored in the Photoassimilates is partitioned into Structural foliage (from here on: Str. Foliage), Wood (including branches and coarse roots) and Fine roots (from here on: Roots), with the constant rates A_f, A_w, and A_r; part of the C stored in these three compartments also leaves the system with constant rates L_f, L_w, and L_r, which comprise all the carbon released through respiration and other losses (e.g. litterfall). For the models with storage, the C is transferred from the Photoassimilates to the fast cycling storage, from which it is then partitioned to the rest of the compartments. In addition to the fast cycling storage, the model with 2 storage compartments also has a slow cycling compartment. ~~All C transfers between these compartments depend on constant rates. These models are autonomous (no time dependencies) linear systems with multiple interconnected compartments, and therefore they meet the requirements of equation 1. They were implemented in R using the R package `SoilR` (Sierra et al., 2012).~~

~~Three carbon allocation strategies in vegetation models. These strategies differ in the number of storage compartments, Storage: 0, Storage: 1, and Storage: 2. Adapted from (Richardson et al., 2013). The parameters are the rates at which the carbon cycles into and out of the compartments; thus, the fluxes are proportional to the C stocks and the rates.~~

2.2 Optimization procedures

~~The initial parameter values were obtained from the literature, from a deciduous-evergreen model with similar carbon allocation schemes (Fox et al., 2009). However, in order to find the best set of parameters for these models~~To obtain results that can be related to a particular ecosystem, we performed a parameter estimation procedure using published measurements of two ecosystem C compartments from the ~~, ,~~ Harvard Forest Environmental Measurement Site (see links 1, 2, and 3, and the raw LAI) . Harvard Forest is a regenerating temperate forest located in Petersham, Massachusetts (42.54°N, 72.18°W, 340 m asl)~~; among~~ . Among the tree species that are found in this 65- to 85-year-old mixed deciduous forest are: red oak, red maple, white and red pine, yellow and white birch, beech, ash, sugar maple and hemlock (Wofsy et al., 1993). ~~We performed a two-step optimization procedure using first~~

We calculated C stocks in wood and foliage from the above mentioned aboveground biomass, LAI and LMA, using allometric equations. Although in previous studies performed at the same site the use of woody biomass increment and LAI reduced the uncertainties in the predictions of net C sequestration and foliage dynamics, respectively (Keenan et al., 2013), and the wood stock observations have been used to successfully constrain fine root mass simulations (Smallman et al., 2017), we obtained unrealistic simulations of C stocks in storage and roots because these pools were not well constrained. Thus, we estimated C stocks of roots based on the assumption that shoot:root ratio = 1:5, and we also used published ratios to calculate NSC from Wood carbon (p. 5 Richardson et al., 2010).

We were also interested in observing the uncertainty of the model simulations (see section below), so we performed a Bayesian optimization, which gave us alternative parameter sets after exploring the parameter space. This optimization procedure was started using the result of a classical optimization method ~~and plugging in the results into a Bayesian optimization~~ using the *R* package `FME` (Soetaert and Petzoldt, 2010). Taking into account that data uncertainties have a direct influence on the fit of the outcome of the parameter estimation (Richardson et al., 2010), we accounted for the uncertainty in the data using the standard deviation of the measurements in the cost functions.

~~Since the lack of constraints for the other compartments resulted in unrealistic C stocks, we used the above mentioned data to calculate the C stocks in Roots and Storage and further constrain the parameter estimation. For this purpose, we relied on the assumption that shoot:root ratio = 1:5, and the NSC calculations from Wood (Richardson et al., 2010, p. 5).~~

As means to evaluate whether the parameters could be estimated from the given data sets, i.e. parameter identifiability, we performed a local sensitivity analysis and estimated the collinearity of the parameter sets with the package `FME`. The obtained collinearity index γ expresses the degree at which pairs of parameters are linearly related. Values of $\gamma > 20$ indicate high collinearity among parameters and poor identifiability of the model given the available data (Soetaert and Petzoldt, 2010). ~~Since~~

Given the high correlation between some of the parameters~~were highly correlated, we ran~~, we decided to run all model simulations using the ~~posterior parameter set with the highest frequency.~~

~~With the obtained parameter sets, we~~ parameter set that was most frequently chosen by the Bayesian optimization method. We then calculated C stocks, release fluxes, radiocarbon values based on the bomb spike, ages, and transit times, using functions implemented in `SoilR`. The functions that calculate age and transit time distributions are based on the formulas proposed by Metzler and Sierra (2017).

2.3 Uncertainty analysis

In order to explore model predictions that could result from different parameter sets that were possible and likely, we extracted a random sample of 1000 posterior parameter sets from the Bayesian optimization that used Markov chain Monte Carlo. We ran the models with the unique sets, and calculated the weighted mean and standard deviation of the C stocks, the released C from each compartment, and the system's mean age and transit time. The weights corresponded to the number of ~~repetitions of each unique parameter set~~ times that each parameter set was repeated in the sample.

~~In summary, we implemented three models with carbon allocation strategies that varied depending on the number of storage compartments: *Storage: 0*, *Storage: 1*, and *Storage: 2*. We estimated the parameter values using available published data, and then we calculated the local sensitivity analysis and collinearities of the parameters. Finally we assessed the influence of the carbon allocation strategies on ecosystem level C cycling using different metrics.~~

Code and data availability. All of the simulations and figures for this work can be reproduced using the code and data provided in the supplementary material.

3 Results

3.1 ~~Parameter estimation, local sensitivity analysis, and collinearities~~Simulations of carbon stocks

~~Model prediction using the best parameter sets estimated from both optimization procedures (Table 1) provided a good fit to~~ The C stocks simulations obtained from the three models were within the uncertainty range of the available data (Figures 3 and 4). The simulations of ~~C stocks in Wood~~ Wood and Foliage C (Photoassimilates + Structural) stocks were in accordance with the stocks estimated from the aboveground biomass inventory data ~~(Figure 3) . Predictions of Cstocks in Foliage (Photoassimilates + Structural) were also as expected from the data (4).~~ and the LAI, respectively.

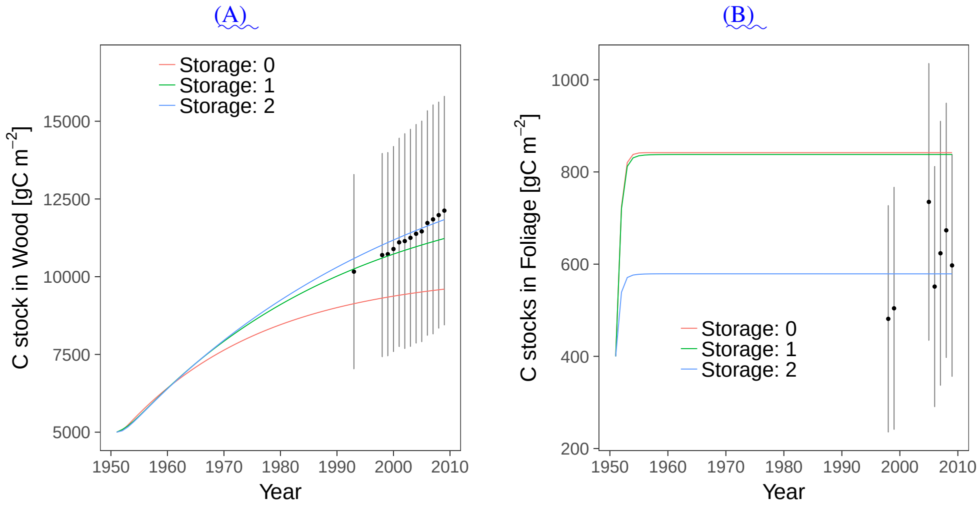


Figure 3. Carbon stocks estimated for each model, comparing the observed data and the model predictions of C stocks in Wood (A) and Foliage (B). The C stocks of foliage from the models Storage: 0 and 1 have overlaps.

All of these predictions were obtained using the parameter set that was most frequently chosen by the Bayesian optimization method (Table 1). This table also shows two quantiles of the distribution of each parameter value after exploring the parameter space.

~~Carbon stocks estimated for each model, comparing the observed data and model predictions for the years with available data. These simulations were obtained using the best parameter set.~~

~~(A) (B) (C) (D) Carbon stocks estimated for each compartment and their uncertainties. Carbon in the (A) Foliage (Photoassimilates + Structural), (B) Wood, (C) Roots, and (D) Storage compartments.~~

~~Despite the good fit~~Interestingly, some of the parameters were strongly correlated among each other. For the three model structures and the available empirical data, the number of parameters that can be simultaneously estimated with a collinearity index < 20 was 3 for the model without storage, and 5 for the other models. The correlations can be seen in the pairwise plots of sensitivity functions (Figures ~~A1-A3~~A1-A3). Table A1 summarizes the number of parameter correlations that we observed ~~in the lower-diagonal~~ under the diagonal of the pairwise plots of sensitivity functions~~:~~

~~Number of positive and negative correlations between parameters Model Possitive correlations Negative correlations Possible combinations *Storage: 0* 5 12 21 *Storage: 1* 4 10 28 *Storage: 2* 16 16 45~~. For this reason, the results presented here need to be interpreted within the context of predicted uncertainties.

