Divergent predictions of carbon storage between two global land models: Attribution of the causes through traceability analysis

Authors: Rashid Rafique[1,2], Jianyang Xia[1], Oleksandra Hararuk[1,3], Ghassem R. Asrar[2], Yingping Wang[4] ,Yiqi Luo[1]

Affiliation: [1]Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK, USA

[2] Joint Global Change Research Institute, Pacific Northwest National Lab, College Park, MD, USA

[3] Pacific Forestry Centre, Victoria, BC, Canada

[4] CSIRO Ocean and Atmosphere Flagship, PMB 1, Aspendale, Victoria 3195 Australia

Corresponding Authors: Rashid Rafique and Yiqi Luo

Phone: (301) 405-5353; Fax: (301) 314-6719; Email: rashidbao@gmail.com; yluo@ou.edu

Key Words: Traceability framework, Partitioning coefficients, Transfer coefficients, Carbon storage, Ecosystem carbon residence time, environmental scalars.

Abstract

Representations of the terrestrial carbon cycle in land models are becoming increasingly complex. It is crucial to develop approaches for critical assessment of the complex model properties in order to understand key factors contributing to models' performance. In this study, we applied a traceability analysis, which decomposes carbon cycle models into traceable components, for two global land models (CABLE and CLM-CASA') to diagnose the causes of their differences in simulating ecosystem carbon storage capacity. Driven with similar forcing data, CLM-CASA' predicted ~31% larger carbon storage capacity than CABLE. Since ecosystem carbon storage capacity is a product of net primary productivity (NPP) and ecosystem residence time (τ_E), the predicted difference in the storage capacity between the two models results from differences in either NPP or τ_E or both. Our analysis showed that CLM-CASA' simulated 37% higher NPP than CABLE. On the other hand, τ_E , which was a function of the baseline carbon residence time (τ'_E) and environmental effect on carbon residence time, was on average 11 years longer in CABLE than CLM-CASA'. This difference in τ_E was mainly caused by longer τ'_E of woody biomass (23 vs. 14 years in CLM-CASA'), and higher proportion of NPP allocated to woody biomass (23% vs. 16%). Differences in environmental effects on carbon residence times had smaller influences on differences in ecosystem carbon storage capacities compared to differences in NPP and τ'_E. Overall, the traceability analysis showed that the major causes of different carbon storage estimation were found to be parameters setting related to carbon input and baseline residence times between two models.

Introduction

Terrestrial ecosystems play a central role in the global carbon cycle as both a reservoir for carbon and as a regulator of atmospheric concentrations of carbon dioxide (CO_2) (Sitch *et al.*, 2015). Future concentrations of atmospheric CO_2 strongly depend on the feedbacks between terrestrial ecosystems and atmosphere; particularly the balance of carbon uptake, driven primarily by CO_2 in simulations; and loss of carbon from the ecosystems, driven primarily by temperature in simulations (Luo, 2007; Luo *et al.*, 2009; Thornton *et al.*, 2009). Improving our understanding of the processes by which ecosystems interact with the atmosphere is of fundamental importance for improving models' predictions (Zhou, *et al.*, 2012). Global land models are the major tools for investigating the climate impacts on terrestrial ecosystem carbon storage capacity (Luo *et al.*, 2012; Rafique *et al.*, 2014). Today's land models have become very sophisticated due to inclusion of multitude of different processes in the hope of simulating the real world more accurately. However, the addition of new processes not only increases the challenge of understanding the complex model behavior but also hinders the diagnosis of uncertainty in model outputs (Luo *et al.*, 2009; Xia *et al.*, 2013; Rafique *et al.*, 2014).

Many studies have been conducted on evaluation and intercomparison of carbon cycle components of land models (Johns *et al.*, 2011; Taylor *et al.*, 2011; Zaehle *et al.*, 2014), and most of these studies show large discrepancies in modeled carbon stocks and fluxes. For example, the Coupled Model Intercomparison Project (C4MIP) reported that carbon uptake responses to a doubling of atmospheric CO_2 concentrations varied from 100 to 800 Gt carbon amongst 11 models for the period 1850-2100 years (Friedlingstein *et al.*, 2006; Arora *et al.*, 2011). Similarly, Todd-Brown et al. (2013) reported that the present day total soil organic carbon simulated by CMIP5 models varied six fold ranging from approximately 510 to 3040 Pg of carbon. Most of these studies use a conventional approach for model intercomparison where models are analyzed by comparing their outputs among each other and with reference data set; however this approach is not sufficient for understanding the causes of discrepancies in model outputs.

There have been a few studies that attempt explaining some of these differences in model outputs by attributing sources of variations. For example, Mishra et al. (2013) identified uncertainties in modeling soil carbon in permafrost regions but insufficiently attributed these variations to different components of their model due to lack of comprehensive tractable approach. Wang et al. (2011) decomposed ecosystem models into several components, such as climate forcing, net primary productivity (NPP) allocation and decomposition rates. This study was partly successful in diagnosing uncertainties in simulated carbon dynamics. However, the framework they used could not adequately address the sources of variations to their origins thoroughly. For example, this framework was not sufficient to explain the variations in respirational fluxes (i.e. whether they were caused by carbon pool sizes or turnover rates). Similarly, Todd-Brown et al. (2013, 2014) explained the model differences based on the variations in NPP, bulk soil decomposition rates and temperature sensitivity. However, they did not describe the effects of parameterizations such as NPP partitioning, carbon transfer coefficients and decomposition rates of individual pools. These shortcomings can only be addressed after gaining more complete understanding of the model’s fundamental structural differences and its traceable components controlling the carbon dynamics.

The traceability framework developed by Xia et al. (2013) provides a powerful method for attributing the sources of variations to different components of models. This framework, based on fundamental properties of the carbon cycle, can be decomposed into few traceable components (Luo *et al.*, 2003; Luo & Weng, 2011). After carbon is fixed by photosynthesis, its further fate can be summarized by ecosystem carbon residence time, which is a length of time a carbon atom spends in ecosystem before leaving it via respiration (Luo *et al.*, 2001). The framework traces modeled ecosystem carbon storage capacity (Xss) to (i) a product of NPP and ecosystem residence time (τ_E). The latter ecosystem residence time can be further traced to (ii) baseline carbon residence times (τ'_E), which are function of model parameters representing vegetation characteristics and soil types, (iii) environmental scalars (ξ) including temperature and water scalars, and (iv) the external climate forcing.

In this study we applied the traceability framework to decompose two commonly used complex land models (CLM-CASA' and CABLE) at global and biome spatial scales into traceable components for better understanding of the sources of variations in modeled carbon storage capacity. The specific objectives of this study were: to (1) quantify the effects of NPP and ecosystem residence time in determining the ecosystem carbon storage and (3) investigate the impact of parameters (relating to NPP partitioning and carbon transfer coefficients) and environmental conditions in determining ecosystem's carbon residence time.

2.0 Methods

2.1 CABLE and CLM-CASA' models

CABLE is an Australian land model used for the simulation of land atmospheric exchanges (Kowalczyk *et al.*, 2006). The biogeochemical model in CABLE is adopted from CASACNP, a model developed by Wang e*t al.* (2010). CASACNP consists of tightly coupled carbon, nitrogen and phosphorus cycles. Like most of other land models, CABLE's carbon cycle also consists of typical pool and flux structure. There are nine carbon pools in the CABLE model: three plant pools, three litter pools and three soil pools. The carbon partitioning of photosynthetically fixed carbon into plant pools is controlled by the availability of light, water and nitrogen. The carbon transfer among pools is determined by the lignin/nitrogen ratio and the lignin fraction. The potential decay rates vary with vegetation types, lignin fraction and soil texture. The environmental scalar regulates the leaf turnover rates via limitations of soil moisture and soil temperature conditions. The more detailed description of CABLE model is given in Wang *et al.* (2011) and Xia *et al.* (2013).

