Dear Dr. Alexey V. Eliseev:

We greatly appreciate your granting us the opportunity to revise our manuscript. We have carefully studied the comments from two referees and M. Freilich. We have revised the manuscript accordingly.

In addition to addressing those comments, we also have made the following changes.

1. We have slightly modified the title to be “Transient Dynamics of Terrestrial Carbon Storage: Mathematical foundation and its applications”. The original term “numeric examples” we think carries less meaning for readers than the term “its applications.”
2. We have deleted the paragraph and relevant sentences about the carbon storage potential to be used as the target quantity of research, trading, and government negotiation. We plan to explain this point in another application paper instead of in this paper.
3. We have also carefully read the manuscript and made many minor changes in the whole manuscript to improving its readability and accuracy of expression.

Hope you will find our revision and responses satisfactory.

Yiqi Luo
On behalf of all the authors
Dear Referee 1:

We greatly appreciate your comments on our manuscript. We have carefully studied your comments and revised the manuscript accordingly.

Hope you will find our revision and responses satisfactory.

Yiqi Luo
On behalf of all the authors

Below we list our point-to-point responses to your comments:

[Comment] General remarks:
The authors present a paper showing that a matrix equation can replicate the output of a comprehensive carbon cycle model. In particular they find that the force driving the ecosystem C storage is the C storage capacity. In general the article is well written and organized and fits into the scope of the journal. Using such a simple matrix equation as a physical emulator of comprehensive models has the potential to save a lot of computing time and gains a deeper understanding of the underlying mechanisms. The authors state in their summary that this would revolutionize model evaluations.

[Response] Thanks for the positive comment.

[Comment] I have some concern about this: The matrix equation has to be fitted to a simulation of the complex model with a specified fixed climate scenario. It would be interesting to know whether this parameter set can be used for a different climate scenario. In particular some parameters in the matrix equation are time-dependent and this time-dependence might change for different climate scenarios. Then the complex model can really be replaced by the matrix equation. Otherwise the matrix equation allows only a more convenient analysis of the model output. Non-linearities in the complex model might lead to a deviation from the linearized matrix representation. It would be nice if the authors could comment on that.

[Response] The physical emulator does not result from fitting the model to simulation of the complex model. It generates by organizing the carbon balance equations in the original model into a matrix form. So the physical emulator is not climate scenario-specific. Once developed, it is applicable to all climate scenarios.

We have revised the manuscript to clarify this point. For example, we revised the title of section 2.2 to be “TECO Model, its physical emulator, and numerical experiments”. We completely rewrote the third paragraph in that section to describe how we have developed the physical emulator of TECO in detail as:

“To support the mathematical analysis using eq. 1, we first developed a physical emulator (i.e., the matrix representation of eq. 1) of the TECO model and then verified that the physical emulator can exactly represent simulations of the original TECO model. We first identified those parameter values in each of the C balance equations in the TECO model that are corresponding
to elements in matrices $A$ and $K$ in eq. 1. The time-dependent variables for $u(t)$, elements in vector $B$, and elements in matrix $\xi(t)$ in the physical emulator were directly from outputs of the original TECO model. Then those parameter values and time-dependent variables were organized into matrices $A$, $\xi(t)$, and $K$; vectors $X(t)$, $X_0$, and $B$; and variable $u(t)$. Note that values of $u(t)$, $B$, and $\xi(t)$ could be different among different climate scenarios. Those matrices, vectors, and variable were entered to matrix calculation to compute $X'(t)$ using eq. 1. The sum of elements in calculated $X'(t)$ is a 100% match with simulated net ecosystem production (NEP) with the TECO model (Fig. 1b).”

Hope this paragraph explains the physical emulator clearly. In addition, we added section 6 Code availability on page 31 and provided a webpage link to both the TECO model and its physical emulator for verification and uses.

[Comment] More specific remarks:
Abstract: The authors are talking about a 3-D parameter space. These 3 parameters, however, are not simple scalars, but are itself vectors (e.g., residence time and storage potential).

[Response] we add elements of the vectors together to get the scalars before we plotted the 3D parameter space. We clarified this point in several places in the manuscript. For example,

One paragraph on page 14 (lines 303-306) on this point is:

“Note that sums of elements in vectors $X(t)$, $X_c(t)$, $X_p(t)$, and $X'(t)$ are corresponding, respectively, to the whole ecosystem C stock, ecosystem C storage capacity, ecosystem C storage potential, and net ecosystem production (NEP). In this paper, we do not use a separate set of symbols to represent those sums rather than express them wherever necessary.”

Also, the legend of Figure 1 explains this point:

“Panel b compares the original TECO model outputs with those from matrix equations for net ecosystem production (NEP = the sum of elements in $X'(t)$ from eq. 1).” “Panel c compares the original TECO model outputs with those from matrix equations for ecosystem C storage (= the sum of elements in $X(t)$ from eq. 2).”

[Comment] Page 4: The authors state that most carbon cycle models follow a mathematical formulation of ordinary differential equations. Many of the dynamic global vegetation models (DGVM) are ab initio formulated as a time discrete model calculating, e.g., NPP on a daily level and carbon allocation to different vegetation pools on annually using some (non-linear) allocation rules. Moreover, the authors should mention these DGVMs.

[Response] Thanks for the comments. It is not very clear with “are ab initio formulated.” That leaves some uncertainty about our understanding of this comment. Nevertheless, the time steps of NPP calculation and allocation do not affect Eq. 1. Indeed, eq. 1 is mainly about C transformation within land ecosystems before the carbon is resired. NPP is input of eq. 1.

We have successfully applied Eqs. 1 and 2 to LPJ-GUESS, a DGVM, as described in line 581.
[Comment] Page 9: The authors should describe which algorithms are necessary in order to develop the matrix equation from the output of the TECO model. In particular how they determined matrix A and K.

[Response] We wrote the physical emulator of the TECO model in matlab. But it can be developed in any other computer language. Basically, we have to understand the original model and identify those carbon balance equations. Then we organize those coefficients and parameters in matrix forms to develop the physical emulator. See our responses to your comment on emulator above. We have completely revised the paragraph in Section 2.2 to describe how we developed the physical emulator of the TECO model.

We also described the physical emulator in paragraph on pages 26-27. Specifically, lines 585-588 state “the physical emulators differ for different models as the elements of each matrix could be differently parameterized or formulized in different models. Also, different models usually have different pool-flux structures, leading to different non-zero elements in the $A$ matrix.”

[Comment] Technical comment:
Page 29, line 586: A “to” is missing: The emulators allow us TO analyze: : :

[Response] Corrected as suggested.

[Comment] In summary the article is suitable for publication if the above-mentioned comments are incorporated.

[Response] Thank the referee for the support.
Dear Referee 2:

We greatly appreciate your time and effort to read, understand, and make comments on our manuscript. We have carefully studied your comments and revised the manuscript accordingly. Hope our responses have adequately addressed your concerns so that we can develop mutual understanding about your concerns and about what we present in the paper.

Please note the line numbers and pages numbers in this letter are all refereed in the revised, marked-up manuscript.

Yiqi Luo
On behalf of all the authors

Below we list our point-to-point responses to your (i.e., referee 2 in this case) comments:

[Comment] In spite of words “mathematical foundation” in the title, the first mistake is contained directly in the first formula (1). Let’s rewrite it in the component form:

\[
\begin{pmatrix}
X_1' \\
\vdots \\
X_n'
\end{pmatrix} = 
\begin{pmatrix}
B_1 \\
\vdots \\
B_n
\end{pmatrix} u(t) - 
\begin{pmatrix}
A_{11} & \cdots & A_{1n} \\
\vdots & \ddots & \vdots \\
A_{n1} & \cdots & A_{nn}
\end{pmatrix} 
\begin{pmatrix}
\xi_1 \\
\vdots \\
\xi_n
\end{pmatrix} 
\begin{pmatrix}
0 & 0 & 0 & K_1 & 0 & 0 \\
0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & \cdots & 0 & 0 & \cdots & 0
\end{pmatrix} 
\begin{pmatrix}
X_1 \\
\vdots \\
X_n
\end{pmatrix}
\]

and see that in this notation all off-diagonal elements of matrix A are useless, and the system (1) is simply a set of trivial linear equations for disconnected variables. Do the authors know that matrix multiplication is non-commutative? My hypothesis is that the matrix A should be stated after other multipliers in the second member of the sum:

\[X'(t) = Bu(t) - \xi KAX(t)\]

Such a formula is at least mathematically correct and allows the following component view:

\[
\begin{pmatrix}
X_1' \\
\vdots \\
X_n'
\end{pmatrix} = u(t) - 
\begin{pmatrix}
0 & 0 & 0 & K_1 & 0 & 0 \\
0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & \cdots & 0 & 0 & \cdots & 0
\end{pmatrix} 
\begin{pmatrix}
A_{11} & \cdots & A_{1n} \\
\vdots & \ddots & \vdots \\
A_{n1} & \cdots & A_{nn}
\end{pmatrix} 
\begin{pmatrix}
\xi_1 \\
\vdots \\
\xi_n
\end{pmatrix} 
\begin{pmatrix}
X_1 \\
\vdots \\
X_n
\end{pmatrix}
\]

Consequently all next formulas should be corrected according to the new form of (1). It’s completely unclear why “all off-diagonal values \(a_{ji}\) are negative” (page 8).

[Response] We are grateful to you for your time and effort to examine mathematical formulas. We agree with you that it is critical to make sure the mathematical expression of biological processes should be correct before we do any analysis.
and the carbon dynamics in pool 1 will be described by:

\[
\begin{bmatrix}
X'_1 \\
X'_2 \\
\vdots \\
X'_n
\end{bmatrix} =
\begin{bmatrix}
B_1 \\
B_2 \\
\vdots \\
B_n
\end{bmatrix} u(t) -
\begin{bmatrix}
A_{11} \xi_1 K_1 & A_{12} \xi_2 K_2 & \cdots & A_{1n} \xi_n K_n \\
A_{21} \xi_1 K_1 & A_{22} \xi_2 K_2 & \cdots & A_{2n} \xi_n K_n \\
\vdots & \vdots & \ddots & \vdots \\
A_{n1} \xi_1 K_1 & A_{n2} \xi_2 K_2 & \cdots & A_{nn} \xi_n K_n
\end{bmatrix}
\begin{bmatrix}
X_1 \\
X_2 \\
\vdots \\
X_n
\end{bmatrix}
\]

Then the carbon dynamics in pool 1 will be described by:

\[
X'_1 = B_1 u(t) - (A_{11} \xi_1 K_1 X_1 + A_{12} \xi_2 K_2 X_2 + \cdots + A_{1n} \xi_n K_n X_n)
\]

The above equation states that change in carbon content in pool 1 equals carbon influx from a fraction of NPP (i.e., \(u(t)\) times partitioning coefficient \(B_1\)) minus decomposition expressed by \((A_{11} \xi_1 K_1 X_1 + A_{12} \xi_2 K_2 X_2 + \cdots + A_{1n} \xi_n K_n X_n)\). Since \(K\) is decomposition coefficient, the term \(K_1 X_1\) describes that decomposition of carbon in pool 1 equals \(K_1\) times \(X_1\), so on for \(K_2 X_2\), and \(K_n X_n\). Environmental scalar \(\xi_i\) modifies its corresponding \(K_i\). Transfer coefficient \(A_{ij}\) in the above equation describes carbon transfer from pool \(j\) to pool 1. In the real world, no carbon is transferred from other plant, litter, and soil pools to leaf pool. Thus \(A_{1j} = 0, j \neq 1\).