3.2 Influence of carbon allocation strategies on ecosystem level C cycling

To assess the impact of different carbon allocation strategies on the ecosystem C cycling, we used the following metrics: 1) C release fluxes, 2) dynamics of radiocarbon for individual compartments, 3) transit time distribution of C through the system,

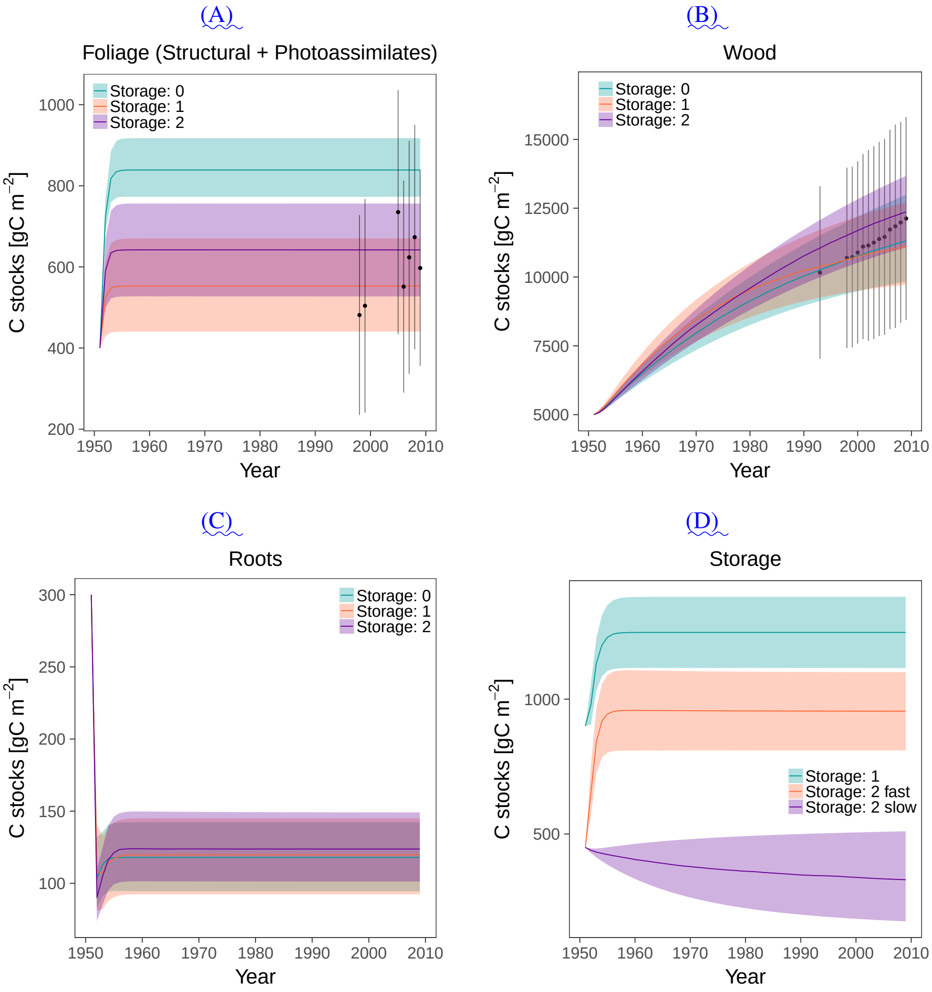


Figure 4. Carbon stocks estimated for each compartment and their uncertainties. Carbon in the (A) Foliage (Photoassimilates + Structural), (B) Wood, (C) Roots, and (D) Storage compartments.

and 4) age distribution of C in the system and in each compartment. The calculations required for these metrics were performed using the ~~best values of the parameter sets that resulted from~~ parameter set that was most frequently chosen by the Bayesian optimization method for each model, unless otherwise noted.

3.2.1 Fluxes of C released from the compartments

The ~~tree~~ three models predicted different mean fluxes of C released from each compartment at steady state (Figure 5). The only exception was the Roots, which had large uncertainties and overlaps among the flux distributions of the three models. This means that certain combinations of model structures with parameter sets result in similar predictions of Root C release fluxes.

Table 1. Parameter values obtained from the optimization procedures [year^{-1}].

Model	Parameter	Final	Best1	Best2	Median	q_{25}	q_{75}
Storage: 0	Ra	0.28	0.7	0.7	0.6	0.52	0.66
	Af	0.48	0.5	0.5	0.4	0.33	0.46
	Ar	0.44	0.5	0.5	0.39	0.31	0.45
	Aw	0.49	0.5	0.5	0.41	0.33	0.46
	Lf	23.32	35.35	35.35	19.82	12.65	27.72
	Lr	3.03	2.71	2.71	2.67	2.23	3.09
	Lw	0.04	0.02	0.02	0.02	0.02	0.03
Storage: 1	~~Ra1~~ Ra	0.67	0.5	0.33	0.45	0.34	0.56
	~~Af1~~ Af	0.27	0.19	0.4	0.36	0.24	0.44
	~~Ar1~~ Ar	0.12	0.4	0.32	0.25	0.18	0.32
	~~Aw1~~ Aw	0.25	0.38	0.24	0.38	0.31	0.44
	~~Lf1~~ Lf	10.43	18.25	28.93	16.12	8.87	24.97
	~~Lr1~~ Lr	1.63	3.27	3.39	2.73	2.03	3.2
	~~Lw1~~ Lw	0.03	0.03	0.02	0.04	0.03	0.05
	As	1.07	2.02	2.23	2.29	1.92	2.89
Storage: 2	~~Ra2~~ Ra	0.65	0.64	0.52	0.58	0.5	0.65
	~~Af2~~ Af	0.36	0.3	0.5	0.43	0.38	0.47
	~~Ar2~~ Ar	0.26	0.47	0.48	0.4	0.33	0.45
	~~Aw2~~ Aw	0.24	0.49	0.47	0.34	0.27	0.4
	~~Lf2~~ Lf	12.88	35.63	27.44	16.48	10.28	26.96
	~~Lr2~~ Lr	2.47	2.19	3.15	2.96	2.52	3.3
	~~Lw2~~ Lw	0.02	0.01	0.03	0.02	0.02	0.03
	~~As1~~ As	1.92	1.99	1.96	1.75	1.5	2.13
	As12	0	0.15	0.02	0.01	0.01	0.02
	As21	0.02	0.01	0.03	0.04	0.03	0.07

Final: Parameter set that was most frequently chosen by the Bayesian optimization method and was used for all of the simulations, unless otherwise noted.

Best1: Parameter set obtained from the classical optimization procedure.

Best2 and the remaining columns were the result of the Bayesian optimization.

However, for the compartments related to the foliage the differences were ~~stronger~~larger. Thus, regardless of the parameter sets, the differences in model structure lead to the prediction of different C release fluxes at steady state.

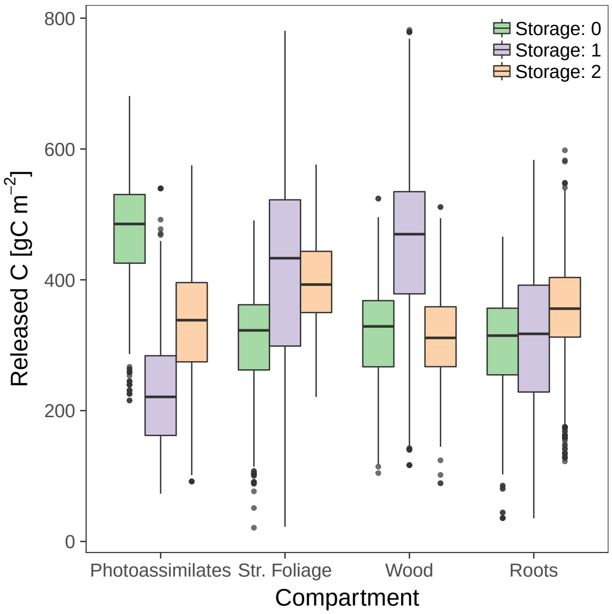


Figure 5. C release fluxes from the compartments at steady state, with uncertainty ranges obtained from the set of posterior parameters obtained by Bayesian optimization.

3.2.2 Radiocarbon content in each compartment

The simulated radiocarbon content of fast cycling compartments (e.g. Photoassimilates, Str. Foliage, Storage (fast), and Roots) had a stronger resemblance to the atmospheric Δ^{14}C values than the slower cycling compartments (Figure 6). However, for the Str. Foliage of the models with storage there was a time lag of about 3 years with respect to the peak that corresponds to the 'bomb-spike'. ~~This short time lag might be the result of that lack of phenology in the models.~~ Furthermore, the accumulation of radiocarbon in slow cycling compartments such as Wood and the Storage (slow), was characterized by a slow incorporation of radiocarbon that resulted in large Δ^{14}C values of the last part of the curve.

