

# UCLA

## UCLA Previously Published Works

### Title

Sources of uncertainty in modeled land carbon storage within and across three MIPs:  
Diagnosis with three new techniques

### Permalink

<https://escholarship.org/uc/item/2ph8705w>

### Journal

Journal of Climate, 31(7)

### ISSN

0894-8755

### Authors

Zhou, Sha  
Liang, Junyi  
Lu, Xingjie  
[et al.](#)

### Publication Date

2018-04-01

### DOI

10.1175/jcli-d-17-0357.1

Peer reviewed

# Sources of Uncertainty in Modeled Land Carbon Storage within and across Three MIPs: Diagnosis with Three New Techniques

SHA ZHOU,<sup>a,p</sup> JUNYI LIANG,<sup>b</sup> XINGJIE LU,<sup>c,g,b</sup> QIANYU LI,<sup>d,b</sup> LIFEN JIANG,<sup>c,g</sup> YAO ZHANG,<sup>b,e</sup> CHRISTOPHER R. SCHWALM,<sup>f,g</sup> JOSHUA B. FISHER,<sup>h</sup> JERRY TJIPUTRA,<sup>i</sup> STEPHEN SITCH,<sup>j</sup> ANDERS AHLSTRÖM,<sup>k,l</sup> DEBORAH N. HUNTZINGER,<sup>m,n</sup> YUEFEI HUANG,<sup>a,o</sup> GUANGQIAN WANG,<sup>a</sup> AND YIQI LUO<sup>c,g,d</sup>

<sup>a</sup> State Key Laboratory of Hydrosience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing, China <sup>b</sup> Department of Microbiology and Plant Biology, University of Oklahoma, Norman, Oklahoma <sup>c</sup> Department of Biological Sciences, Northern Arizona University, Flagstaff, Arizona <sup>d</sup> Department of Earth System Science, Tsinghua University, Beijing, China <sup>e</sup> Center for Spatial Analysis, University of Oklahoma, Norman, Oklahoma <sup>f</sup> Woods Hole Research Center, Falmouth, Massachusetts <sup>g</sup> Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, Arizona <sup>h</sup> Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California <sup>i</sup> Uni Research Climate, Bjerknes Centre for Climate Research, Bergen, Norway <sup>j</sup> College of Life and Environmental Sciences, University of Exeter, Exeter, United Kingdom <sup>k</sup> Department of Earth System Science, Stanford University, Stanford, California <sup>l</sup> Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden <sup>m</sup> School of Earth Sciences and Environmental Sustainability, Northern Arizona University, Flagstaff, Arizona <sup>n</sup> Department of Civil Engineering, Construction Management, and Environmental Engineering, Northern Arizona University, Flagstaff, Arizona <sup>o</sup> College of Ecological and Environmental Engineering, Qinghai University, Qinghai, China <sup>p</sup> Current affiliation: Department of Earth and Environmental Engineering, Columbia University, New York, New York.

Corresponding author: Sha Zhou, sz2766@columbia.edu; Yiqi Luo, yiqi.luo@nau.edu

## Abstract

Terrestrial carbon cycle models have incorporated increasingly more processes as a means to achieve more-realistic representations of ecosystem carbon cycling. Despite this, there are large cross-model variations in the simulation and projection of carbon cycling. Several model intercomparison projects (MIPs), for example, the fifth phase of the Coupled Model Intercomparison Project (CMIP5) (historical simulations), Trends in Net Land-Atmosphere Carbon Exchange (TRENDY), and Multiscale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP), have sought to understand intermodel differences. In this study, the authors developed a suite of new techniques to conduct post-MIP analysis to gain insights into uncertainty sources across 25 models in the three MIPs. First, terrestrial carbon storage dynamics were characterized by a three-dimensional (3D) model output space with coordinates of carbon residence time, net primary productivity (NPP), and carbon storage potential. The latter represents the potential of an ecosystem to lose or gain carbon. This space can be used to measure how and why model output differs. Models with a nitrogen cycle generally exhibit lower annual NPP in comparison with other models, and mostly negative carbon storage potential. Second, a transient traceability framework was used to decompose any given carbon cycle model into traceable components and identify the sources of model differences. The carbon residence time (or NPP) was traced to baseline carbon residence time