However, not all \(A_{ij} = 0, j \neq i\). In TECO model with carbon transfer pathways as depicted in Figure 1a, there are many zero but several non-zero elements in matrix \(A\) to represent carbon transfers among pools as:

\[
A =
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
-\alpha_{41} & -\alpha_{42} & -\alpha_{43} & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & -\alpha_{53} & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & -\alpha_{64} & -\alpha_{65} & 1 & -\alpha_{67} & -\alpha_{68} \\
0 & 0 & 0 & 0 & -\alpha_{75} & -\alpha_{76} & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & -\alpha_{86} & -\alpha_{87} & 1
\end{pmatrix}
\]

For example, \(A_{41} \neq 0\) as it represents litterfall from leaf pool to litter pool. Some of those none-zero transfer coefficients as represented by \(A_{ij}\) are related to microbial carbon use efficiency.

Let us look at the equation you suggested (i.e., Equation 1-a). After the multiplication of \(\xi, K\) and \(A\), Equation 1-a will become:

\[
\begin{bmatrix}
X'_1 \\
X'_2 \\
\vdots \\
X'_n
\end{bmatrix} = u(t)
\begin{bmatrix}
B_1 \\
B_2 \\
\vdots \\
B_n
\end{bmatrix} -
\begin{bmatrix}
A_{11} \xi_1 K_1 & A_{12} \xi_1 K_1 & \cdots & A_{1n} \xi_1 K_1 \\
A_{21} \xi_2 K_2 & A_{22} \xi_2 K_2 & \cdots & A_{2n} \xi_2 K_2 \\
\vdots & \vdots & \ddots & \vdots \\
A_{n1} \xi_n K_n & A_{n2} \xi_n K_n & \cdots & A_{nn} \xi_n K_n
\end{bmatrix}
\begin{bmatrix}
X_1 \\
X_2 \\
\vdots \\
X_n
\end{bmatrix}
\]

and the carbon dynamics in pool 1 will be described by:
\[ X'_1 = B_1 u(t) - (A_{11}\xi_1 K_1 X_1 + A_{12}\xi_1 K_1 X_2 + \cdots + A_{1n}\xi_1 K_1 X_n) \]

In the above equation, the term \( K_1 X_2 \) or \( K_1 X_n \) biologically does not make sense as it describes the decomposition of carbon in pool 2 by coefficient \( K_1 \). The latter describes the relative rate of decomposition of pool 1. Nor does \( K_1 X_n \) biologically make sense.

Thus, we hope that you can see that our original equation still works.

The statement “all off-diagonal values \( a_{ji} \) are negative” (page 9) because carbon transfer from pool \( i \) to pool \( j \) to be positive by having negative coefficient multiplied with negative sign for this term. We have clarified this point by revising the sentences on line 175-176 as:

“In eq. 1, all the off-diagonal \( a_{ji} \) values are negative to reverse the minus sign and indicate positive C influx to the receiving pools”

[Comment] But the more essential question is concerned to it’s biological correctness and sense. According to (1, 1-a) matrix \( A \) consists of transfer coefficients and does not depend on system variables \( X \) making all the system non-autonomous and linear. There is no biological foundation for such strong universality of the form (1, 1-a) for all temporal and spatial scales and no mathematical proof in the paper. In particular, it’s not clear how mass-balance relations are connected with that form.

[Response] Thanks for your question about the biological basis of the mathematical equation. The two paragraphs from line 143 to 161 describe the biological basis as below:

“Hundreds of models have been developed to simulate terrestrial C cycle (Manzoni and Porporato, 2009). All the models have to simulate processes of photosynthetic C input, C allocation and transformation, and respiratory C loss. It is well understood that photosynthesis is a primary pathway of C flow into land ecosystems. Photosynthetic C input is usually simulated according to carboxylation and electron transport rates (Farquhar et al., 1980). Ecosystem C influx varies with time and space mainly due to variations in leaf photosynthetic capacity, leaf area index of canopy, and a suite of environmental factors such as temperature, radiation, and relative humidity (or other water-related variables) (Potter et al., 1993; Sellers et al., 1996; Keenan et al., 2012; Walker et al., 2014, Parolari and Porporato 2016).

Photosynthetically assimilated C is partly used for plant biomass growth and partly released back into the atmosphere through plant respiration. Plant biomass in leaves and fine roots usually lives for several months up to a few years before death, while woody tissues may persist for hundreds of years in forests. Dead plant materials are transferred to litter pools and decomposed by microorganisms to be partially released through heterotrophic respiration and partially stabilized to form soil organic matter (SOM). SOM can store C in the soil for hundreds or thousands of years before it is broken down to CO\(_2\) through microbial respiration (Luo and Zhou, 2006). This series of C cycle processes has been represented in most ecosystem models with multiple pools linked by C transfers among them (Jenkinson et al., 1987; Parton et al., 1987; 1988; 1993), including those embedded in Earth system models (Ciais et al., 2013). “

Moreover, we have conducted many synthesis studies to examine different aspects of the
biological basis. The carbon input via canopy photosynthesis as described by
\[
\begin{bmatrix}
B_1 \\
B_2 \\
\vdots \\
B_n
\end{bmatrix}
\] 
has been well accepted. Scientists in the community have questioned carbon transformation through \( A \xi K_X \) in equation 1. We examine six assumptions of those carbon cycle models and the validity of our analysis in section 4.1 on pages 19-23. We would be happy to answer any specific questions you would have regarding those assumptions.

[Comment] Page 9 gives us an example of a risky statements made in the paper. Authors say that almost all world models of carbon cycle in terrestrial ecosystems have the form (1). They refer to the work (Manzoni, Porporato, 2009) and state that there is a review of 250 models of carbon cycling in it! First, Table A2 in this work has 200 references to papers describing different versions of a smaller number of models. Second, I have a very strong doubt that all of them can be presented in the form (1) because they were made for various time scales, different set of compartments and different details of biogeochemical processes accounted for. Interesting is the fact that the model of Manzoni and Porporato (2009) themselves is nonlinear and does not look like the system (1)! As well as another model of soil organic carbon and microbial dynamics made by Hararuk et al. (2015) also referred to by the authors!

[Response] Thanks for your comment. We agree with you that the nonlinear microbial models by Manzoni and Porporato (2009) or Hararuk et al. (2015) could not be represented by equation 1. This issue is pointed out in section 4.1 regarding those microbial models (i.e., assumption 1). We also pointed out that thousands of datasets we have reviewed do not seem to support those nonlinear microbial models as described on pages 19-20. Paper by Sierra and Müller (2015) also stated that most of the land carbon cycle models can be represented by equation 1.

Indeed, we have worked with many modeling groups and organized those carbon balance equations in their models into the matrix equations. It has been demonstrated that the matrix equation can represent those original models well as described in paragraph from lines 580-592. Please see another publication by Ahlström et al. (2015) with LPJ-GUESS for the application of eq. 1.

[Comment] In part 2.2 (pages 9-11) authors carry out comparison of the TECO terrestrial ecosystem model results and the system (1) calculations. Their statement on a 100% match of NEE calculations for TECO and (1) seem strange. If TECO is independent of the system (1) this is unbelievable result, in the opposite case the comparison has no sense.

[Response] This is the case. We organized those carbon balance equations of TECO into the matrix equation, eq. 1. Dynamics of eq. 1 should be mathematically equivalent to TECO outputs. However, they might be mismatch due to the errors from numerical simulations. Therefore, we run the matrix equation to verify that the physical emulator can exactly represent simulation outputs from the original TECO model. We have done that with CABLE (Xia et al. 2012, 2013), CLM3.5 (Rafiquee et al. 2016), CLM4.5 (Shi et al. in prep.), BEPS (Chen et al. 2016), and LPJ-GUESS (Ahlström et al. 2015). In all the cases, the matrix equation can reproduce simulations of those original models. In this paper, the matrix equation can 100% match NEP simulation, but minor mismatch in estimated C storage.
Introducing two new definitions – the C storage capacity and C storage potential – could be a good idea of this paper if authors would explain their biological interpretation and mathematical correctness. First, we should make correspondence to (1-a) and note that \( \tau_{ch} = (\xi K A)^{-1} \) instead of (3). Second, study of existence for this inverse matrix is needed to state mathematical correctness of these definitions because inverse matrix serves as a foundation for all math terms in the following text. There is no such study in the paper. Another question arises about chasing time \( \tau_{ch} \): why it’s formula \( \tau_{ch} = (\xi K A)^{-1} \) should have physical dimension of time?

There are no explanations in the text.

[Response] The biological interpretation of C storage capacity is given in Abstract (Lines 42-50), Results (lines 271-278), Discussion section 4.2, and Conclusions. For example, sentences on lines 638-641 in the Conclusion section state:

“The capacity, which is the product of C input and residence time, represents their instantaneous responses to a state of external forcing at a given time. Thus, the C storage capacity quantifies the maximum amount of C that an ecosystem can store at the given environmental condition at a point of time.”

Similarly, C storage potential is also biologically explained in Abstract, Results, Discussion, and Conclusions sections. For example, the first paragraph in section 4.3 is:

“The C storage potential represents the internal capability to equilibrate the current C storage with the capacity. Bogochemically, the C storage potential represents re-distribution of net C pool change, \( X'(t) \), of individual pools through a network of pools with different residence times as connected by C transfers from one pool to the others through all the pathways. The potential is conceptually equivalent to the magnitude of disequilibrium as discussed by Luo and Weng (2011).”

Thanks for your comment. We have added the time dimension for chasing time on lines 251-252 as:

“In eq. 2, we name the term \( (A \xi (t) K)^{-1} \) the chasing time, \( \tau_{ch} (t) \), with a time unit used in exit rate \( K \).”

[Comment] All inputs in the model (1) are supposed constant or time-dependent. In particular on page 15 plant photosynthesis is declared only time-dependent. But for some temporal scales (a year, for example) it can essentially dependent on the plant carbon content and in that case the model (1) should have another form (Parolari, Porporato, 2016).

Reference

[Response] We have carefully studied the paper by Parolari and Porporato (2016), particularly
that paragraph on NPP on page 66. That study differentiated the productivity regime into C-limited and N-limited. The C-limit regime accounts for limitation of light, temperature and moisture whereas the N-limited regime accounts for nitrogen limitation. Both of the regimes have been discussed in relation with eq. 1. Please see sentences on lines 147-151 and lines 474-475 for more explanation. In the revised manuscript, we cited the paper and explained those environmental factors as represented by scalars on line 151.