CLM-CASA' model combines the biogeophysics of the CLM with Carnegie-Ames-Stanford Approach (CASA) biogeochemistry module (Oleson, *et al.*, 2008). The CLM, released in 2008, is a component of the Community Climate System Model (CCSM) (Oleson, *et al.*, 2007). CLM examines the physical, chemical, and biological processes through which terrestrial ecosystems interact with climate. CASA' simulates carbon dynamics at the plant functional type (PFT) level beginning with carbon assimilation via photosynthesis, to mortality and decomposition, and the release of CO_2 to the atmosphere. There are three plant carbon pools, six litter pools and three soil pools. A more detailed description of the model is provided by Doney *et al.* (2006).

Biomes for both CABLE and CLM-CASA' were constructed from the 1-km International Geosphere–Biosphere Program Data and Information System (IGBP DISCover) dataset (Loveland *et al.*, 2000). In CLM-CASA', however, the above dataset was combined with 1-km tree cover dataset published by the University of Maryland (DeFries *et al.*, 2000). The CABLE model has 9 biomes (8 used in this study), and CLM-CASA' has 16 plant functional types. We aggregated the CLM-CASA' output from plant functional types to the scale of biomes as defined in CABLE. The aggregation of CLM-CASA' plants functional types into CABLE biomes are described in supplementary material for this paper. Furthermore, the photosynthetic parameters, rate of carboxylation (Vcmax) and specific leaf areas (SLA) were taken from the input files included in models' packages. The preset value of Q10 in CABLE was 1.72, 14 % lower than the Q10 value used in CLM-CASA'. The Q10 plays an important role in determining the temperature sensitivity of soil respiration (Zhou et al., 2009).

2.2 Mathematical description of carbon cycle and traceability framework

The carbon cycle in most models share four common properties: (1) photosynthesis as the starting point of carbon flow in an ecosystem, (2) partitioning of assimilated carbon into different vegetation components, (3) carbon transfer is controlled by donor pool, and, (4) first order decay of litter and soil organic matter. These fundamental properties of the terrestrial carbon cycle can be described using the following equation (Luo *et al.*, 2003; Luo & Weng, 2011).

$$\frac{dX(t)}{dt} = \boldsymbol{B}U(t) - \boldsymbol{A}(\boldsymbol{\xi}(E)\boldsymbol{C})X(t) \qquad (1)$$

Where, $X(t) = (X_1(t), X_2(t), \ldots, X_n(t))^T$ is a vector of length n representing the carbon pool sizes. B is an $n \times 1$ vector representing the partitioning coefficients of the photosynthetically fixed carbon into plant pools. $U(t)$ is the photosynthetically fixed carbon (NPP). $\boldsymbol{A}$ is an $n \times n$ matrix representing the carbon transfer between pools. $\boldsymbol{\xi}(E)$ is an $n \times n$ diagonal matrix of environmental scalars representing the effects temperature and moisture on decomposition rates. $\boldsymbol{C}$ is an $n \times n$ diagonal matrix representing the carbon losses through respiration at each time step.

The mutually independent properties of all these elements (B, $\boldsymbol{A}$, $\boldsymbol{C}$ *and* $\boldsymbol{\xi}(E)$) enable us to implement the analytical framework by decomposing the total ecosystem carbon storage capacity into its traceable components as described in Xia, *et al.* (2013). The elements in $\boldsymbol{\xi}(E)$and $U(t)$ in equation (1) vary with time and climatic conditions, but their long-term averages can be used to calculate steady state carbon pool sizes, X_{ss}, by letting equation (1) equal zero for a given U_{ss} and ξ_{ss}, as described in Xia et al. (2013):

$$X_{ss} = [\boldsymbol{A}\boldsymbol{\xi}_{\boldsymbol{ss}}\boldsymbol{C})]^{-1}\boldsymbol{B}U_{SS} \qquad (2)$$

The vector X_{ss} represents the steady state carbon pools. U_{ss} is the steady state carbon influx in an ecosystem. The partitioning (***B***), transfer coefficients and respirational losses (***A*** and ***C***) in equation (2) together determine the baseline carbon residence time (τ'_E):

$$\tau_E^{'} = (\boldsymbol{AC})^{-1}\boldsymbol{B} \qquad (3)$$

The baseline carbon residence time (τ'_E) in equation (3) and environmental scalar values describe the total ecosystem residence time (τ_E):

$$\tau_E = \ \boldsymbol{\xi}_{\boldsymbol{ss}}{}^{-1}\tau_E^{'} \qquad (4)$$

Thus the ecosystem carbon storage capacity is jointly determined by the ecosystem residence time (τ_E) and steady state carbon influx (U_{ss}):

$$X_{ss} = \quad \tau_E\, U_{ss} \qquad (5)$$

Equation (5) also defines the total ecosystem residence time as the ratio of carbon storage (Xss) to steady state carbon influx (Uss) ($\tau_E = Xss/Uss$)

The environmental scalar is further separated into the temperature ($\boldsymbol{\xi_T}$) and water ($\boldsymbol{\xi_W}$) scalar components which can be represented as:

$$\xi_{ss} = \quad \xi_W\, \xi_T \qquad (6)$$

The set of equations (2-6) not only decomposes the carbon storage capacity into different traceable components in a systematic way, but also explains the mutual relationships among them. The additional information on the description of traceability components can be found at http://ecolab.ou.edu/?research_info&id=36.

2.3 Model simulations and diagnosis

Modeled carbon dynamics heavily depends on the initial conditions of state variables (carbon pools), which, in land models, are customarily assumed to be steady state pools (in the year 1850). In this study, for the estimation of modeled carbon storage capacity and other traceable components, the steady state of the models was obtained through spin up simulations. The process of spin up was carried out using the semi analytical solution (SAS) method developed by Xia et al. (2012). For spin up, the models were simulated until the mean changes in carbon pools over each loop (1 year) were smaller than 0.01 % per year in each cycle. The CLM-CASA and CABLE models were forced with the climate forcing data reported in Qian et al. (2006) and Wang et al. (2010), respectively. The CO_2 concentration was set at 375 ppm for both models' runs. Inputs for soil texture in both models were taken from IGBP-DIS dataset (IGBP-DIS, 2000). For both models, the lignin content and CN ratios were assigned for each plant functional type in the source code (therefore there was no map of them) and lignin to nitrogen ratios were calculated from PFT-level CN ratios and lignin content. The models were run on two spatial resolutions of 2.81^o x 2.81^o (CLM-CASA') and 1^o x 1^o (CABLE). After the spin up simulations, elements of $\boldsymbol{A}$, $\boldsymbol{C}$, $\boldsymbol{B}$, and $\boldsymbol{\xi(E)}$, *as well as* $U(t)$ were stored to calculate their mean values. The obtained averages were used to calculate the carbon residence time and steady state carbon pools (Eqs. 2-4).