(or baseline NPP related to the maximum carbon input), environmental scalars, and climate forcing. Third, by applying a variance decomposition method, the authors show that the intermodel differences in carbon storage can be mainly attributed to the baseline carbon residence time and baseline NPP (>90% in the three MIPs). The three techniques developed in this study offer a novel approach to gain more insight from existing MIPs and can point out directions for future MIPs. Since this study is conducted at the global scale for an overview on intermodel differences, future studies should focus more on regional analysis to identify the sources of uncertainties and improve models at the specified mechanism level.

Keywords: Carbon cycle; Land surface model; Model evaluation/performance

## 1. Introduction

To better understand the past, present, and future role of the terrestrial biosphere in the global carbon cycle, terrestrial carbon cycle models have become increasingly complex. These models are continuously developed and updated based on improved understanding of mechanisms controlling the carbon cycle, such as the improvement of the carbon-nitrogen cycling and dynamic global vegetation from earlier Community Land Model (CLM) to current CLM4.5 (Oleson et al. 2013). Compared to the Coupled Climate-Carbon Cycle Model Intercomparison Project (C<sup>4</sup>MIP) (Friedlingstein et al. 2006), the fifth phase of the Coupled Model Intercomparison Project (CMIP5) comprises models that include improved processes, components, or forcing (Knutti and Sedláček 2013; Taylor et al. 2012). However, large uncertainties remain in the simulation and prediction of carbon uptake and storage among different models. The simulated global soil carbon varied by 5.9-fold across 11 models from CMIP5, resulting from the difference in the simulated net primary productivity and the parameterization of soil heterotrophic respiration (Todd-Brown et al. 2013), as well as the integration over long spinup procedures (Exbrayat et al. 2014). To improve future projection of carbon storage dynamics and constrain its uncertainties, it is essential to understand the underlying key mechanisms of the global carbon cycle.

Model intercomparison studies have been conducted to assess differences between model output and explain the uncertainties among models (Fisher et al. 2014; Friend et al. 2014; Nishina et al. 2014, 2015). Schwalm et al. (2010) examined the ability of 22 terrestrial biosphere models to simulate the seasonal variability in biosphere-atmosphere exchange of CO<sub>2</sub> using data from 44 flux tower sites. Model performance was generally poor, and a large divergence between observations and simulations (~10 times observational error) was found, especially for nonforested sites. Keenan et al. (2012) compared the interannual variability of CO<sub>2</sub> exchange from 16 terrestrial biosphere models against 11 long-term eddy-covariance forest sites in North America. They found that the large biases in the modeled interannual variability are related to the poor representation of spring phenology, soil thaw and snowpack melting, and the lagged response to

extreme climatic events. Ichii et al. (2010) showed that the terrestrial biosphere models, which have been calibrated using eddy flux data, can successfully capture the seasonal and interannual variations in the terrestrial carbon cycle, indicating that the eddy flux observations are critical to improve model simulations and reduce uncertainties. Although observations can evaluate model performance and constrain model uncertainties to a certain degree, the sources of uncertainties among models are still hard to quantify (De Kauwe et al. 2014).