[Comment] Therefore, since all other formulas and descriptions are based on the terms introduced above with mistakes as well as statements made without sufficient biological basis, the conclusion at page 25 (part 4.4, first sentence) about novel approach suggested by the authors to understand, evaluate, diagnose and improve carbon cycle models is represented as inadequate and seems early and premature.

[Response] We hope our responses to your comments above can help us communicate well with you and then gain mutual understanding on what we presented in this paper and what you were concerned.

We thank you for the valuable comments, some of which led us to improve the text and better communicate our points to the reader. We hope our responses above also demonstrate that our formulation did not include mistakes and that the terms we introduced were founded on sound biological principles. Thus, we stand by our conclusion that the presented approach enables one to understand, evaluate, diagnose and improve carbon cycles models.
Dear M. Freilich, maraf@mit.edu:

We greatly appreciate your comments on our manuscript. We have carefully studied your comments and revised the manuscript accordingly. Please note the line numbers and pages numbers in this letter are all refereed in the revised, marked-up manuscript.

Hope you will find our revision and responses satisfactory.

Yiqi Luo
On behalf of all the authors

Below we list our point-to-point responses to your comments:

[Comment] Luo et al provide an excellent mathematical framework for studying the dynamics of the carbon cycle in terrestrial ecosystems. The focus on transient dynamics makes clear which aspects of carbon storage and sequestration are most important to consider in order to understand the functioning of forests are carbon reservoirs. The reduction of the models to a 3D parameter space is seemingly very useful for a mechanistic understanding of the effects of global change on terrestrial carbon storage.

[Response] We greatly appreciate your positive comments.

[Comment] The modeling assumptions could use further clarification. In particular, the assumption that short-term disturbances can be well represented by the matrix equation (assumption 5) and the assertion that this assumption is unlikely to affect the results need further support. Disturbances may be very important for the carbon cycling of terrestrial systems and can affect ecosystem dynamics and carbon cycling for decades, in addition to causing C fluxes that greatly exceed those from annual cycles.

[Response] We agree. Disturbances can substantially affect ecosystem carbon cycling

[Comment] Presumably, disturbance events could be incorporated in the time varying factors u(t) and _ (t). However, there are a number of well-developed non-linear models for pest outbreaks that might violate the assumption that transfer between pools can be represented by a linear model (assumption 1) if outbreaks were to be incorporated into these factors.

[Response] We appreciate for your point that there are many non-linear models for pest outbreaks. Pest outbreaks affect tree mortality, which usually is in proportion to the severity of pest outbreaks. Tree mortality can be non-linearly responding to pest outbreaks as decomposition of soil organic carbon to temperature. Such non-linear responses still do not affect fundamental properties of the carbon cycle as discussed in Assumption 4 on response functions.

[Comment] While one aspect of pest outbreaks is a reduction in GPP or NPP, which may be sufficiently represented by u(t), even a linear approximation of the rapid change in the transfer of biomass between classes cannot be represented by this model without making the matrix A of transfer coefficients also time-dependent. One way this may be overcome is by setting limits on
the timescale of applicability of this mathematical framework, so as to assume that transfer coefficients are not changing. Further, abiotic disturbances such as fire or disturbances that remove carbon from the ecosystem completely such as harvesting would be outside the scope of this model.

[Response] Matrix A can be time-dependent. Equation 1 does not explicitly include abiotic disturbances in influencing carbon cycle. Weng et al. (2012) developed a disturbance regime model that explicitly incorporates disturbances into equation 1 for their influences of terrestrial carbon cycle. This paper focuses on understanding of fundamental properties of equation 1.

To clarify this point, we have revised the second half of the paragraph on Assumption 5 (lines 485-490) as:

“Those disturbance influences can be represented in terrestrial C cycle models through changes in parameter values, environmental scalars, and/or discrete C transfers among pools of eq. 1 (Luo and Weng 2011). While eq. 1 does not explicitly incorporate disturbances for their influences on land C cycle, Weng et al. (2012) developed a disturbance regime model that combines eq. 1 with frequency distributions of disturbance severity and intervals to quantify net biome exchanges.”

[Comment] The authors show that X’(t) in this model is the net ecosystem production (NEP), but non-biotic transformation from organic and inorganic carbon is not included in NEP, nor is transfer between ecosystems. This may just require a clarification of terminology in order to include fire, other abiotic oxidation, and harvesting in the X(t) term of the model.

[Response] Yes, you are very sharp to point out the omission of this analysis. We did not explicitly include disturbances in the analysis but state that disturbances do not alter fundamental properties of the system. As explained above, Weng et al. (2012) developed a model that explicitly combines disturbances with equation 1 to quantify net biome production on lines 488-490.

[Comment] Finally, it may be useful to clarify on what scale the results apply. Based on the assumptions about linear decay smoothing small scale fluctuations and the neglect of lateral C fluxes, it seems important to point out that this is model applies only at the ecosystem scale. The parameters are calibrated based on one grid cell of the TECO model; would the same procedure be expected to scale up to larger spatial scales?

[Response] Thanks for your comment. Equation 1 has been also applied to several global models, such as National Center for Atmosphere Research (NCAR) Community Land Model (CLM) and LPJ-GUESS. See a published paper by Ahlström et al. (2015) for the application of equation to the global model LPJ-GUESS. Fundamentally equation 1 fully represents carbon balance equations in matrix form for almost all the land carbon cycle models. Equation 1 does not do any more smoothing of small-scale fluctuations than do the original models. The paragraph on pages 26-27 about physical emulators explains it.

Yes, equation 1 does not apply to the models with lateral fluxes.
[Comment] In the conclusion, the authors state that this model is consistent with complex dynamics including tipping points, which they say are “caused by multiple environmental forcing variables interacting with relatively simple internal processes over different temporal and spatial scales.” Tipping point behavior crucially depends on non-linear dynamics and so seems inconsistent with this model. However, a clarification that this method can evaluate the transient dynamics in a given state but does not reproduce more complex behavior may be more accurate.

[Response] You are right that the eq. 1 does not cause some of the complex dynamics such as tipping points. Tipping points occur in carbon cycle mainly due to complex behaviors in external forcings. Luo and Weng (2011) and Luo et al. (2015) have explained this phenomenon in detail. While this paper could not explain this in detail again, we revised the manuscript by pointing readers to those papers for detailed discussion as on pages 29-30:

“The two components of land C storage dynamics represent interactions of external forces (via changes in the capacity) and internal capability of the land C cycle (via changes in the C storage potential) to generate complex phenomena of C cycle dynamics, such as fluctuations, directional changes, and tipping points, in the terrestrial ecosystems. From a system perspective, these complex phenomena could not be generated by relatively simple internal processes but are mostly caused by multiple environmental forcing variables interacting with internal processes over different temporal and spatial scales as explained by Luo and Weng (2011) and Luo et al. (2015). Note that while those internal processes can be mathematically represented with a relatively simple formula, their ecological and biological underpinnings can be very complex.”
Transient Dynamics of Terrestrial Carbon Storage: Mathematical foundation and its applications

Yiqi Luo\textsuperscript{1,2}, Zheng Shi\textsuperscript{1}, Xingjie Lu\textsuperscript{3}, Jianyang Xia\textsuperscript{4}, Junyi Liang\textsuperscript{1}, Jiang Jiang\textsuperscript{1}, Ying Wang\textsuperscript{5}, Matthew J. Smith\textsuperscript{6}, Lifen Jiang\textsuperscript{1}, Anders Ahlström\textsuperscript{7,8}, Benito Chen\textsuperscript{9}, Oleksandra Hararuk\textsuperscript{10}, Alan Hastings\textsuperscript{11}, Forrest Hoffman\textsuperscript{12}, Belinda Medlyn\textsuperscript{13}, Shuli Niu\textsuperscript{14}, Martin Rasmussen\textsuperscript{15}, Katherine Todd-Brown\textsuperscript{16}, Ying-Ping Wang\textsuperscript{3}

\textsuperscript{1}Department of Microbiology and Plant Biology, University of Oklahoma, Norman, Oklahoma, USA,  \textsuperscript{2}Center for Earth System Science, Tsinghua University, Beijing, China,  \textsuperscript{3}CSIRO Oceans and Atmosphere, Aspendale, Victoria, Australia,  \textsuperscript{4}School of Ecological and Environmental Sciences, East China Normal University, Shanghai, China,  \textsuperscript{5}Department of Mathematics, University of Oklahoma, Norman, Oklahoma, USA,  \textsuperscript{6}Computational Science Laboratory, Microsoft Research, Cambridge, UK,  \textsuperscript{7}Department of Earth System Science, Stanford University, Stanford, California, USA,  \textsuperscript{8}Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden,  \textsuperscript{9}Department of Mathematics, University of Texas, Arlington, TX, USA,  \textsuperscript{10}Department of Natural Resource Sciences, McGill University, Montreal, Canada,  \textsuperscript{11}Department of Environmental Science and Policy, University of California, One Shields Avenue, Davis, CA 95616, USA,  \textsuperscript{12}Computational Earth Sciences Group, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA,  \textsuperscript{13}Hawkesbury Institute for the Environment, Western Sydney University, Penrith NSW 2751, Australia,  \textsuperscript{14}Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, China,  \textsuperscript{15}Department of Mathematics, Imperial College, London, UK,  \textsuperscript{16}Biological Sciences Division, Pacific Northwest
National Laboratory, Richland, Washington, USA,

Running Title: Land carbon storage dynamics

Correspondence author: Yiqi Luo

Email: yluo@ou.edu

Key words Carbon cycle, carbon sequestration, dynamic disequilibrium, model intercomparison, terrestrial ecosystems, traceability analysis

Type of paper: Primary Research Article
Abstract Terrestrial ecosystems have absorbed roughly 30% of anthropogenic CO₂ emissions over the past decades, but it is unclear whether this carbon (C) sink will endure into the future. Despite extensive modeling, experimental, and observational studies, what fundamentally determines transient dynamics of terrestrial C storage under global change is still not very clear. Here we develop a new framework for understanding transient dynamics of terrestrial C storage through mathematical analysis and numerical experiments. Our analysis indicates that the ultimate force driving ecosystem C storage change is the C storage capacity, which is jointly determined by ecosystem C input (e.g., net primary production, NPP) and residence time. Since both C input and residence time vary with time, the C storage capacity is time-dependent and acts as a moving attractor that actual C storage chases. The rate of change in C storage is proportional to the C storage potential, the difference between the current storage and the storage capacity. The C storage capacity represents instantaneous responses of the land C cycle to external forcing, whereas the C storage potential represents the internal capability of the land C cycle to influence the C change trajectory in the next time step. The influence happens through redistribution of net C pool changes in a network of pools with different residence times.