3.0 Results

3.1 Carbon storage in CABLE and CLM-CASA'

The ecosystem carbon storage capacity differed substantially between CABLE and CLM-CASA' at both global and biome level. CLM-CASA' had 31 % higher global carbon storage capacity compared to CABLE (Circled in Fig. 1). In both models, evergreen needleleaf forest and evergreen broadleaf forest showed the highest carbon storage capacity. However, evergreen needleleaf forest and evergreen broadleaf forest in CLM-CASA' had 63 % and 47 % higher carbon storage capacity compared to respective biomes in CABLE. Shrub land, C3G and C4G showed the most agreement between two models. A substantial variation was observed in the simulated NPP and estimated ecosystem residence time at both global and biome level between CABLE and CLM-CASA'. All biomes in CLM-CASA' produced higher NPP compared to the respective biomes in CABLE. The minimum value of NPP (250 g C m^{-2} yr^{-1} for deciduous needleleaf forest) in CLM-CASA' was much higher than the minimum value of NPP (61 g C m^{-2} yr^{-1} for tundra) in CABLE. A similar diverse trend was also observed for the ecosystem residence time. In CLM-CASA', three biomes (deciduous needleleaf forest, evergreen needleleaf forest and tundra) showed ecosystem residence time of >100 years compared to CABLE. However, C4G in both models represented the shortest ecosystem residence time in CLM-CASA' (13 years) and CABLE (18 years).

3.2 Baseline carbon residence time and its components

Both CABLE and CLM-CASA' showed large variations in baseline carbon residence times at both global and biome level (Fig. 2). The global baseline residence time of 20 years in CABLE was approximately five fold larger than the global baseline carbon residence time of

CLM-CASA’. The deciduous needleleaf forest and evergreen needleleaf forest in both models showed the highest baseline carbon residence times. The tundra in CABLE showed the minimum baseline carbon residence time, whereas, it was ranked third highest in CLM-CASA’. Similarly, the baseline carbon residence time of shrub land in CABLE was 89 % higher than the baseline carbon residence time of tundra in CLM-CASA’. In general, five biomes (evergreen needleleaf forest, evergreen broadleaf forest, deciduous needleleaf forest, deciduous broadleaf forest, shrub land) in CABLE showed baseline residence times of >15 years compared to the maximum baseline carbon residence time of 9 years for deciduous needleleaf forest in CLM-CASA’.

The baseline carbon residence time is dependent on NPP partitioning coefficients (vector B), carbon transfer coefficients (matrix $\boldsymbol{A}$) and decomposition rates (matrix $\boldsymbol{C}$) (Eq. 4). All these components of $\boldsymbol{B}$, $\boldsymbol{A}$, and $\boldsymbol{C}$ showed substantial differences between the two models. CABLE allocated 61 % of NPP to roots, 23 % to wood and 16 % to leaves (Fig. 3A). CLM-CASA’ allocated 43 % of NPP to leaves, 16 % to wood and 41 % to roots (Fig. 3B). Similarly, a large difference in carbon transfers from live plants to litter and soil was also observed. In CABLE, the live tissues were partitioned into three litter pools (including CWD). 59 % of leaf carbon partitioned to metabolic litter and 41% to structural litter pools, while roots transferred 61 % of their carbon to metabolic and 39 % to structural litter. A major portion of litter carbon was released into the atmosphere through respiration losses, while the remaining was transferred into the soil organic matter pools (Fig. 3A). In CLM-CASA’, the plant tissues dispersed to six litter pools (including CWD) after mortality. The leaves allocated 62 % of its carbon to surface metabolic litter and 38 % to surface structural litter. Likewise, the fine roots allocated 62 % of its carbon to soil metabolic litter and 38 % to soil structural litter. All of the litter pools contributed to three soil carbon pools which were then interlinked for back and forth movement of carbon until it was respired completely (Fig. 3B). CLM-CASA' and CABLE also differed in representing their ***C*** matrix which was a fraction of carbon leaving from each pool with values in CLM-CASA' being higher than in CABLE, in general.

3.3 Photosynthetic parameters

The magnitude of NPP is one of the two factors that control ecosystem carbon storage capacity in CLM-CASA' and CABLE. Differences in NPP between the two models could've been caused by differences in model structures, forcing, and in model parameterization of photosynthesis process. As illustrated in Figure 4, there were no significant differences in models' climatic forcing, whereas, photosynthetic parameters differed substantially. For most biomes CLM-CASA' had higher Vcmax and SLA values (Table 1), which caused the NPP to be higher than in CABLE. However, NPP simulated by CLM-CASA' was higher than NPP simulated by CABLE for all biomes, therefore differences in the photosynthetic model formulations were likely the most significant contributor to the differences in NPP between the two models .

3.4 Climate forcing data

The mean air temperature (11.2 ± 4.9 ^{o}C) and precipitation (973 ± 457 mm) in CABLE was comparable to mean air temperature (11.7 ± 5.1 ^{o}C) and precipitation (967 ± 490 mm) in CLM-CASA' (Fig. 4). A strong agreement between climate forcing was also observed between the biomes of both models. However, a few biomes showed substantial variations in climate forcing between CABLE and CLM-CASA'. The maximum difference between mean air temperatures of both models was observed for deciduous broad leaf forest followed by tundra and deciduous needleleaf forest, respectively (Fig 4). CLM-CASA' showed 18 % higher mean air temperature for deciduous broad leaf forest compared to CABLE. In both models, tundra (-8.0 ± 5.2 ^{o}C in CABLE; -5.5 ± 5.2 ^{o}C in CLM-CASA') and deciduous needleleaf forest (-7.0 ± 1.4 ^{o}C in CABLE; -9.8 ± 1.2 ^{o}C in CLM-CASA') showed much lower air temperature compared to all other biomes. The maximum differences in precipitation data between both models were found in C4G, tundra and deciduous needleleaf forest respectively. In CABLE, C4G (1018 ± 491 mm) presented 59 % lower precipitation compared to C4G (1622 ± 765 mm) in CLM-CASA'. However, CABLE exhibited 46 % and 43 % more precipitation for tundra and deciduous needleleaf forest, respectively, compared to that of comparable biomes in CLM-CASA'.

3.5 Environmental scalars

The lower environmental scalar limits decomposition rates and turnover time result in increases of the final ecosystem residence time. The environmental scalars at global and biome level differed substantially between two models (Fig 5). The global average of environmental scalar in CABLE (0.34) was considerably lower compared to that of CLM-CASA' (0.42). In general, CLM-CASA' simulated higher environmental scalar values for most of the biomes compared to CABLE. C4G, shrub land and evergreen broadleaf forest were least limited by temperature and moisture with environmental scalars of 0.65 and 0.49, respectively. Both models simulated tundra with the highest temperature and moisture limitation of organic matter decomposition.

The global temperature and water scalars in CLM-CASA' were found to be 16 % and 4 % higher than that of CABLE. The temperature scalars were strongly dependent on the Q10 value, which was 14 % higher in CLM-CASA' than in CABLE. The C4G, evergreen broadleaf forest and shrubs in CABLE and C4G, shrubs and evergreen broadleaf forest in CLM-CASA', respectively, showed the highest temperature scalar values amongst all other biomes, (Fig. 5). The minimum temperature scalar was observed for tundra in both CABLE and CLM-CASA'. Overall, organic matter decomposition (across the biomes) in CABLE was more dependent on temperature than the organic matter decomposition in CLM-CASA'. The same diverse pattern of biome level water scalars was observed in both models (Fig. 5). The deciduous needleleaf forest (0.87) in CABLE and EBF (0.98) in CLM-CASA' showed the maximum water scalar values. Similarly, evergreen broad leaf forest (0.65) in CABLE and tundra (0.16) in CLM-CASA' showed the minimum environmental scalar values. Overall, the lowest water scalar was observed in the deciduous needleleaf forest for CLM-CASA' and the lowest temperature scalar was observed in Tundra for CABLE. In general, CLM-CASA' presented higher values of water scalars for most biomes compared to CABLE. Furthermore, environmental scalars were mainly determined by temperature rather than water scalar in both models.