Several model intercomparison projects (MIPs) have been established to identify the sources of model uncertainties and improve process representation in models. CMIP provides a standard experiment protocol to evaluate output from coupled ocean-atmosphere-cryosphere-land general circulation models (Meehl et al. 2005). One of the most important targets of CMIP5 is to assess the mechanisms responsible for the spread in model projections when the same set of “external” forcing, such as greenhouse gas forcing in historical simulations, is used (Taylor et al. 2012). The Trends in Net Land-Atmosphere Carbon Exchange (TRENDY) (Sitch et al. 2015) and Multiscale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP) (Huntzinger et al. 2013; Wei et al. 2014) projects facilitate the comparison of model output by using prescribed environmental and meteorological drivers shared among all models. Thus, the role of model structure and parameters in the uncertainty of land-atmosphere carbon exchange can be systematically evaluated. For instance, by investigating the difference in model output under a series of common scenarios, the contribution of environmental drivers (e.g., changing CO<sub>2</sub>, climate, nitrogen, and land use) to the trend and variability of carbon exchange can be diagnosed (Ahlström et al. 2012; Nishina et al. 2015). Despite the fact that these projects can evaluate the impact of environmental drivers on carbon storage based on sensitivity simulations, they have led to little understanding of the underlying mechanisms of carbon storage variations across different models. Even so, the simulated terrestrial carbon storage dynamics from these MIPs can be used to identify the sources of intermodel differences.

Based on the biogeochemical principles of the terrestrial carbon cycle, Xia et al. (2013) proposed a framework to decompose a complex carbon cycle model into traceable components. In the framework, the modeled ecosystem carbon storage capacity is decomposed into the product of carbon residence time and net primary productivity (NPP). The carbon residence time refers to the mean duration of carbon in terrestrial ecosystems from its input via photosynthesis to its release via respiration (Luo et al. 2003). A three-dimensional (3D) model output space proposed by Luo et al. (2017) extends the approach of Xia et al. (2013) by involving carbon storage potential to represent the difference between carbon storage capacity and carbon storage itself. The 3D model output space can be used to evaluate the terrestrial carbon storage dynamics by decomposing the carbon storage into carbon residence time, NPP, and carbon storage potential (Jiang et al. 2018).

Thus, the simulated terrestrial carbon storage can be placed into the 3D model output space to attribute differences in model outputs to the three variables.

The three variables can be further decomposed into their traceable components to track the sources of model uncertainty. The traceability framework developed by Xia et al. (2013) suggested that the carbon residence time can be traced to 1) baseline carbon residence time, which is related to vegetation characteristics and soil types, 2) environmental scalar, including temperature and water scalars, and 3) climate variables, such as temperature and precipitation. The baseline carbon residence time is inversely related to the maximum decomposition rate, which is modified by temperature and moisture conditions. The environmental scalar expressed as a function of environmental variables, such as temperature and precipitation, links the baseline carbon residence time to actual carbon residence time. This framework decomposes carbon residence time into its traceable components; however, the traceability analysis for NPP has not been performed.

In terrestrial carbon cycle models, NPP is generally estimated using two basic approaches. Most models, such as BIOME-BGC (Running and Hunt 1993) and HYBRID (Friend et al. 1997), estimate NPP as the difference between gross primary productivity (GPP) and autotrophic respiration ( $R_a$ ), while the others directly simulate NPP as influenced by vegetation and environmental variables, such as CASA (Potter et al. 1993) and CENTURY (Parton 1996). Despite different representations of physical and biological processes in different models, the concept of light use efficiency (LUE) underpins the simulation of NPP across most models (Cramer et al. 1999). That is, NPP (or GPP) can be expressed as the product of LUE, photosynthetically active radiation (PAR), and the fraction of PAR absorbed by vegetation (fPAR) (Zhang et al. 2016). LUE is regulated by climatic conditions (e.g., temperature and precipitation), and fPAR is the most important vegetation characteristics in controlling potential photosynthetic capacity of vegetation (Schloss et al. 1999). Thus, NPP can be traced to a “baseline NPP,” which is related to vegetation characteristics and an environmental scalar determined by environmental variables, in a similar fashion to carbon residence time. The baseline NPP corresponds to the maximum carbon input when the environmental conditions are favorable for carbon assimilation. The environmental stress on NPP is evaluated by the environmental scalar, which converts the baseline NPP to actual NPP. Following this traceability analysis, the variation in terrestrial carbon storage can be quantitatively attributed to its sources to evaluate the intermodel differences based on variance decomposition.