Moreover, this and our other studies have demonstrated that one matrix equation can exactly replicate simulations of most land C cycle models (i.e., physical emulators). As a result, simulation outputs of those models can be placed into a three-dimensional (3D) parameter space to measure their differences. The latter can be decomposed into traceable components to track the origins of model uncertainty. In addition, the physical emulators make data assimilation computationally feasible so that both C flux- and pool-related datasets can be used to better constrain model predictions of land C sequestration. Overall, this new mathematical framework offers new approaches to understand, evaluate, diagnose, and improve land C cycle models.
1 Introduction

Terrestrial ecosystems have been estimated to sequester approximately 30% of anthropogenic carbon (C) emission in the past three decades (Canadell et al., 2007). Cumulatively, land ecosystems have sequestered more than 160 Gt C from 1750 to 2015 (Le Quéré et al., 2015).

Without land C sequestration, the atmospheric CO₂ concentration would have increased by additional 95 parts per million and result in more climate warming (Le Quéré et al., 2015).

During one decade from 2005 to 2014, terrestrial ecosystems sequestrated 3±0.8 Gt C per year (Le Quéré et al., 2015), which would cost billion dollars if the equivalent amount of C was sequestrated using C capture and storage techniques (Smith et al., 2016). Thus, terrestrial ecosystems effectively mitigate global change through natural processes with minimal cost.

Whether this terrestrial C sequestration would endure into the future, however, is not clear, making the mitigation of global change greatly uncertain. To predict future trajectories of C sequestration in the terrestrial ecosystems, it is essential to understand fundamental mechanisms that drive terrestrial C storage dynamics.

To predict future land C sequestration, the modeling community has developed many C cycle models. According to a review by Manzoni and Porporato (2009), approximately 250 biogeochemical models have been published over a time span of 80 years to describe carbon and nitrogen mineralization. The majority of those 250 models follow some mathematical formulations of ordinary differential equations. Moreover, many of those biogeochemical models incorporate more and more processes in an attempt to simulate C cycle processes as realistically as possible (Oleson et al., 2013). As a consequence, terrestrial C cycle models have become increasingly complicated and less tractable. Almost all model intercomparison projects (MIPs), including those involved in the last three IPCC assessments, indicate that C cycle models have
consistently projected widely spread trajectories of land C sinks and were also found to fit observations poorly (Todd-Brown et al., 2013; Luo et al., 2015). The lack of progress in uncertainty analysis urges us to understand mathematical foundation of those terrestrial C models so as to diagnose causes of model spreads and improve model predictive skills.

Meanwhile, many countries have made great investments on various observational and experimental networks (or platforms) in hope to quantify terrestrial C sequestration. For example, FLUXNET has been established about 20 years ago to quantify net ecosystem exchange (NEE) between the atmosphere and biosphere (Baldocchi et al., 2001). Orbiting Carbon Observatory 2 (OCO-2) satellite was launched in 2014 to quantify carbon dioxide concentrations and distributions in the atmosphere at high spatiotemporal resolution to constrain land surface C sequestration (Hammerling et al., 2012). Networks of global change experiments have been designed to uncover processes that regulate ecosystem C sequestration (Rustad et al., 2001; Luo et al., 2011; Fraser et al., 2013; Borer et al., 2014). Massive data have been generated from those observational systems and experimental networks. They offer an unprecedented opportunity for advancing our understanding of ecosystem processes and constraining model prediction of ecosystem C sequestration. Indeed, many of those networks were initiated with one goal to improve our predictive capability. Yet the massive data have been rarely integrated into earth system models to constrain their predictions. It is a grand challenge in our era to develop innovative approaches to integration of big data into complex models so as to improve prediction of future ecosystem C sequestration.

From a system perspective, ecosystem C sequestration occurs only when the terrestrial C cycle is in a transient state, under which C influx into one ecosystem is larger than C efflux from the ecosystem. Olson (1963) is probably among the first to examine organic matter storage in...
forest floors from the system perspective. His analysis approximated steady-state storage of organic matter as a balance of litter producers and decomposers for different forest types. However, global change differentially influences various C cycle processes in ecosystems and results in transient dynamics of terrestrial C storage (Luo and Weng, 2011). For example, rising atmospheric CO$_2$ concentration primarily stimulates photosynthetic C uptake while climate warming likely enhances decomposition. When ecosystem C uptake increases in a unidirectional trend under elevated [CO$_2$], terrestrial C cycle is at disequilibrium, leading to net C storage. The net gained C is first distributed to different pools, each of which has a different turnover rate (or residence time) before C is eventually released back to the atmosphere via respiration. Distribution of net C exchange to multiple pools with different residence times is an intrinsic property of an ecosystem to gradually equalize C efflux with influx (i.e. internal recovery force toward an attractor). In contrast, global change factors that causes changes in C input and decomposition is considered external forces that create disequilibrium through altering internal C processes and pool sizes. The transient dynamics of terrestrial C cycle at disequilibrium is maintained by interactions of internal processes and external forces (Luo and Weng, 2011). Although the transient dynamics of terrestrial C storage have been conceptually discussed, we still lack a quantitative formulation to estimate transient C storage dynamics in the terrestrial ecosystems.

This paper was designed to address a question: what determines transient dynamics of C storage in terrestrial ecosystems from a system perspective? We first reviewed the major processes that most models have incorporated to simulate terrestrial C sequestration. The review helps establish that terrestrial C cycle can be mathematically represented by a matrix equation. We also described the Terrestrial ECOsystem (TECO) model with its numerical experiments in
support of the mathematical analysis. We then presented results of mathematical analysis on
determinants of the terrestrial C storage, direction and magnitude of C storage at a given time
point, and numerical experiments to illustrate climate impacts on terrestrial C storage. We
carefully discussed assumptions of those terrestrial C cycle models as represented by the matrix
equation, the validity of this analysis, and two new concepts introduced in this study, which are
the C storage capacity and C storage potential. We also discussed the potential applications of
this analysis to model uncertainty analysis and data-model integration. Moreover, we proposed
that the C storage potential be a targeted variable for research, trading, and government
negotiation for C credit.

2 Methods
2.1 Mathematical representation of terrestrial C cycle
This study was conducted mainly with mathematical analysis. We first established the basis of
this analysis, which is that the majority of terrestrial C cycle models can be represented by a
matrix equation.

Hundreds of models have been developed to simulate terrestrial C cycle (Manzoni and
Porporato, 2009). All the models have to simulate processes of photosynthetic C input, C
allocation and transformation, and respiratory C loss. It is well understood that photosynthesis is
a primary pathway of C flow into land ecosystems. Photosynthetic C input is usually simulated
according to carboxylation and electron transport rates (Farquhar et al., 1980). Ecosystem C
influx varies with time and space mainly due to variations in leaf photosynthetic capacity, leaf
area index of canopy, and a suite of environmental factors such as temperature, radiation, and
relative humidity (or other water-related variables) (Potter et al., 1993; Sellers et al., 1996;
Photosynthetically assimilated C is partly used for plant biomass growth and partly released back into the atmosphere through plant respiration. Plant biomass in leaves and fine roots usually lives for several months up to a few years before death, while woody tissues may persist for hundreds of years in forests. Dead plant materials are transferred to litter pools and decomposed by microorganisms to be partially released through heterotrophic respiration and partially stabilized to form soil organic matter (SOM). SOM can store C in the soil for hundreds or thousands of years before it is broken down to CO$_2$ through microbial respiration (Luo and Zhou, 2006). This series of C cycle processes has been represented in most ecosystem models with multiple pools linked by C transfers among them (Jenkinson et al., 1987; Parton et al., 1987; 1988; 1993), including those embedded in Earth system models (Ciais et al., 2013).

The majority of the published 250 terrestrial C cycle models use ordinary differential equations to describe C transformation processes among multiple plant, litter, and soil pools (Manzoni and Porporato, 2009). Those ordinary differential equations can be summarized into a matrix formula (Luo et al., 2001; 2003; Luo and Weng, 2011; Luo et al., 2015; 2016; Sierra and Müller 2015) as:

\[
\begin{align*}
X'(t) &= Bu(t) - A\xi(t)KX(t) \\
X(t = 0) &= X_0
\end{align*}
\]

where $X'(t)$ is a vector of net C pool changes at time $t$, $X(t)$ is a vector of pool sizes, $B$ is a vector of partitioning coefficients from C input to each of the pools, $u(t)$ is C input rate, $A$ is a matrix of transfer coefficients (or microbial C use efficiency) to quantify C movement along the pathways, $K$ is a diagonal matrix of exit rates (mortality for plant pools and decomposition coefficients of litter and soil pools) from donor pools, $\xi(t)$ is a diagonal matrix of environmental scalars to represent responses of C cycle to changes in temperature, moisture, nutrients, litter
quality, and soil texture, and $X_0$ is a vector of initial values of pool sizes of $X$. In eq. 1, all the off-diagonal elements of matrix $A$, $a_{ij}$, are negative to reverse the minus sign and indicate positive C influx to the receiving pools. The equation describes net C pool change, $X'(t)$, as a difference between C input, $u(t)$, distributed to different plant pools via partitioning coefficients, $B$, and C loss through the C transformation matrix, $A\xi(t)K$, among individual pools, $X(t)$.

Elements in vector $B$, matrices $A$ and $K$ could vary with many factors, such as vegetation types, soil texture, microbial attributes, and litter chemistry. For example, vegetation succession may influence elements in vector $B$, matrices $A$ and $K$ in addition to C input, $u(t)$, and forcing that affects C dynamics through environmental scalars, $\xi(t)$.

After synthesis of all the possible soil C cycle models based on six principles (mass balance, substrate dependence of decomposition, heterogeneity of decay rates, internal transformations of organic matter, environmental variability effects, and substrate interactions), Sierra and Müller (2015) concluded that this form of matrix equation such as eq. 1 represents the majority of terrestrial C cycle models. Similarly, Manzoni and Porporato (2009) concluded their review of 250 models that the majority of them use ordinary differential equations, which can be summarized by eq. 1, to describe land C cycle. Our mathematical analysis in this study used matrix operations of eq. 1 to reveal determinants of transient dynamics of terrestrial C cycle, including direction and rate of C storage changes, in response to global change. We examined assumptions underlying this equation and the validity of our analysis in the Discussion section.