4.0 Discussion

The traceability framework implemented in this study is an effective method to characterize the major components of the carbon cycle represented by two widely used land models, CABLE and CLM-CASA'. We were able to identify the differences in modeled carbon storage capacity in an independent manner through decomposing of the carbon cycle into its major components of NPP, ecosystem residence time and environmental scalars (Eq. 1-6). For example, the global carbon storage capacity in CLM-CASA' was substantially higher (31%) compared to that in CABLE, primarily due to 37% higher simulated NPP slightly offset by lower ecosystem residence time (Fig. 1 and Fig. 6). The higher NPP in CLM-CASA' was partly attributed to the relatively higher rates of carboxylation and specific leaf areas (Table 1) compared to CABLE, but for half of the biomes, the cause of differences in NPP between the two models was not straightforward, and might have been a combination of models formulation and assumptions about autotrophic respiration (Kowalczyk *et al.*, 2006; Oleson, *et al.*, 2004).

Both models showed a distinctive pattern of NPP partitioning and transferring carbon among different pools (Fig. 3) which resulted in different baseline carbon residence times. The baseline carbon residence time in CABLE was longer due to more NPP partitioning into roots and wood, which had higher residence times than in CLM-CASA'. In biomes, deciduous needleleaf and evergreen needleleaf forests showed the highest baseline carbon residence times because they partitioned the largest fraction of NPP to woody biomass. For tundra the baseline residence times differed also, likely due to the partitioning coefficients, because both models simulated similar environmental scalars of 0.1. Previous studies also reported that partitioning of NPP among different pools is a significant factor in determining carbon residence time (Todd-Brown *et al.*, 2013; Rafique *et al.*, 2014). In CABLE, the allocation of NPP into plant pools was mainly driven by the availability of water, nitrogen and light (Xia et al., 2013), whereas, CLM-CASA' considers only water and light (Friedlingstein et al., 1999). CABLE and CLM-CASA' also differed significantly in transferring carbon among pools, and their corresponding respiration loss (Fig. 3). The most obvious difference was the pattern of carbon transfer from live tissues to litter pools. These carbon transfer rates among pools directly influence the carbon pool sizes and residence time (Xia *et al.*, 2013). The more complicated interactions between soil pools in CLM-CASA' slightly increase the residence time but not significantly), because instead of leaving the system, carbon returns to another pool, thus staying in the system longer (results not shown).

Environmental scalars strongly influenced the actual ecosystem residence time and varied substantially across the biomes in both models. Temperature scalars in both models showed more diverse distribution than water scalars, indicating that temperature limitation was more important in determining actual ecosystem residence time than water limitation (Todd-Brown *et al.*, 2014). However, water scalars were more variable across biomes in CLM-CASA' than in CABLE. Despite the similarity of air temperature data used in both models (Fig. 4), the temperature scalars were found to be different between the two models due to the considerable difference in Q10 value, which was higher in CLM-CASA'. Even though the annual values of two climate forcing data were not different, the climate variations in few biomes were moderately different (Fig 4). Therefore, the variation in actual ecosystem residence time by the two models, at the biome level, can also be partly due to the use of two different climate forcing data.

The traceability framework is an effective method for explaining the models variations, a major issue identified by previous studies (Friedlingstein *et al.*, 2006; Wang *et al.*, 2011; Mishra *et al.*, 2013; Todd-Brown *et al.*, 2013; Rafique *et al.*, 2014; Zaehle *et al.*, 2014). Overall, our results showed that the major factors contributing to the differences between the two models were primarily due to parameter settings related to photosynthesis, carbon input, baseline residence times and environmental conditions. This study provides information on the relative importance of model components and source of variations which are useful for model intercomparisons, benchmark analyses and evaluation of additional components in models. Hence, this framework can be applied to other biogeochemical models to better characterize and quantify the processes that contribute to model differences. For example, CLM4, VEGAS and CENTURY share similar structure of carbon cycle modules and thus can be diagnosed through the traceability framework for evaluating the models' performance.

Summary

The modeled total carbon storage capacity in CLM-CASA' was ~31% higher compared to CABLE, due to the combined effect of higher NPP and lower ecosystem residence time. The ecosystem residence time was primarily dependent on the baseline carbon residence time and environmental scalar. Both CABLE and CLM-CASA' showed large variations in baseline carbon residence times, which is largely influenced by NPP partitioning coefficients (vector ***B***), carbon transfer coefficients (matrix ***A***), and decomposition rates (matrix ***C***). The global average of environmental scalar in CABLE (0.34) was lower compared to that of CLM-CASA' (0.42). At biome level, CLM-CASA' exhibited higher environmental scalar values for most of the biomes compared to CABLE. The difference in environmental scalars between CABLE and CLM-CASA' was largely due to the differences in temperature scalars rather than water scalars. Overall, our results suggested that the differences in carbon storage between the two models were largely influenced by parameter settings related to photosynthesis, baseline residence times and temperature limitation of organic matter decomposition. The different NPP values were determined by the differences in Vcmax and SLA, while the differences in baseline carbon residence times were determined by differences in NPP partitioning and carbon transfer coefficients.

Acknowledgement

This work is financially supported by USA Department of Energy, Terrestrial Ecosystem Sciences grant DE SC0008270 and National Science Foundation (NSF) grant DEB 0743778, DEB 0840964, EPS 0919466, and EF 1137293. Partial support for final preparation of this paper was provided by a Laboratory Directed Research and Development grant by the Pacific National Laboratory which is managed by the Battelle Memorial Institute for the US Department of Energy. The data produced and used in this study can be obtained on request from Dr. Yiqi Luo (Email: yluo@ou.edu). However, the source codes for the CABLE and CLM-CASA are located at https://trac.nci.org.au/trac/cable/wiki and http://www.cgd.ucar.edu/tss/clm/distribution/clm3.5, respectively. We are thankful to Lifen Jiang, Katherine Todd-Brown, Xia Xu, Zheng Shi, Junyi Liang and Changting Wang for their suggestions and feedback in conducting this research.

References

Arora, V. K., J. F. Scinocca, G. J. Boer *et al.,*: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases, *Geophysical Research Letters,* 38, L05805, doi:10.1029/2010GL046270, 2011.

DeFries, R. S., M. C. Hansen, J. R. G. Townshend, A. C. Janetos, and T. R. Loveland:. A new global 1-km dataset of percentage tree cover derived from remote sensing. *Global Change Biol.*, 6**,**247–254, 2000b.

Doney, S. C., K. Lindsay, I. Fung and J. John:. Natural Variability in a Stable, 1000-Yr Global Coupled Climate–Carbon Cycle Simulation, *Journal of Climate,* 19, 3033-3054, 2006.

Friedlingstein, P., G. Joel, C. B. Field and I. Y. Fung:. Toward an allocation scheme for global terrestrial carbon models. *Global Change Biology,* 5, 755-770, 1999.

Global Soil Data Task, cited 2000: Global soil data products CD-ROM (IGBP-DIS). International Geosphere–Biosphere Programme—Data and Information Available Services. [Available online at http://www.daac.ornl.gov.].

Friedlingstein, P., P. Cox, R. Betts, *et al:*. Climate–Carbon Cycle Feedback Analysis: Results from the C4MIP Model Intercomparison, *Journal of Climate*, 19, 3337-3353, 2006.

Johns, T. C., J. F. Royer, I. Höschel I, *et al:*. Climate change under aggressive mitigation: the ENSEMBLES multi-model experiment, *Climate Dynamics,* 37, 1975-2003, 2011.