The objective of this study is to compare the annual carbon storage simulated by different models in the three MIPs (i.e., CMIP5, TRENDY and MstMIP), based on the 3D model output space, and identify the sources of carbon storage variation using a transient traceability framework and a

variance decomposition method. First, the terrestrial carbon storage is decomposed into the 3D model output space: carbon residence time, NPP, and carbon storage potential. Second, a transient traceability framework of carbon storage dynamics is proposed to determine what controls the carbon cycle dynamics (e.g., climate factors such as temperature and precipitation). Following the transient traceability framework, the sources of the variation in carbon storage dynamics will be diagnosed. Third, the variation in carbon storage simulations is attributed to its sources by quantifying the relative contributions of them using the variance decomposition method. Our rigorous framework for multimodel assessment facilitates better understanding of the complex behaviors of various terrestrial carbon cycle models and is suggested to be a valuable evaluation method for future model intercomparison projects.

## 2. Methods and materials

### *a. Carbon storage dynamics decomposition*

Based on mathematical analysis of the matrix equation for terrestrial carbon cycle models, Luo et al. (2017) developed a 3D model output space to assess the dynamics of terrestrial carbon storage  $X$ . By decomposing the carbon storage dynamics into three variables, we can better evaluate the responses of terrestrial carbon storage to environmental factors and the capability of ecosystem processes to influence the carbon storage change.

The magnitude and direction of carbon storage change are controlled by the carbon storage capacity  $X_c$ , that is, the capacity of an ecosystem to store carbon or the carbon storage at steady state under current conditions (see Table 1 for symbol definition). The  $X_c$  is jointly determined by the carbon residence time  $\tau_E$  and ecosystem carbon input, for example, NPP (Xia et al. 2013):

$$X_c = \tau_E \times \text{NPP}. \quad (1)$$

Luo et al. (2017) further proposed that the capability of the terrestrial carbon cycle to influence carbon storage can be evaluated by carbon storage potential  $X_p$ , the potential of an ecosystem to store additional carbon or lose carbon. The  $X_p$  is proportional to the rate of carbon storage change  $X'$  and regulated by the chasing time  $\tau_{ch}$ :

$$X_p = \tau_{ch} X'. \quad (2)$$

The  $\tau_{ch}$  is a nonnegative matrix of carbon residence times through the network of individual pools. The  $X_p$  is positive (negative) when the carbon storage capacity is larger (smaller) than current carbon storage. Positive  $X_p$  values indicate an increasing trend of carbon storage, and vice versa. The larger the carbon storage potential, the faster the rate of carbon storage change, and the ratio between the two is determined by the chasing time.

TABLE 1. Symbols and parameters used in this study.

Symbol or parameter	Definition	Unit
$X$	Carbon storage	PgC
$X_c$	Carbon storage capacity	PgC
$X_p$	Carbon storage potential	PgC
$X'$	Rate of carbon storage change	PgC yr <sup>-1</sup>
NPP	Net primary productivity	PgC yr <sup>-1</sup>
$\tau_E$	Carbon residence time	yr
$\tau_{ch}$	Chasing time	yr
$\xi$	Environmental scalar for carbon residence time	—
$\delta$	Environmental scalar for NPP	—
$\tau'_E$	Baseline carbon residence time	yr
NPP'	Baseline NPP	PgC yr <sup>-1</sup>
$\xi_T$	Temperature scalar for carbon residence time	—
$\xi_W$	Water scalar for carbon residence time	—
$\delta_T$	Temperature scalar for NPP	—
$\delta_W$	Water scalar for NPP	—
$Q_{10}$	Respiration temperature sensitivity	—
$T_0$	Reference temperature	°C
$T$	Mean annual temperature	°C
$W_0$	Reference precipitation	mm yr <sup>-1</sup>
$W$	Annual total precipitation	mm yr <sup>-1</sup>

Carbon storage can be expressed as the difference between carbon storage capacity [Eq. (1)] and carbon storage potential (Luo et al. 2017). Thus, dynamics in terrestrial carbon storage can be projected into a 3D model output space:  $\tau_E$ , NPP, and  $X_p$  as

$$X = \tau_E \times \text{NPP} - X_p, \quad (3)$$

The 3D model output space above offers a new framework to quantify differences across land carbon cycle models. Thus, it helps us better understand complex model dynamics, diagnose sources of model differences, and improve model predictive capability.