### 2.2 TECO Model, its physical emulator, and numerical experiments
We conducted numerical experiments to support the mathematical analysis and thus help understand the characteristics of terrestrial C storage dynamics using the Terrestrial ECOSystem...
(TECO) model. TECO has five major components: canopy photosynthesis, soil water dynamics, plant growth, litter and soil carbon decomposition and transformation, and nitrogen dynamics as described in detail by Weng and Luo (2008) and Shi et al. (2016). Canopy photosynthesis is referred from a two-leaf (sunlit and shaded) model developed by Wang and Leuning (1998). This submodel simulates canopy conductance, photosynthesis, and partitioning of available energy. The model combines the leaf photosynthesis model developed by Farquhar et al. (1980) and a stomatal conductance model (Harley et al., 1992). In the soil water dynamic submodel, soil is divided into 10 layers. The surface layer is 10 cm deep and the other 9 layers are 20 cm deep. Soil water content (SWC) in each layer results from the mass balance between water influx and efflux. The plant growth submodel simulates C allocation and phenology. Allocation of C among three plant pools, which are leaf, fine root, and wood, depends on their growth rates (Fig. 1a). Phenology dynamics are related to leaf onset, which is triggered by growing degree days, and leaf senescence, which is determined by temperature and soil moisture. The C transformation submodel estimates carbon transfer from plants to two litter pools and three soil pools (Fig. 1a). The nitrogen (N) submodel is fully coupled with C processes with one additional mineral N pool. Nitrogen is absorbed by plants from mineral soil and then partitioned among leaf, woody tissues and fine roots. Nitrogen in plant detritus is transferred among different ecosystem pools (i.e. litter, coarse wood debris, fast, slow and passive SOM) (Shi et al., 2016). The model is driven by climate data, which include air and soil temperature, vapor-pressure deficit, relative humidity, incident photosynthetically active radiation, and precipitation at hourly steps.

We first calibrated TECO with eddy flux data collected at Harvard Forest from 2006-2009. The calibrated model was spun up to the equilibrium state in pre-industrial environmental conditions by recycling a 10-year climate forcing (1850-1859). Then the model was used to
simulate C dynamics from year 1850 to 2100 with the historical forcing scenario for 1850-2005 and RCP8.5 scenario for 2006-2100 as in the Community Land Model 4.5 (Oleson et al., 2013) in the grid cell where Harvard Forest is located.

To support the mathematical analysis using eq. 1, we first developed a physical emulator (i.e., the matrix representation of eq. 1) of the TECO model and then verified that the physical emulator can exactly represent simulations of the original TECO model. We first identified those parameter values in each of the C balance equations in the TECO model that are corresponding to elements in matrices $A$ and $K$ in eq. 1. The time-dependent variables for $u(t)$, elements in vector $B$, and elements in matrix $\xi(t)$ in the physical emulator were directly from outputs of the original TECO model. Then those parameter values and time-dependent variables were organized into matrices $A$, $\xi(t)$, and $K$, vectors $X(t)$, $X^0$, and $B$, and variable $u(t)$. Note that values of $u(t)$, $B$, and $\xi(t)$ could be different among different climate scenarios. Those matrices, vectors, and variable were entered to matrix calculation to compute $X'(t)$ using eq. 1. The sum of elements in calculated $X'(t)$ is a 100% match with simulated net ecosystem production (NEP) with the TECO model (Fig. 1b).

Once eq. 1 was verified to exactly replicate TECO simulations, we used TECO to generate numerical experiments to support the mathematical analysis on the transient dynamics of terrestrial C storage. To analyze the seasonal patterns of C storage dynamics, we averaged 10 series of three-year seasonal dynamics from 1851-1880. Then we used a 7-day moving window to further smooth the data.

3. Results
3.1 Determinants of C storage dynamics

The transient dynamics of terrestrial carbon storage are determined by two components: the C storage capacity and the C storage potential. The two components of C storage dynamics can be mathematically derived from multiplying both sides of eq. 1 by \((A\xi(t)K)^{-1}\) as:

\[ X(t) = (A\xi(t)K)^{-1}Bu(t) - (A\xi(t)K)^{-1}X'(t) \]  

(2)

The first term on the right side of eq. 2 is the C storage capacity and the second term is the C storage potential. Fig. 2a shows time courses of C storage and its capacity over one year for the leaf pool of Harvard Forest.

In eq. 2, we name the term \((A\xi(t)K)^{-1}\) the chasing time, \(\tau_{ch}(t)\), with a time unit used in exit rate \(K\). The chasing time is defined as:

\[ \tau_{ch}(t) = (A\xi(t)K)^{-1} \]  

(3)

\(\tau_{ch}(t)\) is a matrix of C residence times through the network of individual pools each with a different residence time and fractions of received C connected by pathways of C transfer.

Analogous to the fundamental matrix measuring life expectancies in demographic models (Caswell, 2000), the matrix, \(\tau_{ch}(t)\), here measures expected residence time of a C atom in pool \(i\) when it has entered from pool \(j\). We call this matrix the fundamental matrix of chasing times to represent the time scale at which the net C pool change, \(X'(t)\), is redistributed in the network.

Meanwhile, the residence times of individual pools in network, \(\tau_{pi}(t)\), can be estimated by multiplying the fundamental matrix of chasing times, \((A\xi(t)K)^{-1}\), with a vector of partitioning coefficients, \(B\) as:

\[ \tau_{pi}(t) = (A\xi(t)K)^{-1}B \]  

(4a)

Ecosystem residence time, \(\tau_{E}(t)\), is the sum of the residence time of all individual pools in network as:

Yiqi Luo 12/6/2016 10:23 AM  
Deleted: left

Yiqi Luo 12/6/2016 10:26 AM  
Deleted: capacities as measured by their

Yiqi Luo 12/6/2016 10:28 AM  
Deleted: s

Yiqi Luo 12/6/2016 10:28 AM  
Deleted: by

Yiqi Luo 12/6/2016 10:31 AM  
Deleted: E

Yiqi Luo 12/6/2016 10:32 AM  
Deleted: ,
Thus, the C storage capacity can be defined by:

\[ X_c(t) = (A\xi(t)K)^{-1}Bu(t) \]  

(5a)

Or it can be estimated from C input, \( u(t) \), and residence time, \( \tau_{\text{ch}}(t) \), as:

\[ X_c(t) = \tau_{\text{ch}}(t)u(t) \]  

(5b)

As C input (e.g., Gross or Net Primary Productions, GPP or NPP) and residence times vary with time, the C storage capacity varies with time. It represents instantaneous responses of the terrestrial C cycle to the external forcing. The modeled C storage capacity in the leaf pool (Fig. 2a), for example, increases in spring, reaches the peak at summer, declines in autumn, and becomes minimal in winter largely due to strong seasonal changes in C input (Fig. 2b). Note that either GPP or NPP can be used as C input for analysis of transient C dynamics. Estimated residence times, however, are smaller with GPP as C input than those with NPP as input. In this paper, we mostly used NPP as C input as that fraction of C is distributed among pools.

The C storage potential at time \( t \), \( X_p(t) \), can be mathematically described as:

\[ X_p(t) = (A\xi(t)K)^{-1}X'(t) \]  

(6a)

Or it can be estimated from net C pool change, \( X'(t) \), and chasing time, \( \tau_{\text{ch}}(t) \), as:

\[ X_p(t) = \tau_{\text{ch}}(t)X'(t) \]  

(6b)

Eqs. 6a and 6b suggest that the C storage potential represents re-distribution of net C pool change, \( X'(t) \), of individual pools through a network of pools with different residence times as connected by C transfers from one pool to the others through all the pathways. As time evolves, the net C pool change, \( X'(t) \), is redistributed again and again through the network of pools. The network of redistribution of next C pool change, thus, represents the potential of an ecosystem to store additional C when it is positive and lose C when it is negative. The C storage potential can
also be estimated from the difference between the C storage capacity and the C storage itself at
time $t$ as:

$$X_p(t) = X_c(t) - X(t)$$  \hspace{1cm} (6c)

The C storage potential in the leaf pool, for example, is about zero in winter and early spring
when the C storage capacity is very close to the storage itself (Fig. 2a). The C storage potential is
positive when the capacity is larger than the storage itself from late spring to summer and early
fall. As the storage capacity decreases to the point when the storage equals the capacity on the
265th day of year (DOY), the C storage potential is zero. After that day, the C storage potential
becomes negative.

Dynamics of ecosystem C storage, $X(t)$, can be characterized by three parameters: C
influx, $u(t)$, residence times, $\tau_c(t)$, and the C storage potential $X_p(t)$ as:

$$X(t) = \tau_c(t)u(t) - X_p(t)$$ \hspace{1cm} (7)

Eq. 7 represents a three-dimensional (3D) parameter space within which model simulation
outputs can be placed to measure how and how much they diverge.

Note that sums of elements in vectors $X(t)$, $X_c(t)$, $X_p(t)$, and $X'(t)$ are corresponding,
respectively, to the whole ecosystem C stock, ecosystem C storage capacity, ecosystem C storage
potential, and net ecosystem production (NEP). In this paper, we describe them wherever
necessary rather than use a separate set of symbols to represent those sums,

3.2 Direction and rate of C storage change at a given time

Like studying any moving object, quantifying dynamics of land C storage needs to determine
both the direction and the rate of its change at a given time. To determine the direction and rate
of C storage change, we re-arranged eq. 2 to be:
or re-arranging eq. 6a leads to:

\[ X'(t) = A \xi (t) K X_p(t) \]  

(8b)

As all the elements in \( \tau_{ch} \) are positive, the sign of \( X'(t) \) is the same as for \( X_p(t) \). That means \( X'(t) \) increases when \( X_c(t) > X(t) \), does not change when \( X_c(t) = X(t) \), and decreases when \( X_c(t) < X(t) \) at the ecosystem scale. Thus, the C storage capacity, \( X_c(t) \), is an attractor and hence determines the direction toward which the C storage, \( X(t) \), chases at any given time point.

The rate of C storage change, \( X'(t) \), is proportional to \( X_p(t) \) and also regulated by \( \tau_{ch} \).

When we study C cycle dynamics, we are interested in understanding dynamics of not only a whole ecosystem but also individual pools. Eq. 8a can be used to derive equations to describe C storage change for an \( i \)th pool as:

\[ \sum_{j=1}^{n} f_{ij} \tau_j x'_j(t) = \sum_{j=1}^{n} f_{ij} \tau_j b_j u(t) - x_i(t) = x_{p,i}(t) \]  

(9a)

where \( n \) is the number of pools in a C cycle model, \( f_{ij} \) is a fraction of C transferred from pool \( j \) to \( i \) through all the pathways, \( \tau_i \) measure residence times of individual pools in isolation (in contrast to \( \tau_{ch} \) in the network), \( x'_j \) is the net C change in the \( j \)th pool, \( b_j \) is a partitioning coefficient of C input to the \( j \)th pool, \( x_i(t) \) is the C storage in the \( i \)th pool, and \( x_{p,i}(t) \) is the C storage potential in the \( i \)th pool. Eq. 9a means that the C storage potential of each pool at time \( t \), \( x_{p,i}(t) \), is the sum of all the individual net C pool change, \( x'_j \), multiplied by corresponding residence time spent in pool \( i \) coming from pool \( j \). Through re-arrangement, eq. 9a can be solved for each individual pool net C change as a function of C storage potential of all the pools as:

\[ x'_i(t) = \frac{x_{ci,u}(t) - x_{ci,p}(t) - x_i(t)}{f_{ii} \tau_i} \]  

(9b)
where \( x_{c,i,u}(t) = \sum_{j=1}^{n} f_{ij} \tau_i b_j u(t) \) for the maximal amount of C that can transfer from C input to the \( i^{th} \) pool. \( x_{c,i,p}(t) = \sum_{j=1,j\neq i}^{n} f_{ij} \tau_j x'_j(t) \) for the maximal amount of C that can transfer from all the other pools to the \( i^{th} \) pool. \( f_{ii} = 1 \) for all the pools if there is no feedback of C among soil pools. \( f_{ii} < 1 \) when there are feedbacks of C among soil pools.