Differences in radiocarbon values for the different compartments hint to different levels of mixing of carbon fixed at different times. For the fast cycling compartments such as the Photoassimilates, the degree of mixing is relatively low because most of the radiocarbon reflects the values in the atmosphere. For other compartments that cycle at slower rates, the mix of recent and old radiocarbon results in important divergences from the atmosphere. Mixing of carbon of different ages can be further studied with ages and transit times distributions.

3.2.3 Age and transit time distributions

The age and transit time distributions were calculated assuming that the system was in steady state. These distributions had a wide range, expanding from 0 to several decades old carbon, and their shape varied according to the model structure (Figure

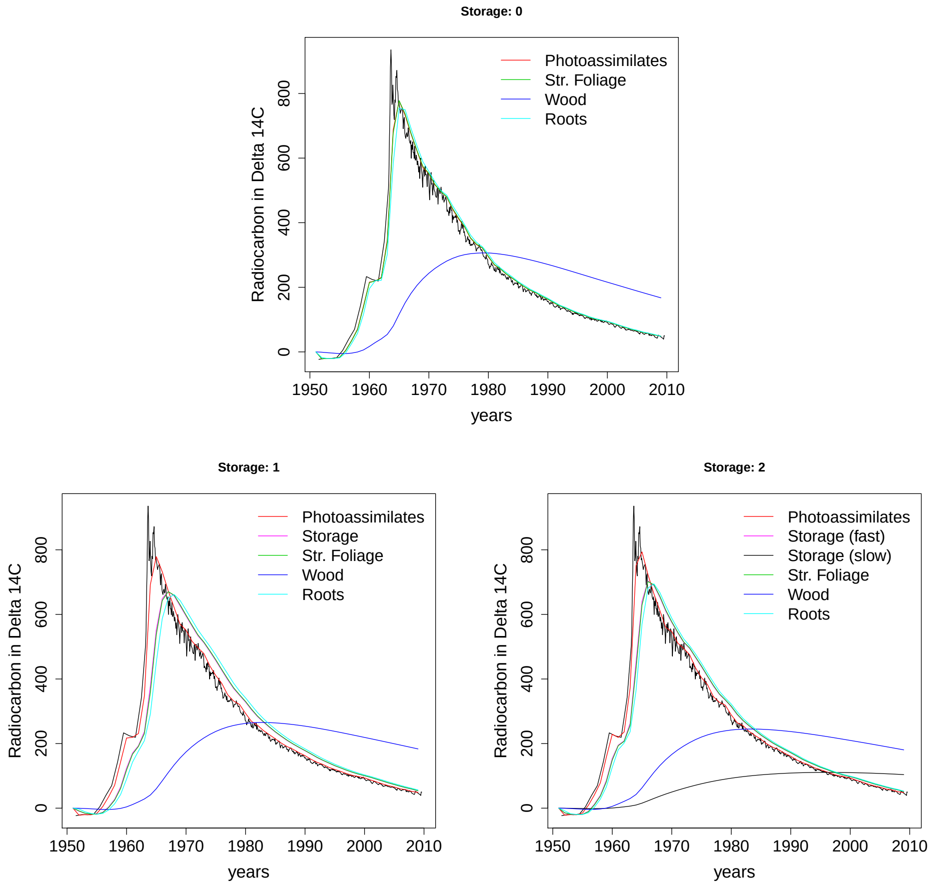


Figure 6. Radiocarbon simulations for the three model structures. The black curve corresponds to the Δ^{14}C in the atmosphere, and the other colors depict the vegetation compartments.

7). ~~Notice that distributions with the highest peak closer to 0 years, and with younger mean and median ages had the youngest C. Thus, the~~ The ascending order of the models according to their mean ~~ages~~ age, from young to old, was: *Storage: 0*, *Storage: 1*, *Storage: 2*. As expected, the model with the oldest ages (*Storage: 2*) had the longest transit time. These trends were partially observed when we analyzed the uncertainties in mean age and mean transit times (Figure 7 (C) and (D)), but the uncertainties for the model without storage were large. The large uncertainties regarding the mean ages and transit times of the model *Storage: 0* might have resulted from the high correlation between its parameters.

~~The~~ At the compartment level, the above mentioned age-dependent ranking of the models only ~~hold~~ holds true for Wood (Figure 8), which was the compartment with the closest resemblance to the overall system age densities because it comprised most of the mass in the system. The inclusion of two storage compartments in *Storage: 2* resulted in a relatively ‘flat’ distribution; with a ~~peak at very young ages, but with a~~ long tail that ~~lead~~ leads to a mean age 10-20 years older than the other two

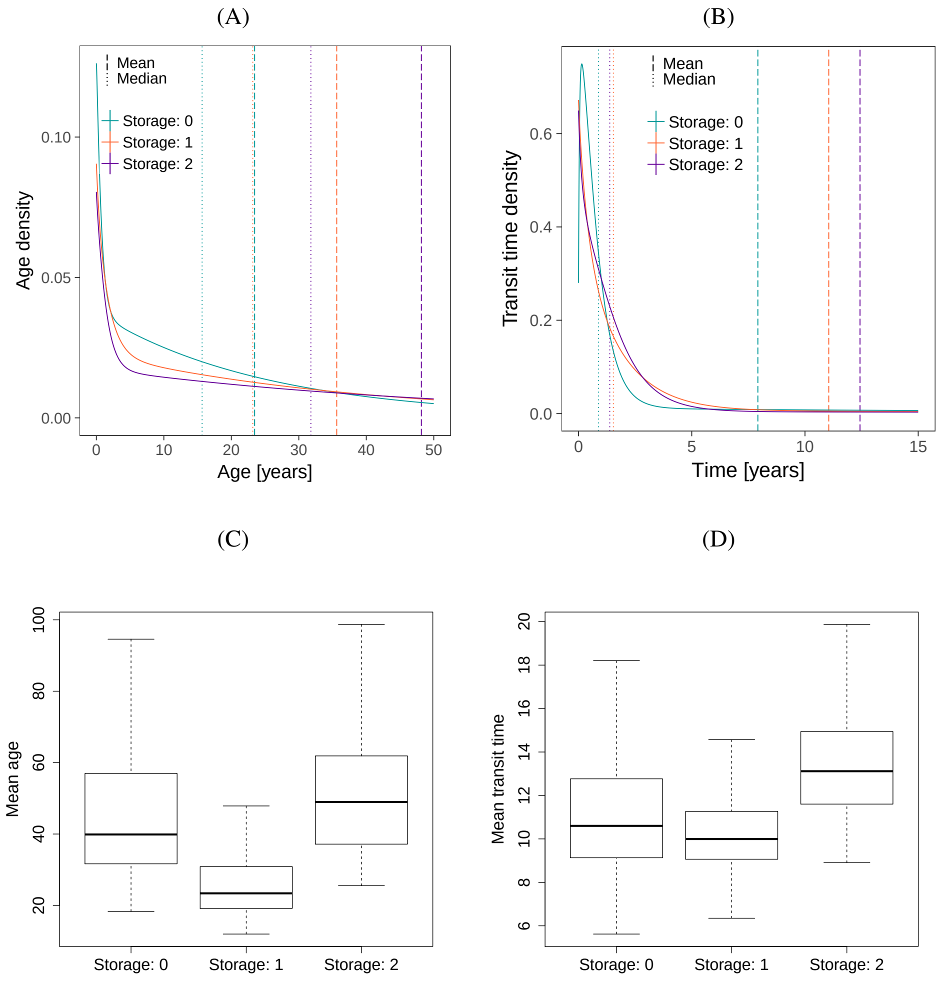


Figure 7. System ages and transit times. (A) Age and transit (B) time density distributions calculated for each of model structure using the best parameter set from the optimization; the dashed and dotted lines mark the mean and median ages, respectively. (C) Spread of mean ages (left) and mean transit times (D) obtained from all posterior parameter sets from the Bayesian optimization.

models, but with a peak at very young ages. This contrasts with the other models, which peaked at around the same time, but had steeper ~~curves~~declines with age.