Kowalczyk, E. A., Y. P. Wang, R. M. Law, H. L. Davies, J. L. McGregor and G. Abramowitz:. The CSIRO atmosphere biosphere land exchange (CABLE) model for use in climate models and as an offline model. *www.cmar.csiro.au/e-print/open/kowalczykea 2006a.pdf, 2006.*

Loveland, T. R., B. C. Reed, J. F. Brown, D. O. Ohlen, Z. Zhu, L. Yang, and J. W. Merchant:. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data.Int. *J. Remote Sens*., 21, 1303–1330, 2000.

Luo, Y:. Terrestrial Carbon–Cycle Feedback to Climate Warming, *Annual Review of Ecology, Evolution, and Systematics*, 38, 683-712, 2007.

Luo, Y., and E. Weng:. Dynamic disequilibrium of the terrestrial carbon cycle under global change, *Trends in Ecology & Evolution*, 26, 96-104, 2011.

Luo, Y., R. Sherry, X. Zhou, and S. Wan:. Terrestrial carbon-cycle feedback to climate warming: experimental evidence on plant regulation and impacts of biofuel feedstock harvest, *GCB Bioenergy*, 1, 62-74, 2009.

Luo, Y., E. Weng, X. Wu, C. Gao, X. Zhou and L. Zhang:. Parameter identifiability, constraint, and equifinality in data assimilation with ecosystem models, *Ecological Applications*, 19, 571-574, 2009.

Luo, Y., L. Wu, J. A. Andrews, L. White, R. V. R. Matamala, K. Schäfer, and W. H. Schlesinger:. ELEVATED CO2 DIFFERENTIATES ECOSYSTEM CARBON PROCESSES: DECONVOLUTION ANALYSIS OF DUKE FOREST FACE DATA, *Ecological Monographs*, 71, 357-376, 2001.

Luo, Y. Q., L. W. White, J. G. Canadell, *et al*:. Sustainability of terrestrial carbon sequestration: A case study in Duke Forest with inversion approach, *Global Biogeochemical Cycles*, 17, 1021, doi:10.1029/2002GB001923, 1, 2003.

Luo, Y. Q., J. T. Randerson, G. Abramowitz, *et al*:. A framework for benchmarking land models, *Biogeosciences*, 9, 3857-3874, 2012.

Mishra, U., J. D. Jastrow, R. Matamala, *et al*:. Empirical estimates to reduce modeling uncertainties of soil organic carbon in permafrost regions: a review of recent progress and remaining challenges, *Environmental Research Letters*, 8, 035020, 2013.

Oleson, K. W., G. Y. Niu, Z. L. Yang, *et al*:. Improvements to the Community Land Model and their impact on the hydrological cycle, *Journal of Geophysical Research: Biogeosciences*, 113, G01021, doi:10.1029/2007JG000563, 2008.

Oleson, K.W., G. Y. Niu, *et al*:.Technical description of the Community Land Model (CLM3.5), *NCAR Technical Note,* National Center for Atmospheric Research, Boulder, CO, 2007.

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

Qian, T., A. Dai, K. E. Trenberth and K. W. Oleson:. Simulation of Global Land Surface Conditions from 1948 to 2004. Part I: Forcing Data and Evaluations, *Journal of Hydrometeorology*, 7, 953-975, 2006.

Rafique, R., J. Xia, O. Hararuk and Y. Luo:. Structural analysis of three global land models on carbon cycle simulations using a traceability framework, *Biogeosciences Discuss.* 11, 9979-10014, 2014.

Sitch, S., P. Friedlingstein, N. Gruber, *et al*:. Trends and drivers of regional sources and sinks of carbon dioxide over the past two decades, *Biogeosciences Discuss.* 10, 20113-20177., 2013.

Taylor KE, Stouffer RJ & Meehl GA:. An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, 93, 485-498, 2011.

Thornton, P. E., and N. E. Zimmermann:. An Improved Canopy Integration Scheme for a Land Surface Model with Prognostic Canopy Structure, *Journal of Climate*, 20, 3902-3923, 2007.

Thornton, P. E., S. C. Doney, K. Lindsay, *et al*:. Carbon-nitrogen interactions regulate climate-carbon cycle feedbacks: results from an atmosphere-ocean general circulation model, *Biogeosciences*, 6, 2099-2120, 2009.

Todd-Brown, K. E. O., J. T. Randerson, W. M. Post, F. M. Hoffman, C. Tarnocai, E. A. G. Schuur and S. D. Allison:. Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations, *Biogeosciences*, 10, 1717-1736, 2013.

Todd-Brown, K. E. O., J. T. Randerson, F. Hopkins, V. Arora, T. Hajima, C. Jones, E. Shevliakova, J. Tjiputra, E. Volodin, T. Wu, Q. Zhang, and S. D. Allison:. Changes in soil organic carbon storage predicted by Earth system models during the 21st century, *Biogeosciences*, 11, 2341-2356, 2014.

Wang, Y. P., R. M. Law, and B. Pak:. A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere, *Biogeosciences*, 7, 2261-2282, 2010.

Wang, Y. P., E. Kowalczyk, R. Leuning, *et al* :. Diagnosing errors in a land surface model (CABLE) in the time and frequency domains, *Journal of Geophysical Research: Biogeosciences*, 116, G01034, doi:10.1029/2010JG001385, 2011.

White, L., and Y. Luo:. Estimation of carbon transfer coefficients using Duke Forest free-air CO2 enrichment data, *Applied Mathematics and Computation*, 130, 101-120, 2002.

Xia, J. Y., Y. Q. Luo, Y. P. Wang and O. Hararuk:. Traceable components of terrestrial carbon storage capacity in biogeochemical models, *Global Change Biology*, 19, 2104-2116, 2013.

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

Zaehle, S., B. E. Medlyn, M. G. De Kauwe, *et al* :. Evaluation of 11 terrestrial carbon–nitrogen cycle models against observations from two temperate Free-Air CO2 Enrichment studies, *New Phytologist*, 202, 803-822, 2014.

Zhou T, PJ Shi, DF Hui and YQ Luo:. Global pattern of temperature sensitivity of soil heterogeneous respiration (Q10) and its implications for carbon-climate feedback. *Journal of Geophysical Research – Biogeosciences*. 114, G02016, doi:10.1029/2008JG000850, 2009.

Zhou, X. H., T. Zhou, and Y. Q. Luo:. Uncertainties in carbon residence time and NPP-driven carbon uptake in terrestrial ecosystems of the conterminous USA: a Bayesian approach, *Tellus*, 64, 17223, doi.org/10.3402/tellusb.v64i0.17223, 2012.

Figure captions

Figure 1

Determination of ecosystem carbon storage (kg C cm^{-2}) capacity (grey contour lines) by carbon influx (*Uss;* X-axis) and ecosystem residence time (τ_E; Y-axis) (at global and biome level) between CABLE and CLM-CASA'. The contour lines show the constant values of ecosystem carbon storage capacity. ENF – Evergreen needleleaf forest, EBF – Evergreen broadleaf forest, DNF – Deciduous needleleaf forest, DBF – Deciduous broadleaf forest, Shrub – Shrub land, C3G – C3 grassland, C4G – C4 grassland. Open squares in the circle show the global values.

Figure 2

Spatial distribution of ecosystem residence time (τ_E) and baseline carbon residence time (τ'_E) (at global and biome level) between CABLE and CLM-CASA'. Abbreviations of biomes are given in Fig 1. Circles separate the biomes of CLM-CASA' and CABLE. Open squares in the circle show the global values.