As plant pools get C only from photosynthetic C input, \( u(t) \), but not from other pools,

the direction and rate of C storage change in the \( i^{th} \) plant pool is determined by:

\[
\begin{align*}
    x'_i(t) &= \frac{x_{c,i}(t) - x_i(t)}{\tau_i} = \frac{x_{p,i}(t)}{\tau_i} \quad \text{for } i = 1, 2, 3 \\
    x_{c,i}(t) &= b_i u(t) \tau_i
\end{align*}
\] (10)

The C storage capacity of plant pools equals the product of plant C input, \( u(t) \) (i.e., net primary production, NPP), partitioning coefficient, \( b_i \), and residence time, \( \tau_i \), of its own pool (Fig. 2b-d). Thus, the C storage capacities of the leaf, root, and wood pools are high in summer and low in winter. Plant C storage, \( x_i(t) \), still chased the storage capacity, \( x_{c,i}(t) \), of its own pool at a rate that is proportional to \( X_{p,i}(t) \). For the leaf pool, the C storage, \( x_1(t) \), increases when \( x_{c,1}(t) > x_1(t) \) (or \( x_{p,1}(t) > 0 \)) from late spring until early fall on the 265th day of year (DOY) and then decreases when \( x_{c,1}(t) < x_1(t) \) (or \( x_{p,1}(t) < 0 \)) from DOY of 265 until 326 during fall (Fig. 2a).

However, the direction of C storage change in litter and soil pools are no longer solely determined by the storage capacity, \( x_{c,i}(t) \), of their own pools or at a rate that is proportional to \( X_{p,i}(t) \). The C storage capacity of one litter or soil pool has two components. One component, \( x_{c,i,u}(t) \) is set by the amount of plant C input, \( u(t) \), going through all the possible pathways, \( f_{ij} b_j \), multiplied by residence time, \( \tau_i \), of its own pool. The second component measures the C exchange of one litter or soil pool with other pools according to net C pool change, \( x'_j(t) \), through pathways, \( f_{ij}, j \neq i \), weighed by residence time, \( \tau_i \), of its own pool. For example, C input to the litter pool is a combination of C transfer from C input through the leaf, root, and
wood pools (Fig. 3c, 3d, and 3e) and C transfer due to the net C pool changes in the leaf, root, and wood pools (Fig. 3f, 3g, and 3h). Thus the first capacity component of the litter pool to store C is the sum of three products of NPP, C partitioning coefficient, and network residence time, respectively, through the leaf, root, and wood pools (Fig. 3c, 3d, and 3e). The second capacity component is the sum of other three products of C transfer coefficient along all the possible pathways, network residence time, and net C pool changes, respectively, in the leaf, root, and wood pools (Fig. 3f, 3g, and 3h). Thus, C storage in the $i^{th}$ pool, $x_i(t)$, chases an attractor,

$$\left(\sum_{i=1}^{n} f_{ij} b_j u(t) - \sum_{j=1}^{n} f_{ij} r_{ij} x_j(t)\right) \tau_i,$$

for litter and soil pools (Fig. 4).

In summary, due to the network of C transfer, C storage in litter and soil pools does not chase the C storage capacities of their own pools in a multiple C pool model (Fig. 4). The capacities for individual litter and soil pools measure the amounts of C that is transferred from photosynthetic C input through plant pools to be stored in those pools. However, those litter and soil pools also exchange C with other pools according to transfer coefficients along pathways of C movement multiplying net C pool change in those pools. Integration of the C input and C exchanges together still sets as a moving attractor toward which individual pool C storage approaches (Fig. 4).

### 3.3 C storage dynamics under global change

In response to a global change scenario that combines historical change and simulated RCP8.5 in the TECO experiment, the modeled ecosystem C storage capacity (the sum of all elements in vector $X_e(t)$) at Harvard Forest increases from 27 kg C m$^{-2}$ in 1850 to approximately 38 kg C m$^{-2}$ in 2100 with strong interannual variability (Fig. 5a). The increasing capacity results from a combination of a nearly 44% increase in NPP with a ~2% decrease in ecosystem residence times.
variability in the modeled capacity is attributable to the variability in NPP and residence times,
both of which directly respond to instantaneous variations in environmental factors. In
comparison, the ecosystem C storage (the sum of all elements in vector \( X(t) \)) itself gradually
increases, lagging behind the capacity, with much dampened interannual variability (Fig. 5a).
The dampened interannual variability is due to smoothing effects of pools with various residence
times. In response to global change scenario RCP8.5, the ecosystem C storage potential (the sum
of all elements in vector \( X_p(t) \)) in the Harvard Forest ecosystem increases from zero at 1980 to
3.5 kg C m\(^{-2}\) in 2100 with strong fluctuation over years (Fig. 5a). Over seasons, the potential is
high during the summer and low in winter, similarly with the seasonal cycle of the C storage
capacity.

Since chasing time, \( \tau_{ch} \), is a matrix and net C pool change, \( X'(t) \), is a vector, eq. 6a or 6b
(i.e., the C storage potential) can not be analytically separated into the chasing time and net C
pool change as can the capacity into C input and residence time in eq. 5a or 5b for traceability
analysis. The relationships among the three quantities can be explored by regression analysis.
The ecosystem C storage potential fluctuates in a similar phase with NEP from 1850 to 2100
(Fig. 5c). Consequently, the C storage potential is well correlated with NEP at the whole
ecosystem scale (Fig. 5d). The slope of the regression line is a statistical representation of
ecosystem chasing time. In this study, we find that \( r^2 \) of the relationship between the storage
potential and NEP is 0.79. The regression slope is 28.1 years in comparison with the ecosystem
residence time of approximately 22 years (Fig. 5b).

The capacity and storage itself of individual pools display similar long-term trends and
interannual variability to those for the total ecosystem C storage dynamics (Fig. 6). Noticeably,
19

the deviation of the C storage from the capacity, which is the C storage potential, is much larger for pools with long residence times than those with short residence times. For individual pools, the potential is nearly zero for those fast turnover pools and becomes very large for those pools with long residence time (Fig. 6).

For individual plant pools, eq. 10 describes the dependence of the C storage potential, \( x_{p,i}(t) \), on the pool-specific residence time, \( \tau_i \), \( i = 1, 2, \) and \( 3, \) and net C pool change of their own pools, \( x'_i(t), i = 1, 2, \) and \( 3. \) Thus, one value of \( x_{p,i}(t) \) is exactly corresponding to one value of \( x'_i(t) \) at slope of \( \tau_i \), leading to the correlation coefficient in Fig. 7 being 1.00 for leaf, root, and wood pools. For a litter or soil pool, however, the C storage potential is not solely dependent on the residence time and net C pool change of its own pool but influenced by several other pools. Thus, the potential of one litter or soil pool is correlated with net C pool changes of several pools with different regression slopes (Fig. 7).

4 Discussion

4.1 Assumptions of the C cycle models and validity of this analysis

This analysis is built upon eq. 1, which represents the majority of terrestrial C cycle models developed in the past decades (Manzoni and Porporato, 2009; Sierra and Müller, 2015). Those models have several assumptions, which may influence the validity of this analysis. First, those models assume that donor pools control C transfers among pools and decomposition follows 1\(^{st}\)-order decay functions (Assumption 1). This assumption is built upon observations from litter and SOC decomposition. Analysis of data from nearly 300 studies of litter decomposition (Zhang et al., 2008), about 500 studies of soil incubation (Schädel et al., 2014; Xu et al., 2016), more than 100 studies of forest succession (Yang et al., 2011), and restoration
(Matamala et al., 2008) almost all suggest that the 1st-order decay function captures macroscopic patterns of land C dynamics. Even so, its biological, chemical and physical underpinnings need more study (Luo et al., 2016). This assumption has recently been challenged by a notion that microbes are actively involved in decomposition processes. To describe the active roles of microbes in organic C decomposition, a suite of nonlinear microbial models has been proposed using Michaelis-Menten or reverse Michaelis-Menten equations (Allison et al., 2010; Wieder et al., 2013). Those nonlinear models exhibit unique behaviors of modeled systems, such as damped oscillatory responses of soil C dynamics to small perturbations and insensitivity of the equilibrium pool sizes of litter or soil carbon to inputs (Li et al., 2014; Wang et al., 2014; 2016). Oscillations have been documented for single enzymes at timescales between \(10^{-4}\) to 10 seconds (English et al., 2006; Goldbeter, 2013; Xie, 2013). Over longer timescales with mixtures of large diversity of enzyme-substrate complexes in soil, oscillations may be likely averaged out so that the 1st order decay functions may well approximate these average dynamics of organic matter decomposition (Sierra and Müller, 2015).

Second, those models all assume that multiple pools can adequately approximate transformation, decomposition, and stabilization of SOC in the real world (Assumption 2). The classic SOC model, CENTURY, uses three conceptual pools, active, slow, and passive SOC, to represent SOC dynamics (Parton et al., 1987). Several models define pools that are corresponding to measurable SOC fractions to match experimental observation with modeling analysis (Smith et al., 2002; Stewart et al., 2008). Carbon transformation in soil over time has also been described by a partial differential function of SOM quality (Bosatta and Ågren, 1991; Ågren and Bosatta, 1996). The latter quality model describes the external inputs of C with certain quality, C loss due to decomposition, and the internal transformations of the quality of
It has been shown that multi-pool models can approximate the partial differential function or continuous quality model as the number of pools increases (Bolker et al., 1998; Sierra and Müller, 2015).

Assumption 3 is on partitioning coefficients of C input (i.e., elements in vector $B$) and C transformation among plant, litter, and soil pools (i.e., elements in the matrix, $A\xi(t)K$). Some of the terrestrial C cycle models assume that elements in vector $B$, and matrices $A$ and $K$ are constants. All the factors or processes that vary with time are represented in the diagonal matrix $\xi(t)$. In the real world, C transformation are influenced by environmental variables (e.g., temperature, moisture, oxygen, N, phosphorus, and acidity varying with soil profile, space, and time), litter quality (e.g., lignin, cellulose, N, or their relative content), organomineral properties of SOC (e.g., complex chemical compounds, aggregation, physiochemical binding and protection, reactions with inorganic, reactive surfaces, and sorption), and microbial attributes (e.g., community structure, functionality, priming, acclimation, and other physiological adjustments) (Luo et al., 2016). It is not practical to incorporate all of those factors and processes into one model. Only a subset of them is explicitly expressed while the majority is implicitly embedded in the C cycle models. Empirical studies have suggested that temperature, moisture, litter quality, and soil texture are primary factors that control C transformation processes of decomposition and stabilization (Burke et al., 1989; Adair et al., 2008; Zhang et al., 2008; Xu et al., 2012; Wang et al., 2013). Nitrogen influences C cycle processes mainly through changes in photosynthetic C input, C partitioning, and decomposition. It is yet to identify how other major factors and processes, such as microbial activities and organomineral protection, regulate C transformation.
Assumption 4 is that terrestrial C cycle models use different response functions (i.e., different $\xi(t)$ in eq. 1) to represent C cycle responses to external variables. As temperature modifies almost all processes in the C cycle, different formulations, including exponential, Arrhenius, and optimal response functions, have been used to describe C cycle responses to temperature changes in different models (Lloyd and Taylor, 1994; Jones et al., 2005; Sierra and Müller, 2015). Different response functions are used to connect C cycle processes with moisture, nutrient availability, soil clay content, litter quality, and other factors. Different formulations of response functions may result in substantially different model projections (Exbrayat et al., 2013) but unlikely change basic dynamics of the model behaviors.