To apply the 3D model output space to those three MIPs, we consider the terrestrial ecosystem as one pool. According to Luo et al. (2017),  $\tau_{ch}$  equals  $\tau_E$  when there is only one pool. Thus, Eq. (3) can be transformed into

$$X = \tau_E (\text{NPP} - X'), \quad (4)$$

where  $(\text{NPP} - X')$  represents the total carbon losses from the terrestrial ecosystem, mainly through heterotrophic respiration. Generally,  $X$  and NPP are directly available from model output,  $X'$  can be calculated as the difference of carbon storage between time step  $(t + 1)$  and  $t$ . So the carbon residence time and carbon storage potential can be calculated as follows:

$$\tau_E = \frac{X}{\text{NPP} - X'}, \quad \text{and} \quad (5a)$$

$$X_p = \tau_E \times \text{NPP} - X. \quad (5b)$$

Although the structure varies in different models, a one-pool model can effectively estimate the three variables. For example, we have reproduced the model output in CanESM2 and CESM1(BGC) using a five-pool model and found that the derived  $\tau_E$  is close to that calculated using the one-pool model (see Fig. S1 in the online supplemental material).

*b. Traceability analysis of carbon storage dynamics*

Xia et al. (2013) developed a framework for traceability analysis of steady-state carbon storage. This study expands the framework to transient dynamics of terrestrial carbon storage (Fig. 1), by incorporating the third dimension of carbon cycle dynamics (i.e., the carbon storage potential). This transient traceability framework can decompose the land carbon cycle into traceable components. The framework first traces the simulated terrestrial carbon storage to carbon storage capacity and potential. The former can be traced to a product of carbon residence time and NPP. The carbon residence time and NPP are further traced to 1) their baseline values, which are determined by soil properties and vegetation characteristics, 2) the environmental scalars, including temperature and water scalars, and ultimately 3) the climate forcing.

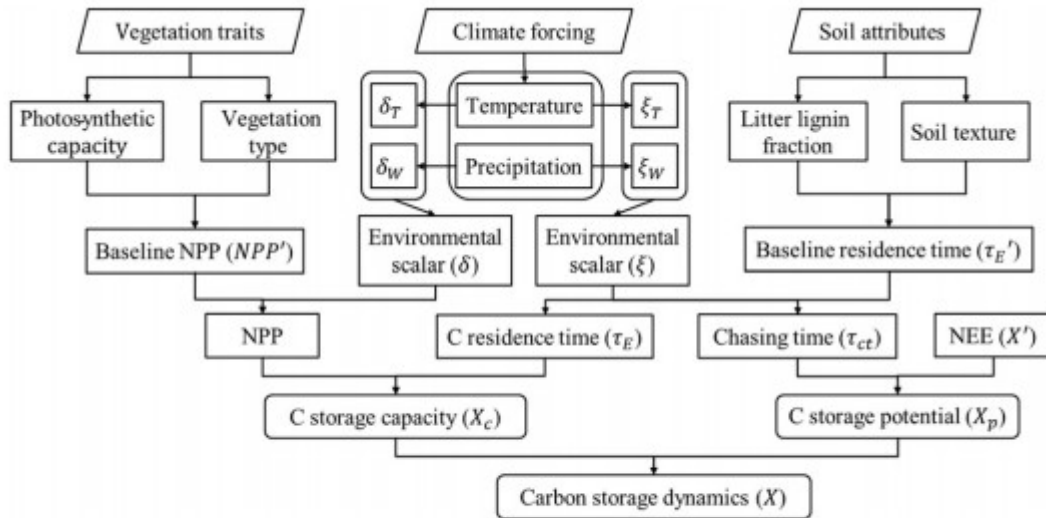


FIG. 1. Schematic diagram of the transient traceability framework. This framework traces the modeled transient carbon storage dynamics to carbon residence time, NPP, carbon storage potential, and their source factors.