Figure 3

Schematic diagram showing the carbon cycle in CABLE (A) and CLM-CASA' (B). Carbon enters the system through photosynthesis and is partitioned among live pools. From live pools, carbon is transferred to litter pools, and from litter pools it is transferred to soil carbon pools. Values in boxes show the pools residence times. Values outside the boxes show the partitioning and transfer coefficients. The full names of the abbreviated carbon pools are coarse woody debris (CWD), structural litter (surface and soil), metabolic litter (surface and soil), surface microbial litter, soil microbial carbon, fast soil organic matter, slow, and passive soil organic matter.

Figure 4

Distribution of climate forcing data (at global and biome levels) used for CABLE and CLM-CASA' simulations. Open square show the global values. Abbreviations of biomes are given in Fig 1.

Figure 5

Determination of environmental scalars by the temperature and water scalars (at global and biome level) between CABLE and CLM-CASA'. Open squares show the global values. The contour lines show the constant value of environmental scalars. Abbreviations of biomes are given in Fig 1.

Figure 6

Schematic diagram of the traceability framework along with the summary of the results obtained in this study. The numerical values show the percentage increase between two models. Xss - ecosystem carbon storage capacity; τ_E - ecosystem carbon residence time; τ'_E - baseline carbon residence time; ξ - environmental scalar; ξ_T - temperature scalar; ξ_W - water scalar.

Table Caption

Table 1

Photosynthesis parameter values for different biomes in CLM-CASA' and CABLE. Abbreviations of biomes are given in Fig 1. The relative difference is calculated by CLM-CASA' minus CABLE and then divided by CLM-CASA'.

Figure 1

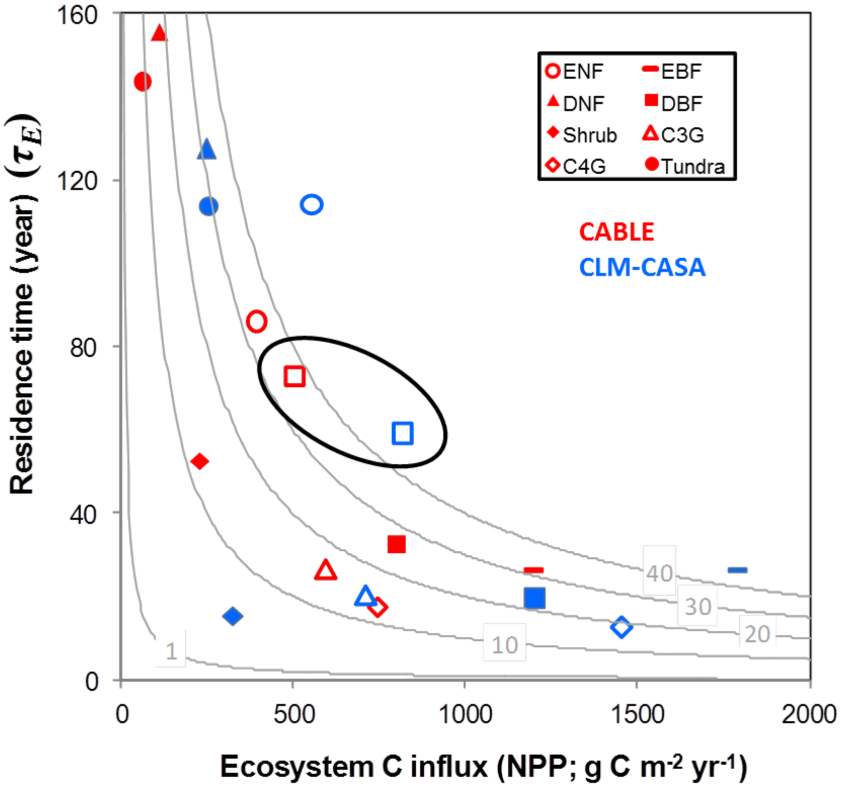

160
120
80
40
0
0
500
1000
1500
2000
Residence time (year) (τ_E)
Ecosystem C influx (NPP; g C m^{-2} yr^{-1})
○ENF
▬EBF
▲DNF
■DBF
◆Shrub
△C3G
◇C4G
●Tundra
CABLE
CLM-CASA
40
30
20
10
1

Figure 2

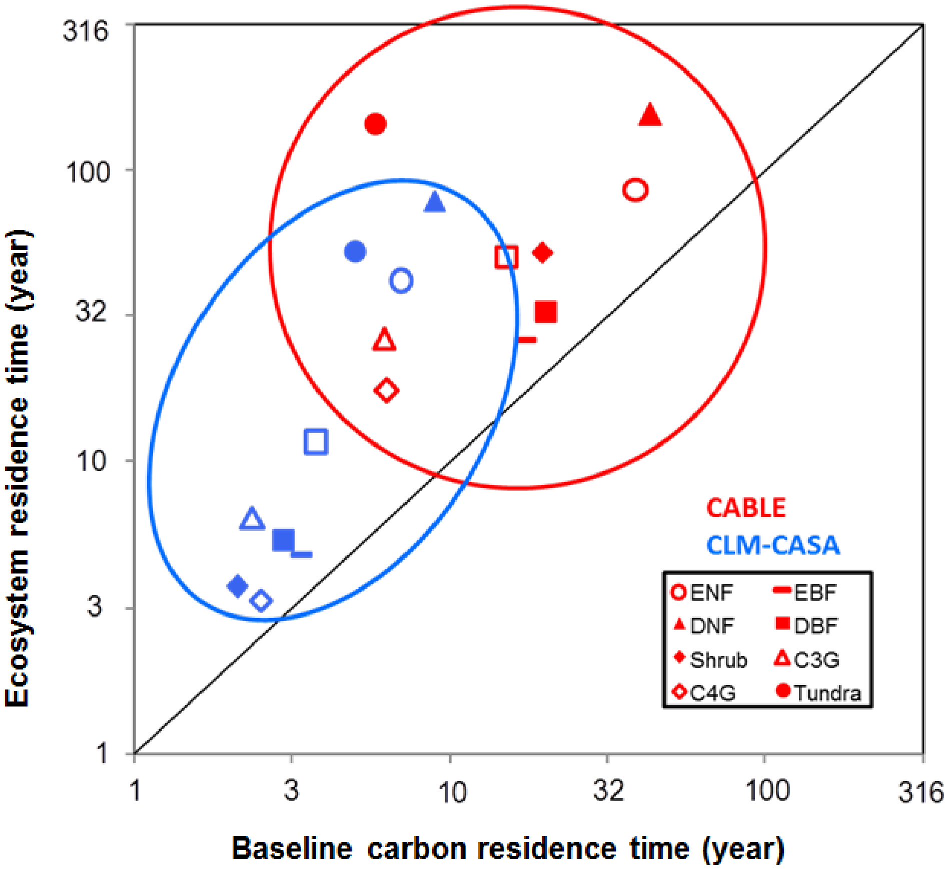

316
100
32
10
3
1
1
3
10
32
100
316
Ecosystem residence time (year)
Baseline carbon residence time (year)
CABLE
CLM-CASA
ENF
EBF
DNF
DBF
Shrub
C3G
C4G
Tundra