Assumption 5 is that disturbance events are represented in models in different ways (Grosse et al., 2011; West et al., 2011; Goetz et al., 2012; Hicke et al., 2012). Fire, extreme drought, insect outbreaks, land management, and land cover and land use change influence terrestrial C dynamics via 1) altering rate processes, for example, gross primary productivity (GPP), growth, tree mortality, or heterotrophic respiration; 2) modifying microclimatic environments; 3) transferring C from one pool to another (e.g., from live to dead pools during storms or release to the atmosphere with fire) (Kloster et al., 2010; Thonicke et al., 2010; Luo and Weng, 2011; Prentice et al., 2011; Weng et al., 2012). Those disturbance influences can be represented in terrestrial C cycle models through changes in parameter values, environmental scalars, and/or discrete C transfers among pools of eq. 1 (Luo and Weng 2011). While eq. 1 does not explicitly incorporate disturbances for their influences on land C cycle, Weng et al. (2012) developed a disturbance regime model that combines eq. 1 with frequency distributions of disturbance severity and intervals to quantify net biome exchanges.

The sixth assumption that those models make is that the lateral C fluxes through erosion
or local C drainage is negligible so that eq (1) can approximate terrestrial C cycle over space. If
soil erosion is substantial enough to be modeled with horizontal movement of C, a third
dimension should be added in addition to two-dimensional transfers in classic models.

Our analysis on transient dynamics of terrestrial C cycle is valid unless some of the
assumptions are violated. Assumption 1 on the 1st-order decay function of decomposition
appears to be supported by thousands of datasets. It is a burden on microbiologists to identify
empirical evidence to support the nonlinear microbial models. Assumption 2 may not affect the
validity of our analysis no matter how C pools are divided in the ecosystems. Our analysis in this
study is applicable no matter whether elements are time-varying or constant in vector B and
matrices A and K as in assumption 3. Neither assumption 4 nor 5 would affect the analysis in this
study. The environmental scalar, , as related to assumption 4 can be any forms in the derived
equations (e.g., eq. 2). Disturbances of fire, land use, and extreme drought change rate processes
but do not alter the basic formulation of eq. 1. If soil erosion and lateral transportation of C
become a major research objective, Eq. (1) can no longer be analyzed to understand the
mathematical foundation underlying transient dynamics of terrestrial C cycle.

4.2 Carbon storage capacity

One of the two components this analysis introduces to understand transient dynamics of
terrestrial C storage is the C storage capacity (Eq. 2). Olson (1963) is probably among the first
who systematically analyzed C storage dynamics at forest floor as functions of litter production
and decomposition. He collected data of annual litter production and approximately steady-state
organic C storage at forest floor, from which decomposition rates were estimated for a variety of
ecosystems from Ghana in the tropics to alpine forests in California. Using the relationships
among litter production, decomposition, and C storage, Olson (1963) explored several issues, such as decay without input, accumulation with continuous or discrete annual litter fall, and adjustments in production and decay parameters during forest succession. His analysis approximated the steady-state C storage as the C input times the inverse of decomposition (i.e., residence time). The steady-state C storage is also considered the maximal amount of C that a forest can store.

This study is not only built upon Olson’s analysis but also expands it at least in two aspects. First, we similarly define the C storage capacity (i.e., eqs. 5a and 5b). Those equations can be applied to a whole ecosystem with multiple C pools while Olson’s analysis is for one C pool. Second, Olson (1963) treated the C input and decomposition rate as yearly constants at a given location even though they varied with locations. This study considers both C input and rate of decomposition being time dependent. A dynamical system with its input and parameters being time dependent mathematically becomes a nonautonomous system (Kloeden and Rasmussen, 2011). As terrestrial C cycle under global change is transient, we need to treat it a nonautonomous system to better understand the properties of transient dynamics. Olson (1963) approximated the non-autonomous system at the yearly time scale without global change so as to effectively understand properties of the steady-state C storage at the forest floor. In comparison, eqs. 5a and b are not only more general but also essential for understanding transient dynamics of the terrestrial C cycle in response to global change.

Under the transient dynamics, the C storage capacity as defined by eqs 5a and b still sets the maximal amount of C that one ecosystem can store at time t. This capacity represents instantaneous responses of ecosystem C cycle to external forcing via changes in both C input and residence time, and thus varies within one day, over seasons of a year, and interannually over...
longer time scales as forcings vary. The variation of the C storage capacity can result from cyclic environmental changes (e.g., dial and seasonal changes), directional \textit{global change} (e.g., rising atmospheric CO$_2$, nitrogen deposition, altered precipitation, and warming), disturbance events, disturbance regime shifts, and changing vegetation dynamics (Luo and Weng, 2011). As the capacity sets the maximal amount of C storage (Fig. 2a), it is a moving attractor toward which the current C storage chases. When the capacity is larger than the C storage itself, C storage increases. Otherwise, the C storage decreases.

4.3 Carbon storage potential

The C storage potential represents the internal capability to equilibrate the current C storage with the capacity. Bogeochemically, the C storage potential represents re-distribution of net C pool change, $X'(t)$, of individual pools through a network of pools with different residence times as connected by C transfers from one pool to the others through all the pathways. The potential is conceptually equivalent to the magnitude of disequilibrium as discussed by Luo and Weng (2011). Extensive studies have been done to quantify terrestrial C sequestration. The most commonly estimated quantities for C sequestration include net ecosystem exchange (NEE), C stocks in ecosystems (i.e., plant biomass and SOC) and their changes (Baldocchi et al., 2001; Pan et al., 2013). This study, for the first time, offers the theoretical basis to estimate the terrestrial C storage potential in at least two approaches: (1) the product of chasing time and net C pool change with eqs. 6a and 6b; and (2) the difference between the C storage capacity and the C storage itself with eqs. 6c. Since the time-varying C storage capacity is fully defined by residence time and C input at any given time, C storage potential can be estimated from three
To effectively quantify the C storage potential in terrestrial ecosystems, we need various data sets from experimental and observatory studies to be first assimilated into models. For example, data from Harvard Forest were first used to constrain the TECO model. The constrained model was used to explore changes in ecosystem C storage in response to global change scenario, RCP8.5. That scenario primarily stimulated NPP, which increased from 1.06 to 1.8 kg C m\(^{-2}\) yr\(^{-1}\) in the Harvard Forest (Fig. 5b). Although climate warming decreased residence time in the Harvard Forest, the substantial increases in NPP resulted in increases in the C storage potential over time.

### 4.4 Novel approaches to model evaluation and improvement

Our analysis of transient C cycle dynamics offers new approaches to understand, evaluate, diagnose, and improve land C cycle models. We have demonstrated that many global land C cycle models can be exactly represented by the matrix equation (Eqs. 1 and 2) (i.e., physical emulators). As a consequence, outputs of all those models can be placed into a three-dimensional (3D) space (Eq. 7) to measure their differences. In addition, components of land C cycle models are simulated in a mutually independent fashion so that modeled C storage can be decomposed into traceable components for traceability analysis. Moreover, the physical emulators computationally enable data assimilation to constrain complex models.

*Physical Emulators of land C cycle models* We have developed matrix representations (i.e., physical emulators) of CABLE, LPJ-GUESS, CLM3.5, CLM 4.0, CLM4.5, BEPS, and TECO (Xia et al., 2013; Hararuk et al., 2014; Ahlström et al., 2015; Chen et al., 2015). The emulators can exactly replicate simulations of C pools and fluxes with their original models.
when driven by a limited set of inputs from the full model (GPP, soil temperature, and soil moisture) (Fig. 1b and 1c). However, the physical emulators differ for different models as the elements of each matrix could be differently parameterized or formulized in different models. Also, different models usually have different pool-flux structures, leading to different non-zero elements in the $A$ matrix. Nonetheless, the physical emulators make complex models analytically clear and, therefore, give us a way to understand the effects of forcing, model structures, and parameters on modeled ecosystem processes. They greatly simplify the task of understanding the dynamics of submodels and interactions between them. The emulators allow us to analyze model results in the 3D parameter space and the traceability framework.

**Parameter space of C cycle dynamics** Eq. 7 indicates that transient dynamics of modeled C storage are determined by three parameters: C input, residence time, and C storage potential. The 3D parameter space offers one novel approach to uncertainty analysis of global C cycle models. As global land models incorporate more and more processes to simulate C cycle responses to global change, it becomes very difficult to understand or evaluate complex model behaviors. As such, differences in model projections cannot be easily diagnosed and attributed to their sources (Chatfield, 1995; Friedlingstein et al., 2006; Luo et al., 2009). Eq. 7 can help diagnose and evaluate complex models by placing all modeling results within one common parameter space in spite of the fact that individual global models may have tens or hundreds of parameters to represent C cycle processes as affected by many abiotic and biotic factors (Luo et al., 2016). The 3D space can be used to measure how and how much the models diverge.

**Traceability analysis** The two terms on the right side of eq. 2 can be decomposed into traceable components (Xia et al., 2013) so as to identify sources of uncertainty in C cycle model projections. Model intercomparison projects (MIPs) all illustrate great spreads in projected land
C sink dynamics across models (Todd-Brown et al., 2013; Tian et al., 2015). It has been extremely challenging to attribute the uncertainty to sources. Placing simulation results of a variety of C cycle models within one common parameter space can measure how much the model differences are in a common metrics (Eq. 7). The measured differences can be further attributed to sources in model structure, parameter, and forcing fields with traceability analysis (Xia et al., 2013; Rafique et al., 2014; Ahlström et al., 2015; Chen et al., 2015). The traceability analysis also can be used to evaluate effectiveness of newly incorporated modules into existing models, such as adding the N module on simulated C dynamics (Xia et al., 2013) and locate the origin of model ensemble uncertainties to external forcing vs. model structures and parameters (Ahlström et al., 2015).