The only compartment that had an age maximum at 0 years was the Photoassimilates. Hence, this compartment had a unique distribution curve reflecting the fact that all new carbon (age = 0 years) enters the models only through this compartment~~and it is later transferred to the others~~. Although the other fast cycling compartments (Str. Foliage and Roots) had peaks after 0 years, their C age was distributed towards young ages, with mean ages between 1 and 3 years. This spread in the C age of Str. Foliage

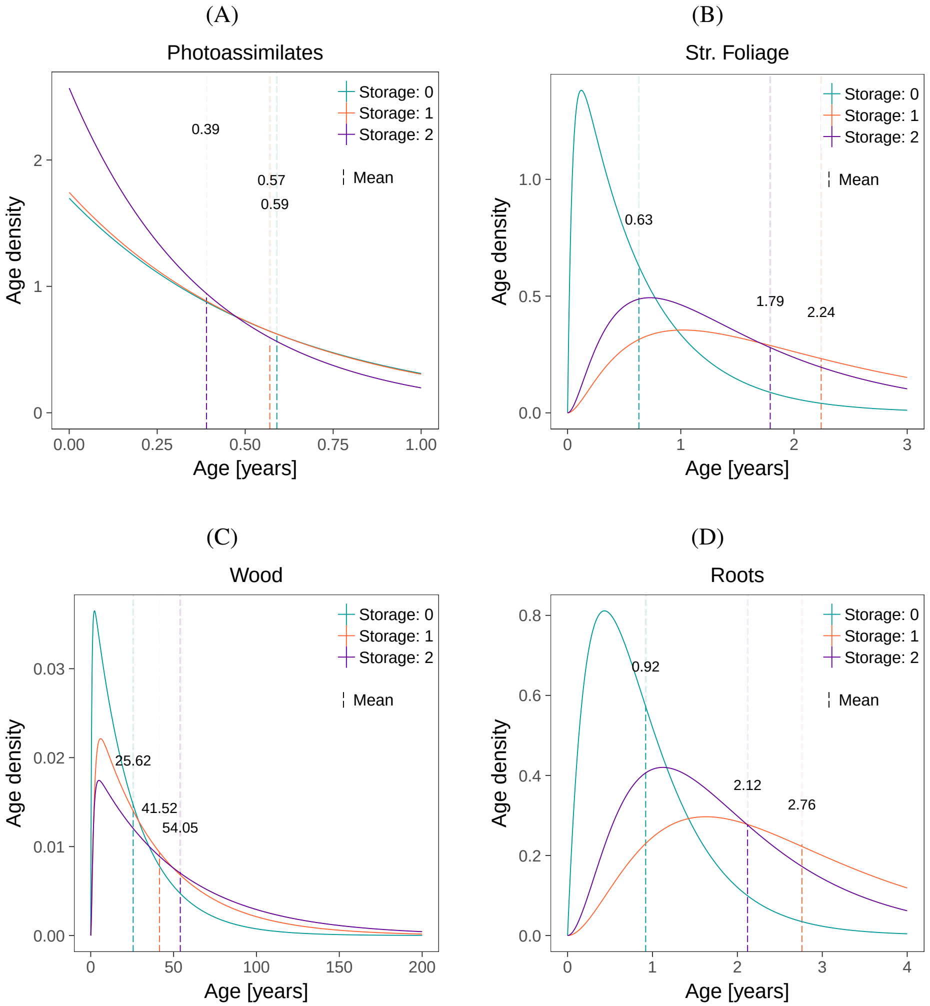


Figure 8. Age densities simulated for the compartments: (A) Photoassimilates, (B) Str. Foliage, (C) Wood and (D) Roots. Each model structure is depicted in a different color. Dashed lines correspond to mean ages.

for the models with storage may suggest either that the cycling rate of this compartment was relatively slow or that it received C from compartments with older carbon.

The age densities of the storage compartments, just as the ones for Foliage, Wood, and Roots, consisted of curves with peaks at young ages and long tails (Figure 9). In the case of the fast cycling compartments, the mean age of the models *Storage: 1* and *Storage: 2* was 1.71 and 2.14 years, respectively, but the long tail indicates that it is also probable to find 5-year-old C in this compartment. The mean age of the slow cycling compartment was 45 years, but the mixing of ages is also observed in the density curve, in which the age of C ranged from 0 to more than 150 years.

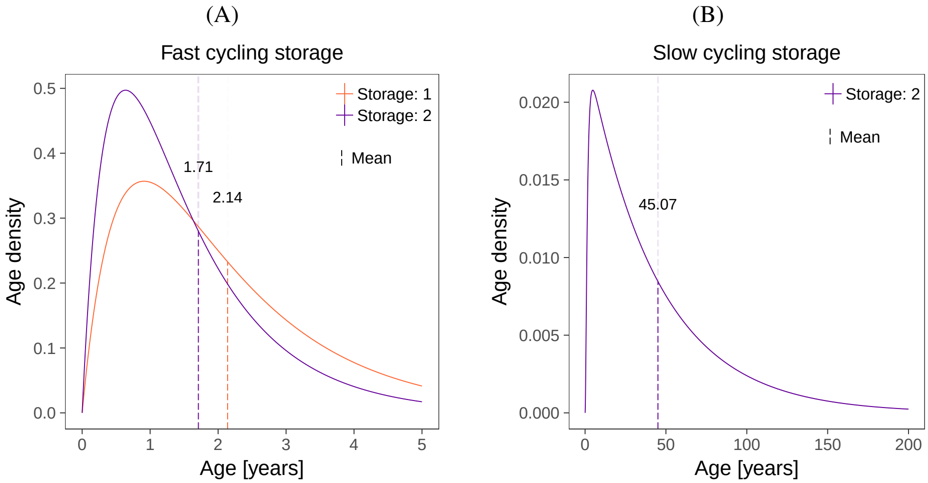


Figure 9. Age densities simulated for the models with storage compartments. (A) Fast cycling compartment of models *Storage: 1* and *Storage: 2*. (B) Slow cycling storage of the only model with 2 storage compartments.

4 Discussion

Our simulation results showed that C cycling in ecosystems can be largely influenced by different carbon allocation strategies, which may result in diverging carbon cycling predictions for specific simulations. However, not all of the different prediction metrics were impacted with the same strength by the assumed number of compartments and values of cycling rates, so results here need to be interpreted within the context of predicted uncertainties.

4.1 Diagnosing model performance with C release fluxes, and age and transit time distributions

The simulated ecosystem properties that were more strongly impacted by the assumptions behind model structure were: the fluxes of C released from each compartment, and the ages and transit time distributions of carbon in the system and in each compartment. This sensitivity to different carbon allocation strategies makes them good candidates for diagnosing model performance.

Given that the C release fluxes from Photoassimilates, Str. Foliage, and Wood were highly sensitive to the three model structures, empirical measurements of these compartments could be used as constraints during the parameter estimation procedure. ~~This is congruent with previous studies performed at the same site, where the use of data regarding woody biomass increment and LAI reduced the uncertainties in the predictions of net C sequestration and foliage dynamics, respectively (Keenan et al., 2013). Wood stock observations had also been proven to constrain well simulations of fine root mass (Smallman et al., 2017)~~

~~Despite less sensitivity to the different carbon allocation schemes, the~~ The radiocarbon accumulation was less sensitive, but can be useful to diagnose the models' performance according to the cycling speed of their compartments. As an example, we could identify the short delay of the Δ^{14}C signature of the Str. Foliage compartment with respect to the current year's atmosphere, in the models with storage (Figure 6). This means that this compartment had a slightly slower C cycling than what is expected for a deciduous forest (Trumbore et al., 2002, 2015). This delay in radiocarbon accumulation was also observed as shifts to older ages in the age distributions (Figure 8), with a resolution of months. In general, radiocarbon measurements can also help to constrain model parameters with respect to how fast or slow different compartments cycle C (Trumbore et al., 2016), but with less resolution than the age and transit time distributions.

Overall, the age and transit distributions were the best candidates to diagnose model performance and potentially constrain parameter estimations, because they were the most sensible to differences in model structure and parameter values. So, what had the highest impact on these distributions, the ~~inclusion of storage compartments or the~~ differences in cycling rates or the inclusion of storage compartments? It seems like these two characteristics had direct and indirect effects on the predictions of the above mentioned distributions, respectively.

On the one hand, as we initially inferred from the equations 2-5, the calculation of age and transit time distributions depends on the C partitioning schemes ($\boldsymbol{\beta}$) and the transfer -cycling- rates between plant compartments ($\mathbf{B}$). Therefore, different parameter values that compose the vector $\boldsymbol{\beta}$ and matrix $\mathbf{B}$ directly result in the different calculations of ages and transit times. This was tested by running the three models using the same parameter set; although they still differed in the number of storage compartments, the differences between the age and transit time distributions in the whole system and in the compartments was minimal (Figures ~~A6 and A7~~A6, A7, and A8). These similarities were also observed for the radiocarbon accumulation (Figure ~~A5~~A5). The only exceptions to the above were the fast cycling compartments of the model *Storage: 0*, which were faster than in the other two models.

On the other hand, model structure had an indirect effect on the predictions of age and transit time, most likely because the addition of storage compartments impacted the outcome of the parameter estimation and these different parameter values then lead to different age and transit time predictions. An illustration of this is the fit of the models to the data (Figure ~~A4~~A4): Running the three models with the parameter set that gave the best fit for the model with 2 storage compartments hampered the fit of the model with no storage compartments to the woody C measurements. Thus, although the system had an external C input that could never be depleted because it was assumed to be a constant flux, the parameter estimation accounted for extra compartments by modifying the cycling rates to optimize the fit of the models to the data.