Figure 3

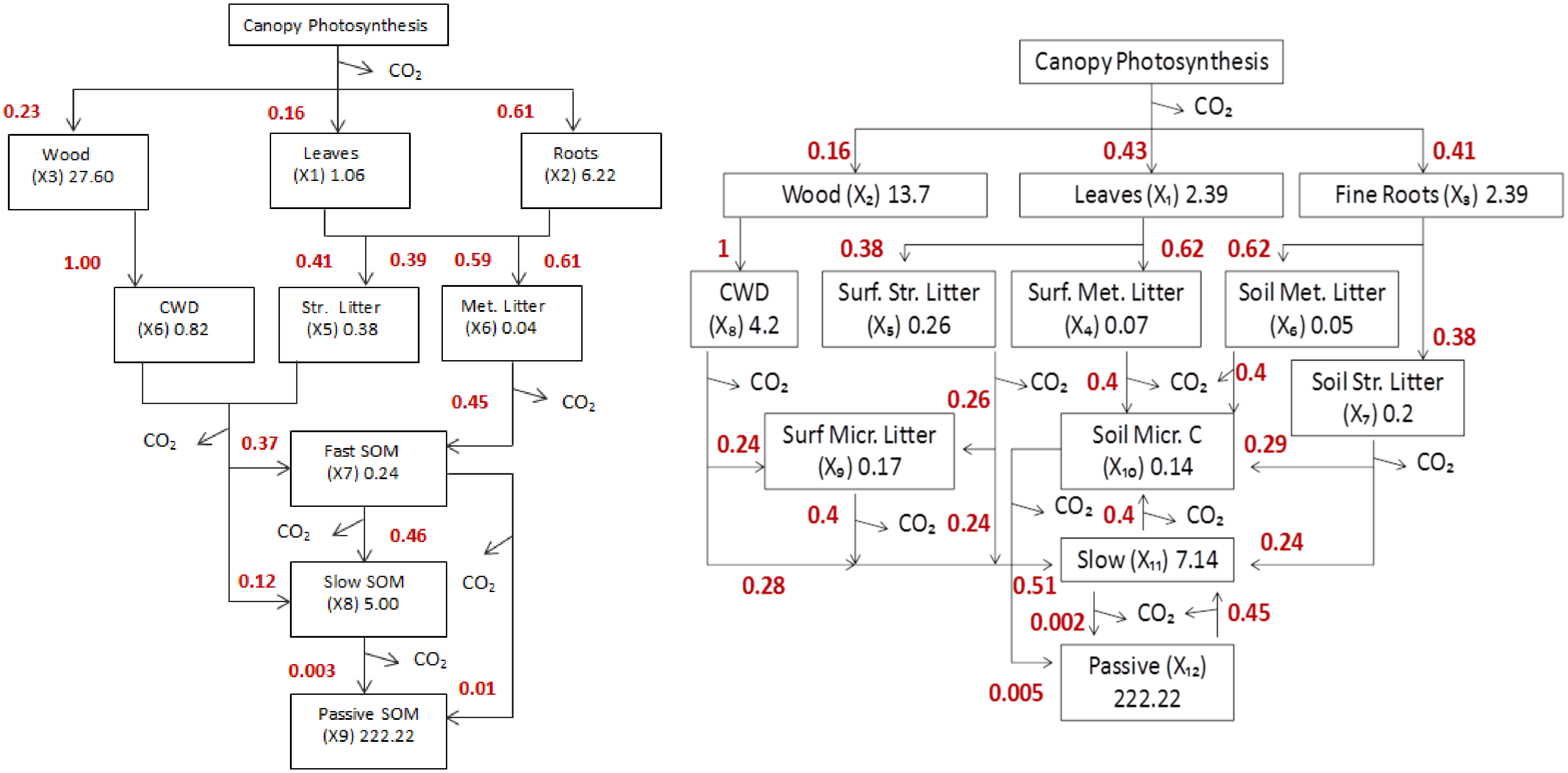

Canopy Photosynthesis
CO₂
0.23
0.16
0.61
Wood (X3) 27.60
Leaves (X1) 1.06
Roots (X2) 6.22
1.00
0.41
0.39
0.59
0.61
CWD (X6) 0.82
Str. Litter (X5) 0.38
Met. Litter (X6) 0.04
0.45
CO₂
CO₂
0.37
Fast SOM (X7) 0.24
CO₂
0.46
CO₂
0.12
Slow SOM (X8) 5.00
CO₂
0.003
0.01
Passive SOM (X9) 222.22
Canopy Photosynthesis
CO₂
0.16
0.43
0.41
Wood (X₂) 13.7
Leaves (X₁) 2.39
Fine Roots (X₃) 2.39
1
0.38
0.62
0.62
CWD (X₈) 4.2
Surf. Str. Litter (X₅) 0.26
Surf. Met. Litter (X₄) 0.07
Soil Met. Litter (X₆) 0.05
0.38
Soil Str. Litter (X₇) 0.2
CO₂
CO₂
0.4
CO₂
0.4
0.26
0.24
Surf Micr. Litter (X₉) 0.17
Soil Micr. C (X₁₀) 0.14
0.29
CO₂
CO₂
0.4
CO₂
0.4
CO₂
0.24
0.24
Slow (X₁₁) 7.14
0.28
0.51
0.002
CO₂
0.45
Passive (X₁₂) 222.22
0.005