**Constrained estimates of terrestrial C sequestration** Traditionally, global land C sink is indirectly estimated from airborne fraction of C emission and ocean uptake. Although many global land models have been developed to estimate land C sequestration, a variety of MIPs indicate that model predictions widely vary among them and do not fit observations well (Schwalm et al., 2010; Luo et al., 2015; Tian et al., 2015). Moreover, the prevailing practices in the modeling community, unfortunately, may not lead to significant enhancements in our confidence on model predictions. For example, incorporating an increasing number of processes that influence the C cycle may represent the real-world phenomena more realistically but makes the models more complex and less tractable. MIPs have effectively revealed the extent of the differences between model predictions (Schwalm et al., 2010; Keenan et al., 2012; De Kauwe et al., 2013) but provide limited insights into sources of model differences (but see Medlyn et al., 2015). The physical emulators make data assimilation computationally feasible for global C cycle models (Hararuk et al., 2014; 2015) and thus offer the possibility to generate independent...
yet constrained estimates of global land C sequestration to be compared with the indirect
estimate from the airborne fraction of C emission and ocean uptake. With the emulators, we can
assimilate most of the C flux- and pool-related datasets into those models to better constrain
global land C sink dynamics.

5 Concluding remarks

In this study we theoretically explored the transient dynamics of terrestrial C storage. Our
analysis indicates that transient C storage dynamics can be partitioned into two components: the
C storage capacity and the C storage potential. The capacity, which is the product of C input and
residence time, represents their instantaneous responses to a state of external forcing at a given
time. Thus, the C storage capacity quantifies the maximum amount of C that an ecosystem can
store at the given environmental condition at a point of time. Thus it varies diurnally, seasonally,
and interannually as environmental condition changes.

The C storage potential is the difference between the capacity and the current C storage
and thus measures the magnitude of disequilibrium in the terrestrial C cycle (Luo and Weng,
2011). The storage potential represents the internal capability (or recovery force) of the
terrestrial C cycle to influence the change in C storage in the next time step through
redistribution of net C pool changes in a network of multiple pools with different residence
times. The redistribution drives the current C storage towards the capacity and thus equilibrates
C efflux with influx.

The two components of land C storage dynamics represent interactions of external forces
(via changes in the capacity) and internal capability of the land C cycle (via changes in the C
storage potential) to generate complex phenomena of C cycle dynamics, such as fluctuations,
directional changes, and tipping points, in the terrestrial ecosystems. From a system perspective, these complex phenomena could not be generated by relatively simple internal processes but are mostly caused by multiple environmental forcing variables interacting with internal processes over different temporal and spatial scales as explained by Luo and Weng (2011) and Luo et al. (2015). Note that while those internal processes can be mathematically represented with a relatively simple formula, their ecological and biological underpinnings can be very complex.

The theoretical framework developed in this study has the potential to revolutionize model evaluation. Our analysis indicates that the matrix equation as in eq. 1 or 2 can adequately emulate most of the land C cycle models. Indeed, we have developed physical emulators of several global land C cycle models. In addition, predictions of C dynamics with complex land models can be placed in a 3D parameter space as a common metric to measure how much model predictions are different. The latter can be traced to its source components by decomposing model predictions to a hierarchy of traceable components. Moreover, the physical emulators make it computationally possible to assimilate multiple sources of data to constrain predictions of complex models.

The theoretical framework we developed in this study can well explain dynamics of C storage in response to cyclic seasonal change in external forcings (e.g., Figs. 2 and 3), climate change, and rising atmospheric CO₂ (Fig. 5). It can also explain responses of ecosystem C storage to disturbances and other global change factors, such as nitrogen deposition, land use changes, and altered precipitation. The theoretical framework is simple and straightforward but able to characterize the direction and rate of C storage change, which are arguably among the most critical issues for quantifying terrestrial C sequestration. Future research should explicitly incorporate stochastic disturbance regime shifts (e.g., Weng et al., 2012) and vegetation changes.
dynamics (Moorcroft et al., 2001; Purves and Pacala, 2008; Fisher et al., 2010; Weng et al., 2015) into this theoretical framework to explore their theoretical issues related to biogeochemistry.

6 Code availability

Computer code of the TECO model and its physical emulator are available at http://ecolab.ou.edu/download/TECO%20Emulator.php

Acknowledgements: This work was partially done through the working group, Nonautonomous Systems and Terrestrial Carbon Cycle, at the National Institute for Mathematical and Biological Synthesis, an institute sponsored by the National Science Foundation, the US Department of Homeland Security, and the US Department of Agriculture through NSF award no. EF-0832858, with additional support from the University of Tennessee, Knoxville. Research in Yiqi Luo EcoLab was financially supported by U.S. Department of Energy grants DE-SC0006982, DE-SC0008270, DE-SC0014062, DE-SC0004601, and DE-SC0010715 and U.S. National Science Foundation (NSF) grants DBI 0850290, EPS 0919466, DEB 0840964, and EF 1137293.

References


De Kauwe, M. G., Medlyn, B. E., Zaehle, S., Walker, A. P., Dietze, M. C., Hickler, T., Jain, A.
K., Luo, Y., Parton, W. J., Prentice, I. C., Smith, B., Thornton, P. E., Wang, S., Wang, Y.-P.,
Wårlind, D., Weng, E., Crous, K. Y., Ellsworth, D. S., Hanson, P. J., Seok Kim, H., Warren, J.
M., Oren, R., and Norby, R. J.: Forest water use and water use efficiency at elevated CO₂: a
model-data intercomparison at two contrasting temperate forest FACE sites, Global Change
Biology, 19, 1759-1779, 2013.

English, B. P., Min, W., Van Oijen, A. M., Lee, K. T., Luo, G., Sun, H., Cherayil, B. J., Kou, S.,
and Xie, X. S.: Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited,

carbon uncertainty in a global model: response of microbial decomposition to temperature,
mochre and nutrient limitation, Biogeosciences, 10, 7095-7108, 2013.

Farquhar, G., von Caemmerer, S., and Berry, J.: A biochemical model of photosynthetic CO₂

Fisher, R., McDowell, N., Purves, D., Moorcroft, P., Sitch, S., Cox, P., Huntingford, C., Meir, P.,
and Ian Woodward, F.: Assessing uncertainties in a second - generation dynamic vegetation

Fraser, L. H., Henry, H. A., Carlyle, C. N., White, S. R., Beierkuhnlein, C., Cahill, J. F., Casper,
B. B., Cleland, E., Collins, S. L., and Dukes, J. S.: Coordinated distributed experiments: an
emerging tool for testing global hypotheses in ecology and environmental science, Frontiers in

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 (CMIP)-M-4 model intercomparison, J Climate, 19,
3337-3353, 2006.

McNulty, S., O'Halloran, T., Harmon, M., Meddens, A. J. H., Pfeifer, E. M., Mildrexler, D., and
Kasischke, E. S.: Observations and assessment of forest carbon dynamics following disturbance

Goldbeter, A.: Oscillatory enzyme reactions and Michaelis–Menten kinetics, FEBS letters, 587,
2778-2784, 2013.

Grosse, G., Harden, J., Turetsky, M., McGuire, A. D., Camill, P., Tarnocai, C., Frolking, S.,
Schuur, E. A. G., Jorgenson, T., Marchenko, S., Romanovsky, V., Wickland, K. P., French, N.,
Waldrop, M., Bourgeau-Chavez, L., and Striegl, R. G.: Vulnerability of high-latitude soil organic


Fig. 1 The Terrestrial ECOsystem (TECO) model and its outputs. Panel a is a schematic representation of C transfers among multiple pools in plant, litter and soil in the TECO model. TECO has feedback loops of C among soil pools. CWD = coarse wood debris, SOM = Soil Organic Matter. Panel b compares the original TECO model outputs with those from matrix equations for net ecosystem production (NEP = the sum of elements in $X'(t)$ from eq. 1). The perfect match between the TECO outputs and NEP from eq. 1 is due to the fact that they are mathematically equivalent. Panel c compares the original TECO model outputs with those from matrix equations for ecosystem C storage (= the sum of elements in $X(t)$ from eq. 2). The C storage values calculated with eq. 2 are close to 1:1 line with $r^2=0.998$ with the modeled values (panel c). The minor mismatch in estimated C storage between the matrix equation calculation and TECO outputs is due to numerical errors via inverse matrix operation with some small numbers.

Fig. 2 Seasonal cycles of the C storage capacity and C storage dynamics for the leaf pool (i.e., pool 1 as shown in Fig. 1). All the components are showed in panels b-d to calculate $x_{c1}(t) = b_1 u(t) \tau_1$ through multiplication, where $u(t) = NPP$ and $\tau_1 = 1/k_1$ for leaf.

Fig. 3 Seasonal cycles of the C storage capacity and C storage dynamics for the litter pool (i.e., pool 4 as shown in Fig. 1). All the components are showed to calculate $x_{cA4}(t) = \sum_{j=1}^{n} f_{4j} r_4 b_j u(t)$ in panels b-e and $x_{cA4}(t) = \sum_{j=1}^{n} f_{4j} r_4 x_j'(t)$ in panels f-i for litter. $x_{cA4}(t)$ is the maximal amount of C that can transfer from C input to the litter pool. $x_{cA4}(t)$ is the maximal amount of C that can transfer from all the other pools to the litter pool. This figure is to illustrate the network of pools through which C is distributed.
Fig. 4 Components of the C storage capacity for litter pool (i.e., pool 4 as shown in Fig. 1). Component, $x_{c,t}(t)$, is the C from C input and component, $x_{c,A,p}(t)$, is the C from all the other pools to the litter pool. The sum of them is the attractor that determines the direction of C storage change in pool 4.

Fig. 5 Transient dynamics of ecosystem C storage in response to global change in Harvard Forest. Panel a shows the time courses of the ecosystem C storage capacity, the ecosystem C storage potential, and ecosystem C storage (i.e., C stock) from 1850 to 2100. Panel b shows time courses of NPP(t) as C input and ecosystem residence times. Panel c shows correlated changes in ecosystem C storage potential and net ecosystem production (NEP). Panel d illustrates the regression between the C storage potential and NEP.

Fig. 6 The C storage capacity ($x_{c,i}(t)$), the C storage potential ($x_{p,i}(t)$), and C storage ($x_i(t)$) of individual pools. The potential is nearly zero for those fast turnover pools with short residence times but very large for those pools with long residence times.

Fig. 7 The C storage potential of individual pools ($x_{p,i}$) as influenced by net C pool change of different pools ($x_{i}'$) in their corresponding rows. The correlation coefficients show the degree of influences of net C pool change in one pool on the C storage potential of the corresponding pool through the network of C transfer. Those empty cells indicate no pathways of C transfer between those pools as indicated in Fig. 1.
Fig. 1
Fig. 2

(a) C storage (g C m$^{-2}$)

(b) NPP (g C m$^{-2}$ d$^{-1}$)

(c) $b$

(d) $\beta$

X

DOY
Fig. 3
Fig. 4

Carbon storage (kg C m$^{-2}$) vs. DOY

- $X_4$
- $X_{c, 4, p}$
- $X_{c, 4, u}$
- $X_{c, 4, u} + X_{c, 4, p}$
Fig. 5
Leaf
Carbon storage (kg C m\(^{-2}\))

Wood

Root

Litter

CWD

Active SOM

Slow SOM

Passive SOM

Fig. 6