As expected, systems with ages distributed towards older values also have older transit time distributions. In fact, their correlation can be confirmed by observing the formulas once again. The calculation of these two properties depends on the matrix of transfer coefficients ($\mathbf{B}$), but there are other factors driving these two distributions~~;~~. The additional factor driving the mean age calculation is the relative amount of C stock at steady state; whereas the mean transit times depend on the C partitioning schemes ($\boldsymbol{\beta}$) (Metzler and Sierra, 2017). So, the mean transit time calculation is only limited by the ~~transfer coefficients~~rates of C transfer, while the mean age of C in the system also depends on its mass. This is why the mean age of C in vegetation is determined by the age of the compartment where the majority of the mass is stored, which in this case is Wood. We further explored this relation in the scatter plot in (Figure ~~A9~~A9), where the distribution of the points below the 1:1

line indicate that the three models have mean ages greater than their mean transit times. This means that they have big masses of old carbon, but they also have highly dynamic compartments through which carbon ~~transists~~ transits very fast.

We also diagnosed the model performance by comparing the predicted ages for the storage compartments (Figure 9) with the mean age of NSC measured in previous empirical studies (Richardson et al., 2013). The mean age of NSCs from red maple cores obtained at Harvard forest, was 7.2 ± 7.7 yr; for the fast cycling storage compartments of the models *Storage: 1* and *Storage: 2*, it was 1.71 and 2.14 yr respectively, and 45 yr for the slow cycling storage. Clearly, the mean ages of fast storage compartments are smaller than the mean value calculated empirically, but given the uncertainty of the measurements, they are still ~~under~~ within the observed range. Furthermore, the density distributions of these storage compartments show that even though they are well-mixed, their C do not have the same age. Thus, it is also likely to find C between 0-5 yr and 0-150 yr in the fast and slow cycling storage compartments, respectively. This resonates with the fact that although stemwood NSC is highly dynamic on seasonal time-scales, it can also be surprisingly old (Richardson et al., 2013). Then, the two hypothesis regarding C mixing -inward mixing of younger and older C in one compartment (Trumbore et al., 2015), and 2 compartments (young and old) that mix (Richardson et al., 2013)- converge in the concept of age distributions because C is simultaneously been fixed and removed from the compartments at different times, a process that results in C age distributions. We can think about these dynamics in the context of a stochastic process. The total amount of C that enters and leaves each compartment is fixed and given by the deterministic model, but the time that each C particle stays in a compartment varies ~~randomly~~ stochastically within them. So, the age distribution of C particles in each compartment is a mix of new and old carbon, with distribution functions emerging from the deterministic model (equation 2).

Another important observation regarding the mean age predictions is the fact that these calculations were performed under the assumption that the system, in this case the forest, was in steady state. Since the C stocks in Harvard forest continue growing, the calculated mean ages and transit times should be interpreted as predictions of the mean age that the carbon may have in this forest once it is in steady state. Based on this, the time that this forest will take to reach the steady state is highly divergent among the three models. As an example, the model with two storage compartments would predict 20 yr more of growth to reach a steady-state ~~than~~ close to that of the model with no compartments. ~~If the environmental forcing was included in the simulations~~

In case of systems that are driven by environmental variables and result in time dependencies of inputs (GPP) and process rates, the mean ages and transit ~~times would then have to be calculated using another set of equations (e.g., Rasmussen et al., 2016). However, to use such equations we~~ time distributions would also change over time. To calculate the means of these time-dependent distributions one would need to know the complete history of inputs and cycling rates for the duration of the simulation ~~, and that information was~~ (Rasmussen et al., 2016), information that is not available for Harvard forest. ~~In any case,~~ Nonetheless, we can predict that if there were external factors influencing the simulations, ~~the predicted~~ there would be a different prediction of mean ages and transit times ~~would be different in every time step~~at each time step, but the model structure (C allocation scheme in particular) would play a major role determining the shape of these distributions .

It is also noteworthy that what we assume to be a compartment, e.g. Wood, does not necessarily meet the well mixed assumption, so its particles may not have the same probability to leave the compartment at all times. Richardson et al. (2015)

found a low concentration of old NSC in old rings of stemwood, and a high concentration of old NSC in coarse roots and fine roots of pine. Additionally, they found young and old C in roots. These dynamics were interpreted as poor C mixing and reserves that ~~turn over~~ cycle on different timescales (Richardson et al., 2015), but they may also obey to physiological constraints~~; for~~. For example, the ~~parenquima~~ parenchyma in heartwood is thought to be dead, so NSC trapped in there may no longer be accessible to the plant (Richardson et al., 2015)~~, but it has some probability of being respired and lost from the system~~. So, to study the physiological significance of these findings with models, we might have to include such details regarding tree physiology. The important point we want to emphasize however, is that mixing of carbon in different vegetation compartments results in C age distributions that have been little studied previously. Our results are a first attempt to obtain these distributions using a number of assumptions, but ~~more detailed models relaxing some of these assumptions would be needed to obtain more accurate C age distributions~~in the future other analysis with more complex models and explicit formulas for time-dependent age distributions would help to obtain better predictions of ages and transit times as affected by specific physiological processes.

Although there are still knowledge gaps regarding plant physiology, and current C-dating methods only measure mean age of C rather than age distributions (Richardson et al., 2013), we expect our results to motivate future work, particularly in the use of isotope tracers and their time evolution to approximate age distributions. Biosphere models can be enhanced with structural adjustments and the uncertainties in the parameter values can be reduced by constraining them with age and transit time distributions. These improved models could then be used to test ~~hypothesis~~ hypotheses regarding physiological questions, and assess the sustainability of current terrestrial C sinks, given changes in environmental forcing.

4.2 Model equifinality (identifiability)

Model equifinality (Medlyn et al., 2005) was evident from the fact that despite having different number of compartments and values of cycling rates, all of the three models had similar simulations of C stocks ~~that fitted the data well~~ (Figures 3 and 4). Along with model equifinality, we obtained a high collinearity between some parameters, implying that they are non-identifiable, i.e., they cannot be uniquely estimated from the given data sets (Soetaert and Petzoldt, 2010). Thus, for these particular models, the time-course measurements from only 2 out of 4-6 vegetation compartments is not sufficient to estimate the values of 7-10 parameters.

Model equifinality as well as the impossibility to uniquely identify certain parameters (parameter non-identifiability) ~~may be the result of the~~ is expressed as high correlations between the ~~parameters~~parameter sets. Positive parameter correlations, may indicate ~~a~~ *practical non-identifiability*, where the insufficiency or poor quality of data is not a good constrain for the parameters. In addition, a negative parameter correlation can be a symptom of *structural non-identifiability*, which is the result of a redundant parameterization (Raue et al., 2009; Timmer, 2011; Raue et al., 2012; Cressie et al., 2009). Thus, the three models had practical and structural non-identifiabilities, which means that they need to be constrained with more and better data, and they need to be restructured in order to avoid compensation of fluxes into and out of the compartments.

Since this study was limited by the availability of relevant empirical data, the parameter values that we used are only one of many possible outcomes of parameter estimations using the same data sets. Therefore, it is possible that none of these models accurately depict the C cycle in the Harvard Forest. However, these problems experienced with parameter non-identifiability are not an isolated case; the process of finding unknown rates of C sequestration by fitting biosphere models to empirical data (Luo et al., 2003) is often hampered by ~~the~~ parameter non-identifiability (Schaber and Klipp, 2011). This is a real problem because parameters such as those that correspond to carbon turnover explain most of the variation in the response of terrestrial vegetation to future climate and CO_2 (Friend et al., 2014) and are highly important in determining C age and transit times. ~~Nevertheless, we succeeded in assessing the influence of model structure in predicted ecosystem processes.~~

5 Conclusions

We obtained age and transit times distributions of carbon for simple vegetation models with contrasting carbon allocation schemes. Our results show that mixing of carbon in different vegetation compartments results in C age distributions not explored before in previous studies. The shape of these distributions depends largely on model structure, and in particular on how carbon allocation is represented in models. ~~While the inclusion of a carbon storage compartment had a small impact on predicting C stocks in vegetation compartments, it resulted in very different predictions of age, transit time distributions, C release, and isotopic composition.~~

Models with none or one storage compartment may fail to explain the mixing of ages found in different vegetation compartments, but they are more parsimonious than the model with 2 storage compartments. Nonetheless, parameter collinearity and model equifinality were persistent problems that might be solved if more constraints are added, since the time series of C in foliage and wood are not enough to parameterize a full vegetation model.

~~Our results suggest that~~ Although all models predicted similar C stocks in vegetation compartments, the inclusion of a carbon storage compartment resulted in very different predictions of age, transit time distributions, C release, and isotopic composition. Thus C ages and transit times, which can be indirectly measured using isotope tracers, can be used to improve biosphere models via examination of their structure and estimation of parameter values, which then can be used to assess the strength of C sources or sinks from vegetation.

Finally, it is advantageous to consider age and transit times as distributions, rather than only mean values; with their distributions we obtain additional insights on the temporal dynamics of carbon use, storage, and allocation, which not only depends on the rate at which C flows into and out of the compartments, but also on the stochastic nature of the process itself.

Appendix A: Appendix

Competing interests. The authors declare that they have no conflict of interest.

Table A1. Number of positive and negative correlations between parameters. Only R^2 values < -0.1 and > 0.1 were assumed to account for correlations.