1) Traceability analysis for carbon residence time and NPP

The carbon residence time is mainly related to carbon release from an ecosystem via decomposition and respiration. The maximum carbon decomposition rate corresponds to the baseline carbon residence time, under optimal temperature and moisture conditions (Xia et al. 2013). The carbon residence time is determined by the baseline carbon residence time  $\tau_E'$  and modified by the environmental scalar  $\xi$ :

$$\tau_E = \xi^{-1} \tau_E'. \quad (6)$$



The baseline carbon residence time is usually preset in a carbon cycle model, according to soil properties and vegetation characteristics (Fig. 1).

NPP has been simulated according to different processes by different models, and a large uncertainty of modeled NPP simulation still exists (Cramer et al. 1999; Schwalm et al. 2010). Almost all models simulate NPP as controlled by vegetation characteristics and regulated by climate variables (Schloss et al. 1999). We assume a “baseline NPP,” which is related to the maximum carbon input when the environmental conditions are favorable for carbon assimilation, and an environmental scalar to convert the baseline NPP to actual NPP. Thus, the modeled NPP can be traced to baseline NPP ( $NPP'$ ) and an environmental scalar  $\delta$ , just as the carbon residence time, for the sake of this analysis:

$$NPP = \delta \times NPP'. \quad (7)$$

The baseline NPP is related to vegetation characteristics, including photosynthetic capacity and vegetation type (Fig. 1).

The environmental scalar usually consists of the temperature and water scalars, which are traced to the climate forcing (i.e., temperature and precipitation) (Xia et al. 2013). The terrestrial carbon storage is affected by various environmental factors, including climate,  $CO_2$  concentration, land cover, nitrogen deposition, and so forth. In this study, we focus on the effect of climate change on the carbon storage, by investigating the responses of carbon residence time and NPP to climate forcing (i.e., temperature and precipitation). Thus, the environmental scalars  $\xi$  and  $\delta$  are further decomposed into temperature and water scalars as follows:

$$\xi = \xi_T \xi_W, \quad \text{and} \quad (8a)$$

$$\delta = \delta_T \delta_W, \quad (8b)$$

where the subscripts  $T$  and  $W$  refer to the temperature and water scalars, respectively.

To estimate the baseline residence time and baseline NPP, we use an optimization method to reproduce the simulation results of carbon residence time and NPP using annual temperature and precipitation as inputs. Here we only show the optimization method for the carbon residence time; the method for NPP is the same. In this method,  $\tau'_E$  is set to be an unknown parameter,  $\xi_T$  and  $\xi_W$  are expressed as functions of temperature  $T$  and precipitation  $W$ , respectively:

$$\xi_T = Q_{10}^{\left(\frac{T-T_0}{10}\right)}, \quad \text{and} \quad (9a)$$

$$\xi_W = \frac{W}{W_0}, \quad (9b)$$

where  $Q_{10}$  is an unknown parameter that is related to the temperature sensitivity of respiration and  $T_0$  and  $W_0$  are the reference temperature and

precipitation, which are set to be the maximum values of annual temperature and precipitation, respectively, across the study period. The two parameters,  $\tau'_E$  and  $Q_{10}$ , are calibrated by comparing the calculated  $(\xi_T \xi_W)^{-1} \tau'_E$  with the  $\tau_E$  derived from model output [Eq. (5a)], according to two indicators: the coefficient of determination  $R^2$  and the root-mean-square error (RMSE). The objective of the optimization method is to maximize  $R^2$  while minimizing RMSE:

$$\text{Max} \left( \frac{R^2}{\text{RMSE}} \right), \quad (10a)$$

$$R^2 = 1 - \frac{\sum_i [\tau_{E,i} - (\xi_{T,i} \xi_{W,i})^{-1} \tau'_{E,i}]^2}{\sum_i (\tau_{E,i} - \bar{\tau}_{E,i})^2}, \quad \text{and} \quad (10b)$$