Figure 4

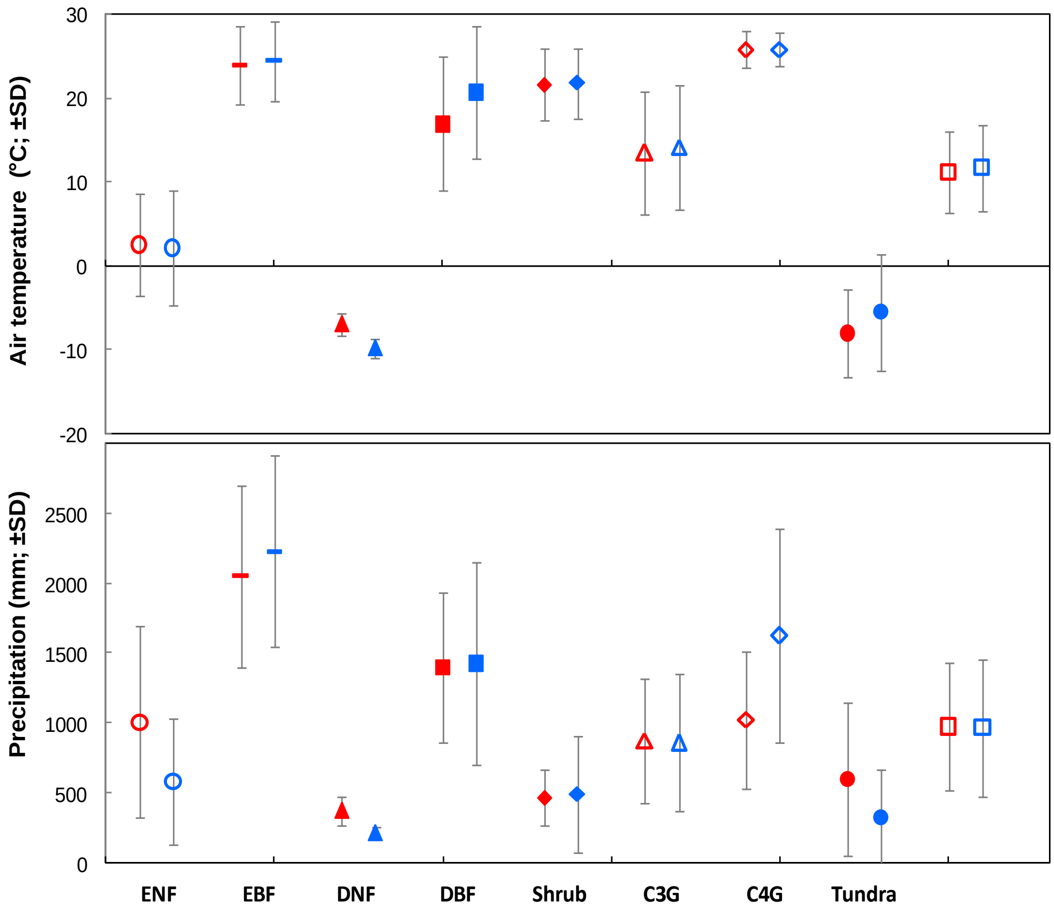
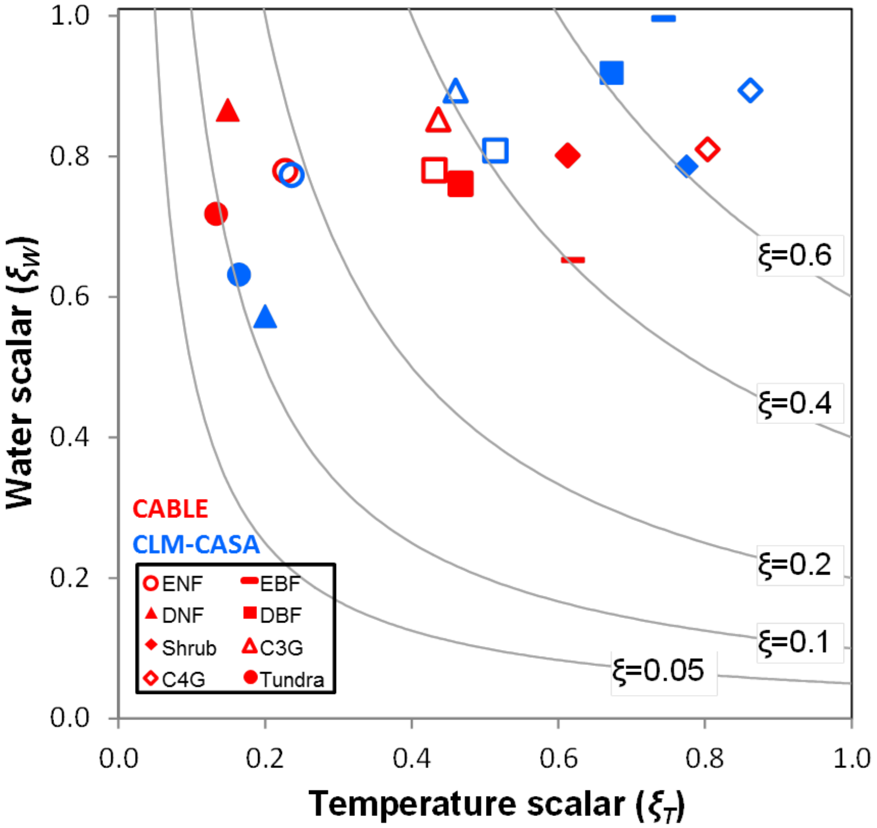

30
20
10
0
-10
-20
Air temperature (°C; ±SD)
2500
2000
1500
1000
500
0
Precipitation (mm; ±SD)
ENF
EBF
DNF
DBF
Shrub
C3G
C4G
Tundra

Figure 5



1.0
0.8
0.6
0.4
0.2
0.0
Water scalar (ξ_W)
0.0
0.2
0.4
0.6
0.8
1.0
Temperature scalar (ξ_T)
ξ=0.6
ξ=0.4
ξ=0.2
ξ=0.1
ξ=0.05
CABLE
CLM-CASA
ENF
EBF
DNF
DBF
Shrub
C3G
C4G
Tundra

Figure 6

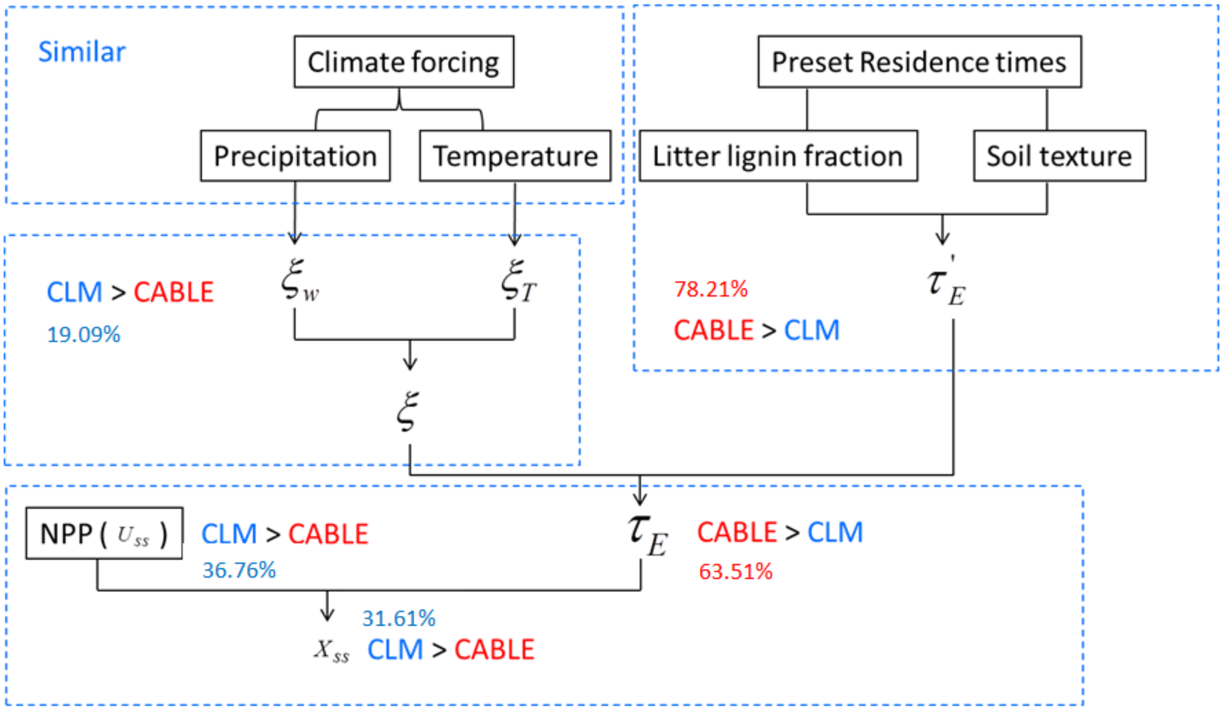

Similar
Climate forcing
Precipitation
Temperature
Preset Residence times
Litter lignin fraction
Soil texture
CLM > CABLE
19.09%
ξ_w
ξ_T
ξ
78.21%
CABLE > CLM
τ'_E
NPP (U_{ss})
CLM > CABLE
36.76%
τ_E
CABLE > CLM
63.51%
31.61%
X_{ss} CLM > CABLE

Table 1

Biomes	CLM-CASA’		CABLE		Difference (%)	
	Vcmax (umol/m^2/s)	SLA (m^2/gC)	Vcmax (umol/m^2/s)	SLA (m^2/gC)	Vcmax (umol/m^2/s)	SLA (m^2/gC)
ENF	47	0.009	40	0.018	14.90	-100
EBF	72	0.006	55	0.021	23.61	-250
DNF	51	0.024	40	0.025	21.57	-4.17
DBF	47	0.03	60	0.025	-26.76	16.67
Shrubland	22	0.024	40	0.025	-79.10	-4.17
C3G	43	0.05	60	0.028	-39.53	44
C4G	24	0.05	10	0.028	58.33	44
Tundra	43	0.05	60	0.028	-39.53	44