Model	Positive correlations	Negative correlations	Possible combinations
Storage: 0	5	12	21
Storage: 1	4	10	28
Storage: 2	16	16	45

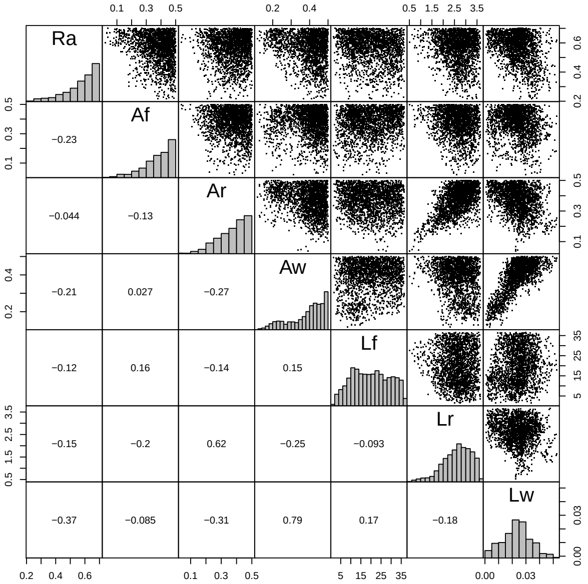


Figure A1. Pairwise plots of sensitivity functions for the model *Storage: 0*

Acknowledgements. Research at Harvard Forest is supported by the National Science Foundation's LTER program (DEB-1237491). This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research. Part of this work was the result of a research visit to the Terrestrial Ecosystems and Global Change group, Department of Organismic and Evolutionary Biology, Harvard University. It was funded by the Max Planck Society and the German Research Foundation through its Emmy Noether Program (SI 1953/2–1).

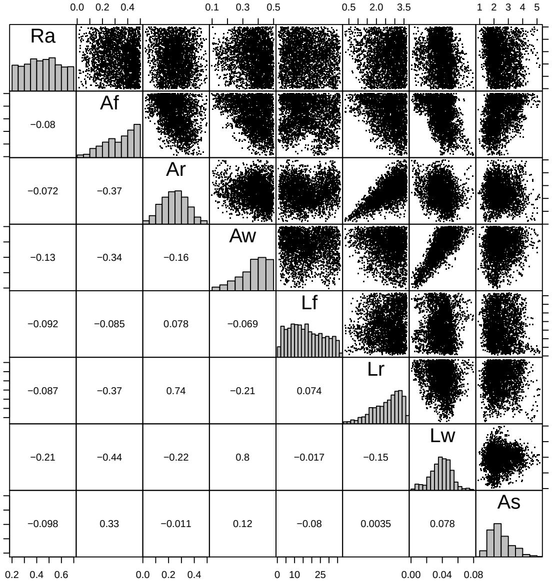


Figure A2. Pairwise plots of sensitivity functions for the model *Storage: 1*

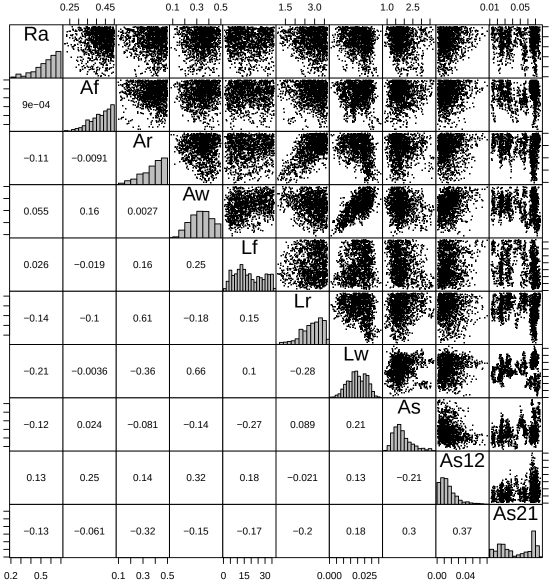


Figure A3. Pairwise plots of sensitivity functions for the model *Storage: 2*

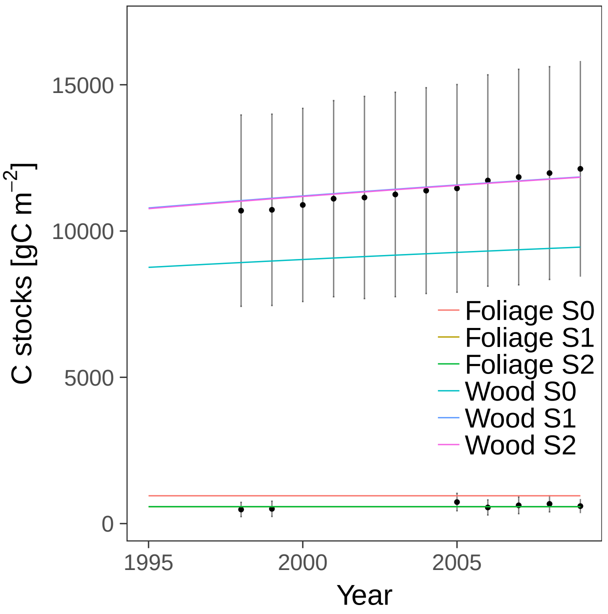


Figure A4. Carbon stocks estimated for each model. The three models were run using the same parameter set.

References

Bolin, B. and Rodhe, H.: A note on the concepts of age distribution and transit time in natural reservoirs, Tellus, 25, 58–62, 1973.

Canadell, J. G., Le Quéré, C., Raupach, M. R., Field, C. B., Buitenhuis, E. T., Ciais, P., Conway, T. J., Gillett, N. P., Houghton, R. A., and Marland, G.: Contributions to accelerating atmospheric CO_2 growth from economic activity, carbon intensity, and efficiency of natural sinks, Proceedings of the National Academy of Sciences, 104, 18 866–18 870, 2007.

Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. V., and Wikle, C. K.: Accounting for Uncertainty in Ecological Analysis: the Strengths and Limitations of Hierarchical Statistical Modeling, Ecological Applications, 19, 553–570, 2009.

Ericsson, T., Rytter, L., and Vapaavuori, E.: Physiology of carbon allocation in trees, Biomass and Bioenergy, 11, 115–127, doi:http://dx.doi.org/10.1016/0961-9534(96)00032-3, 1996.

Fox, A., Williams, M., Richardson, A. D., Cameron, D., Gove, J. H., Quaife, T., Ricciuto, D., Reichstein, M., Tomelleri, E., Trudinger, C. M., and Wijk, M. T. V.: The REFLEX project: Comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data, Agricultural and Forest Meteorology, 149, 1597 – 1615, doi:http://dx.doi.org/10.1016/j.agrformet.2009.05.002, 2009.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., Von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K. G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., and Zeng, N.: Climate-Carbon Cycle Feedback Analysis: Results from the C^{4}MIP Model Intercomparison, Journal of Climate, 19, 3337–3353, 2006.

Friend, A. D., Lucht, W., Rademacher, T. T., Keribin, R., Betts, R., Cadule, P., Ciais, P., Clark, D. B., Dankers, R., Falloon, P. D., Ito, A., Kahana, R., Kleidon, A., Lomas, M. R., Nishina, K., Ostberg, S., Pavlick, R., Peylin, P., Schaphoff, S., Vuichard, N., Warsza-

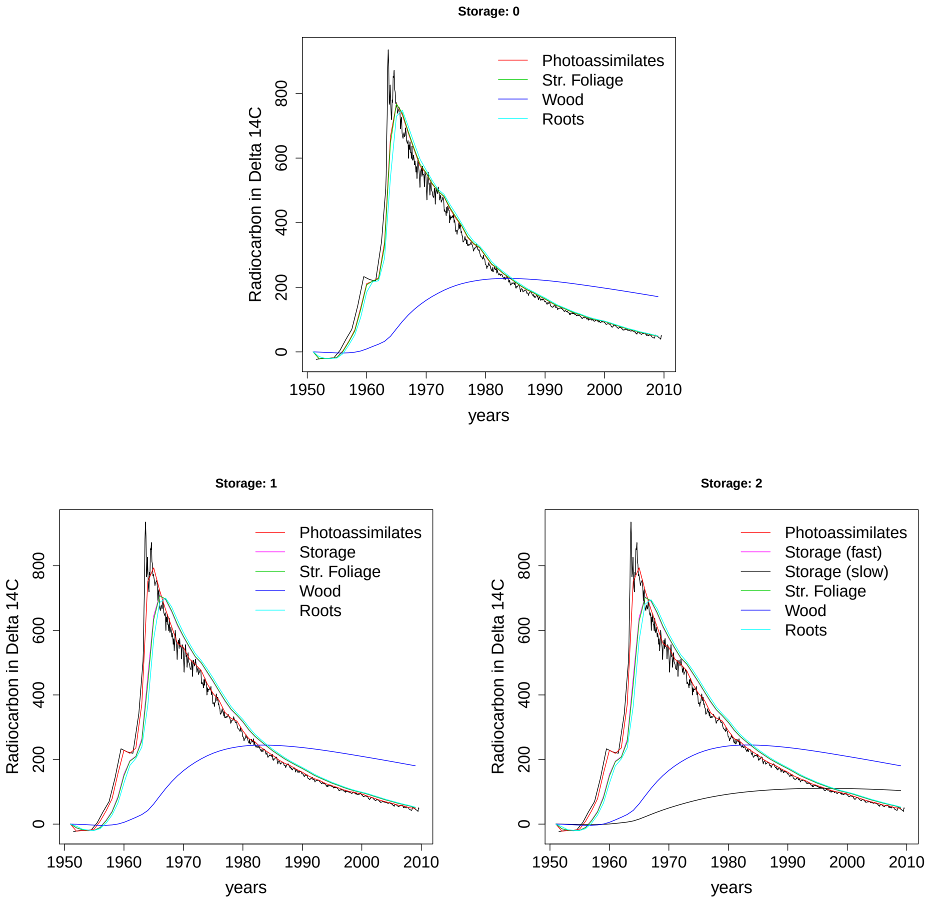


Figure A5. Radiocarbon simulations for the three models. The three models were run using the same parameter set. The black curve corresponds to the Δ^{14}C accumulation in the atmosphere, and the other colors depict the vegetation compartments.

wski, L., Wiltshire, A., and Woodward, F. I.: Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO$_2$., Proceedings of the National Academy of Sciences of the United States of America, 111, 3280–5, doi:10.1073/pnas.1222477110, 2014.