$$\text{RMSE} = \sqrt{\frac{\sum_i [\tau_{E,i} - (\xi_{T,i} \xi_{W,i})^{-1} \tau'_{E,i}]^2}{n}}, \quad (10c)$$

where the subscript  $i$  refers to the time step, and  $n$  is the total time steps. In the optimization method, the parameters  $\tau'_E$  and  $Q_{10}$  are obtained using the generalized reduced gradient (GRG) nonlinear solving method (Drud 1985). For the optimization method for NPP, the two parameters NPP' and  $Q_{10}$  are calibrated using the same method as the carbon residence time.

## 2) Attribution analysis of carbon storage dynamics

After decomposing a complex land carbon cycle model into traceable components, we can better understand model output through attribution. Here we propose a variance decomposition method for the attribution analysis of carbon storage dynamics. This method is based on the covariance allocation principle for capital allocation, which is widely used for portfolio risk decomposition and attribution (Dhaene et al. 2012). According to the covariance allocation principle, the variance of a variable can be decomposed into the sum of the covariances of its individual components and itself (see text section S1 in the online supplemental material).

Three steps are taken to decompose the variance of terrestrial carbon storage into the contributions of the three variables and the source factors. Following the transient traceability framework (Fig. 1), the variance of terrestrial carbon storage is first decomposed into the contributions from carbon storage capacity and potential. The variance of carbon storage capacity is further decomposed into the contributions from carbon residence time and NPP. Finally, the variance of carbon residence time and NPP are decomposed into the contributions from the environmental scalars and the baseline values of them, respectively. To apply the variance decomposition method, the variable to be decomposed should be expressed as the sum of its components. Thus, we perform logarithmic transformation for the carbon storage capacity in Eq. (1), carbon residence time in Eq. (6), and NPP in Eq. (7) to separate them into several components, respectively. Details of

the variance decomposition method for terrestrial carbon storage can be found in text section S2 in the online supplemental material.

*c. The model intercomparison projects*

In this study, we compared the model output from the three MIPs (i.e., CMIP5, TRENDY-v1 and MsTMIP), based on the 3D model output space and the transient traceability framework. For the 3D model output space, carbon storage data (including carbon in the vegetation, soil, litter, and coarse woody debris pools) and NPP were obtained from the three MIPs. For each MIP, several models were selected based on the availability of model output in given historical simulations for our analysis (Table S1 in the online supplemental material). In addition, outlier models, such as SiBCASA with unrealistically strong increase in carbon storage between two continuous years, were excluded.

In CMIP5, output from nine Earth system models (ESMs) for the historical experiment covering a period from mid-nineteenth century to near present (1850 to 2005) was used. The ESMs allow us to explore the comprehensive behaviors of the Earth system through the coupling of ocean-atmosphere-land components. The land components of the ESMs differ in their representations of vegetation types, soil properties, human disturbances, and carbon and nitrogen pools, as well as their spatial resolutions (Anav et al. 2013). In addition, the nitrogen cycle is incorporated in BNU-ESM, CESM1(BGC), and NorESM1-ME, and the latter two ESMs use the same land components as CLM4. The historical simulations are forced by changing conditions that are consistent with observations, including changes in atmospheric chemical composition and land-use change (Taylor et al. 2012). Since the carbon storage and NPP output from ESMs in CMIP5 represent coupled simulations, the climate forcing used in our analysis was also obtained from the output of each ESM.

In TRENDY-v1, global simulations S2 (with historical climate, CO<sub>2</sub> fertilization) over the period 1901–2009 from nine dynamic global vegetation models (DGVMs) were used. The historical climate forcing data are taken from the combined dataset of the climatology data produced by the Climate Research Unit (CRU) and the reanalysis data from National Centers for Environmental Prediction (NCEP)–National Center for Atmospheric Research (NCAR) (Harris et al. 2014; Kalnay et al. 1996). Annual-resolution CO<sub>2</sub> data are sourced from historic atmospheric CO<sub>2</sub> from ice cores, and the National Oceanic and Atmospheric Administration (NOAA) for 1901–2009. Although the simulations S3, which use the historical land-use-change data from the History Database of the Global Environment (HYDE) (Hurtt et al. 2011), are more suitable for comparison with observations, the simulations S2 with a constant land-use mask were employed because the former simulations S3 are only available for a smaller subset of TRENDY models.