Grulke, N. E., Andersen, C. P., and Hogsett, W. E.: Seasonal changes in above- and belowground carbohydrate concentrations of ponderosa pine along a pollution gradient, Tree Physiology, 21, 173, doi:10.1093/treephys/21.2-3.173, 2001.

Hartmann, H. and Trumbore, S.: Understanding the roles of nonstructural carbohydrates in forest trees -from what we can measure to what we want to know, New Phytologist, 211, 386–403, doi:10.1111/nph.13955, 2016.

Hartmann, H., Ziegler, W., and Trumbore, S. E.: Lethal drought leads to reduction in nonstructural carbohydrates in Norway spruce tree roots but not in the canopy, Functional Ecology, 27, 413–427, doi:10.1111/1365-2435.12046, 2013.

Hoch, G. and Körner, C.: The carbon charging of pines at the climatic treeline: a global comparison, Oecologia, 135, 10–21, doi:10.1007/s00442-002-1154-7, 2003.

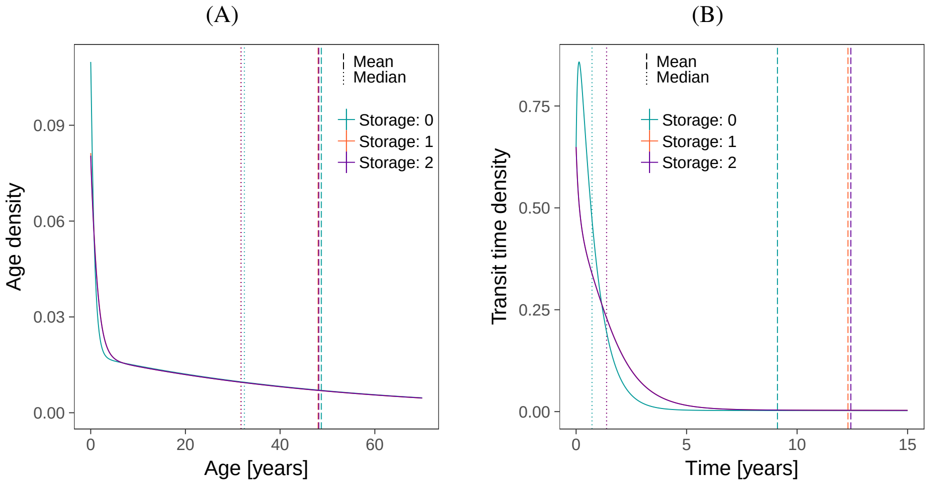


Figure A6. System ages and transit times. (A) Age and transit time (B) distributions calculated for each of the ~~the~~ three models using the same parameter set; the dashed and dotted lines mark the mean and median ages, respectively. Some of the lines may be overlapping others.

Keenan, T. F., Davidson, E. A., Munger, J. W., and Richardson, A. D.: Rate my data: quantifying the value of ecological data for the development of models of the terrestrial carbon cycle, Ecological Applications, 23, 273–286, doi:10.1890/12-0747.1, 2013.

Körner, C.: A matter of tree longevity, Science, 355, 130–131, doi:10.1126/science.aal2449, 2017.

Lacointe, A.: Carbon allocation among tree organs: A review of basic processes and representation in functional-structural tree models, Annals of Forest Science, 57, 521–533, doi:10.1051/forest:2000139, 2000.

Luo, Y., White, L. W., Canadell, J. G., DeLucia, E. H., Ellsworth, D. S., Finzi, A., Lichter, J., and Schlesinger, W. H.: Sustainability of terrestrial carbon sequestration: A case study in Duke Forest with inversion approach, Global Biogeochemical Cycles, 17, n/a–n/a, doi:10.1029/2002GB001923, 2003.

Luo, Y., Shi, Z., Lu, X., Xia, J., Liang, J., Jiang, J., Wang, Y., Smith, M. J., Jiang, L., Ahlström, A., Chen, B., Hararuk, O., Hastings, A., Hoffman, F., Medlyn, B., Niu, S., Rasmussen, M., Todd-Brown, K., and Wang, Y.-P.: Transient dynamics of terrestrial carbon storage: mathematical foundation and its applications, Biogeosciences, 14, 145–161, doi:10.5194/bg-14-145-2017, 2017.

Martínez-Vilalta, J., Sala, A., Asensio, D., Galiano, L., Hoch, G., Palacio, S., Piper, F. I., and Lloret, F.: Dynamics of non-structural carbohydrates in terrestrial plants: a global synthesis, Ecological Monographs, 86, 495–516, doi:10.1002/ecm.1231, 2016.

Medlyn, B. E., Robinson, A. P., Clement, R., and McMurtrie, R. E.: On the validation of models of forest CO2 exchange using eddy covariance data: some perils and pitfalls, Tree Physiology, 25, 839–857, doi:10.1093/treephys/25.7.839, 2005.

Metzler, H. and Sierra, C. A.: Linear Autonomous Compartmental Models as Continuous-Time Markov Chains: Transit-Time and Age Distributions, Mathematical Geosciences, doi:10.1007/s11004-017-9690-1, 2017.

Muhr, J., Messier, C., Delagrange, S., Trumbore, S., Xu, X., and Hartmann, H.: How fresh is maple syrup? Sugar maple trees mobilize carbon stored several years previously during early springtime sap-ascent, New Phytologist, 209, 1410–1416, doi:10.1111/nph.13782, 2016.

Rasmussen, M., Hastings, A., Smith, M. J., Agusto, F. B., Chen-Charpentier, B. M., Hoffman, F. M., Jiang, J., Todd-Brown, K. E. O., Wang, Y., Wang, Y.-P., and Luo, Y.: Transit times and mean ages for nonautonomous and autonomous compartmental systems, J. Math. Biol., 73, 1379–1398, doi:10.1007/s00285-016-0990-8, 2016.

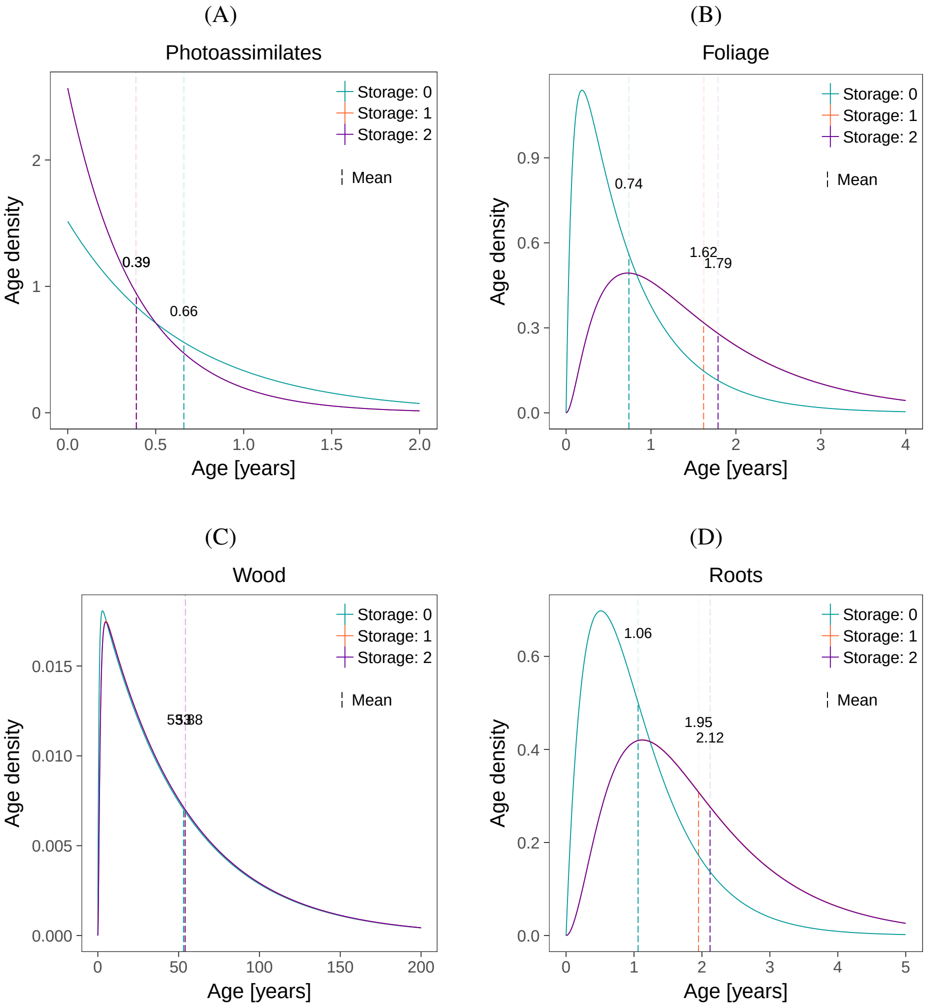
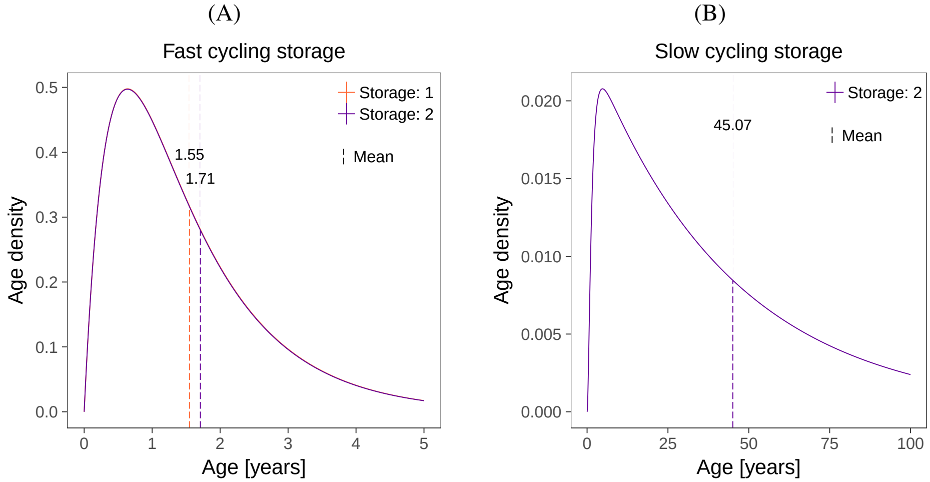


Figure A7. Age densities simulated for the compartments: (A) Photoassimilates, (B) Str. Foliage, (C) Wood and (D) Roots. The three models were run using the same parameter set. Each model is depicted in a different color. The dashed lines correspond to the mean ages of each model for each compartment.

Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., and Timmer, J.: Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood, Bioinformatics, 25, 1923, doi:10.1093/bioinformatics/btp358, 2009.

Raue, A., Kreutz, C., Theis, F. J., and Timmer, J.: Joining forces of Bayesian and frequentist methodology: a study for inference in the presence of non-identifiability, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 371, doi:10.1098/rsta.2011.0544, 2012.


(A)

(B)

Figure A8. Age densities simulated for the models with storage compartments; the models were run using the same parameter set. (A) Fast cycling compartment of models *Storage: 1* and *Storage: 2*. (B) Slow cycling storage of the only model with 2 storage compartments.

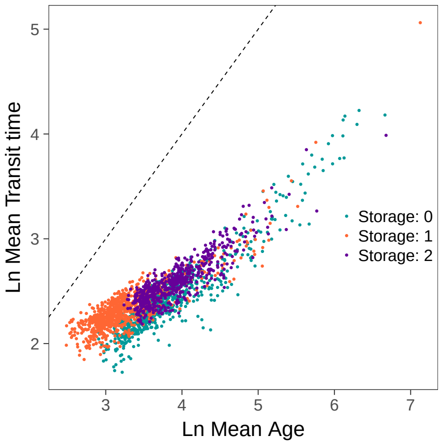


Figure A9. Scatter plot of mean age Vs. mean transit times in log scale. The three models have distributions below the 1:1 line.

Richardson, A. D., Williams, M., Hollinger, D. Y., Moore, D. J. P., Dail, D. B., Davidson, E. A., Scott, N. A., Evans, R. S., Hughes, H., Lee, J. T., Rodrigues, C., and Savage, K.: Estimating parameters of a forest ecosystem C model with measurements of stocks and fluxes as joint constraints, Oecologia, 164, 25–40, doi:10.1007/s00442-010-1628-y, 2010.

Richardson, A. D., Carbone, M. S., Keenan, T. F., Czimczik, C. I., Hollinger, D. Y., Murakami, P., Schaberg, P. G., and Xu, X.: Seasonal dynamics and age of stemwood nonstructural carbohydrates in temperate forest trees, New Phytologist, 197, 850–861, doi:10.1111/nph.12042, 2013.

Richardson, A. D., Carbone, M. S., Huggett, B. A., Furze, M. E., Czimczik, C. I., Walker, J. C., Xu, X., Schaberg, P. G., and Murakami, P.: Distribution and mixing of old and new nonstructural carbon in two temperate trees, New Phytologist, 206, 590–597, doi:10.1111/nph.13273, 2015.

Schaber, J. and Klipp, E.: Model-based inference of biochemical parameters and dynamic properties of microbial signal transduction networks, Current Opinion in Biotechnology, 22, 109–116, doi:10.1016/j.copbio.2010.09.014, 2011.

Schiestl-Aalto, P., Kulmala, L., Mäkinen, H., Nikinmaa, E., and Mäkelä, A.: CASSIA – a dynamic model for predicting intra-annual sink demand and interannual growth variation in Scots pine, New Phytologist, 206, 647–659, doi:10.1111/nph.13275, 2015.

Sierra, C. A., Müller, M., and Trumbore, S. E.: Models of soil organic matter decomposition: The SoilR package, version 1.0, Geoscientific Model Development, 5, 1045–1060, doi:10.5194/gmd-5-1045-2012, 2012.

Sierra, C. A., Müller, M., Metzler, H., Manzoni, S., and Trumbore, S. E.: The muddle of ages, turnover, transit, and residence times in the carbon cycle, Global Change Biology, in print, doi:10.1111/gcb.13556, 2016.

Smallman, T. L., Exbrayat, J. F., Mencuccini, M., Bloom, A. A., and Williams, M.: Assimilation of repeated woody biomass observations constrains decadal ecosystem carbon cycle uncertainty in aggrading forests, Journal of Geophysical Research: Biogeosciences, doi:10.1002/2016JG003520, 2017.

Soetaert, K. and Petzoldt, T.: Inverse Modelling, Sensitivity and Monte Carlo Analysis in R Using Package FME, Journal of Statistical Software, 33, doi:10.18637/jss.v033.i03, 2010.

Timmer, J.: Addressing parameter identifiability by model-based experimentation, IET Systems Biology, 5, 120–130(10), 2011.

Trumbore, S., Gaudinski, J. B., Hanson, P. J., and Southon, J. R.: Quantifying ecosystem-atmosphere carbon exchange with a 14C label, Eos, Transactions American Geophysical Union, 83, 265–268, doi:10.1029/2002EO000187, 2002.

Trumbore, S., Czimczik, C. I., Sierra, C. A., Muhr, J., and Xu, X.: Non-structural carbon dynamics and allocation relate to growth rate and leaf habit in California oaks, Tree Physiology, 35, 1206, doi:10.1093/treephys/tpv097, 2015.

Trumbore, S. E., Sierra, C. A., and Hicks Pries, C. E.: Radiocarbon and Climate Change: Mechanisms, Applications and Laboratory Techniques, chap. Radiocarbon Nomenclature, Theory, Models, and Interpretation: Measuring Age, Determining Cycling Rates, and Tracing Source Pools, pp. 45–82, Springer International Publishing, doi:10.1007/978-3-319-25643-6_3, 2016.

Urbanski, S., Barford, C., Wofsy, S., Kucharik, C., Pyle, E., Budney, J., McKain, K., Fitzjarrald, D., Czikowsky, M., and Munger, J. W.: Factors controlling CO2 exchange on timescales from hourly to decadal at Harvard Forest, Journal of Geophysical Research: Biogeosciences (2005–2012), 112, doi:10.1029/2006JG000293, 2007.

Wofsy, S., Goulden, M., Munger, J., Fan, S.-M., Bakwin, P., Daube, B., Bassow, S., and Bazzaz, F.: Net exchange of CO2 in a mid-latitude forest, Science, 260, 1314–1317, 1993.

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, doi:10.1111/gcb.12172, 2013.

Yizhao, C., Jianyang, X., Zhengguo, S., Jianlong, L., Yiqi, L., Chengcheng, G., and Zhaoqi, W.: The role of residence time in diagnostic models of global carbon storage capacity: model decomposition based on a traceable scheme, Scientific Reports, 5, doi:10.1038/srep16155, 2015.