In MsTMIP, global simulations SG3 (with historical climate, CO<sub>2</sub> fertilization, land-use and land-cover change) from seven terrestrial biosphere models

(TBMs) were used. Four of the seven models are incorporated with nitrogen cycle (i.e., CLM4, CLM4VIC, ISAM, and DLEM). The global simulations were run at 0.5° spatial resolution from 1901 to 2010 (Huntzinger et al. 2013). The standardized environmental driver data are described by Wei et al. (2014) in detail. Similar to TRENDY-v1, the CRU-NCEP dataset is also used as climate forcing in MsTMIP. The atmospheric CO<sub>2</sub> concentration data for MsTMIP are prepared based on the GLOBVIEW-CO<sub>2</sub> product, fossil fuel emissions, and CO<sub>2</sub> observations at Mauna Loa and the South Pole. The land-use and land-cover change are prescribed by merging a statistic satellite-based land-cover product, with the time-varying land-use harmonization data (Wei et al. 2014).

Air temperature and precipitation in the three MIPs were also used for traceability analysis. We used the GCM forcing for individual models in CMIP5 and the CRU-NCEP data for the models in TRENDY and MsTMIP. The monthly output from each model in the three MIPs was processed following three steps. First, the components of all the carbon pools were summed as terrestrial carbon storage. Second, monthly carbon storage and NPP were aggregated into annual totals for each grid cell and then accumulated over all grid cells to calculate the global annual values, respectively. Third, global mean precipitation and temperature over land (excluding Antarctica and Greenland) for each year were similarly obtained from the monthly data at each grid. The global annual data were finally used to derive the 3D model output space and perform traceability analysis.

### 3. Results

#### *a. The 3D model output space*

Figures 2a, 2b, and 2c show the distributions of annual carbon residence time, NPP, and carbon storage potential for the models in CMIP5, TRENDY, and MsTMIP, respectively, at the global scale. The three variables together determine the simulated annual carbon storage by the models in the 3D model output space. Among the nine ESMs in CMIP5, NPP ranges from about 40 to 95 PgC yr<sup>-1</sup> and the carbon residence time from about 20 to 55 years, while the carbon storage potential varies from about -250 to 200 PgC (Fig. 2a). The range of global mean annual carbon residence time is generally small within each model, but annual NPP varies a lot for most models, with those that include nitrogen limitation [i.e., BNU-ESM, CESM1(BGC), and NorESM1-ME] showing smaller mean values (43–44 PgC yr<sup>-1</sup>) than other ESMs (60–80 PgC yr<sup>-1</sup>, see Table S2 in the online supplemental material). The carbon storage capacity (i.e., the product of carbon residence time and NPP) ranges considerably from less than 1200 PgC for CESM1(BGC) and NorESM1-ME to more than 3200 PgC for MPI-ESM-LR. For the seven TBMs in MsTMIP, the global annual carbon storage capacity also shows large variation, which is attributed to the high variability in carbon residence time and NPP (Fig. 2c). In addition, the carbon storage and carbon storage capacity are generally smaller for the four models (i.e., CLM4, CLM4VIC, ISAM, and DLEM), which incorporate a nitrogen cycle and show lower mean annual NPP (38–51

PgC yr<sup>-1</sup>) in comparison with other TBMs (50–73 PgC yr<sup>-1</sup>, see supplemental Table S2). However, the carbon residence time and NPP exhibit smaller variations across the nine DGVMs in TRENDY than those models in CMIP5 and MsTMIP, resulting in less variation in the simulated carbon storage capacity (Fig. 2b).

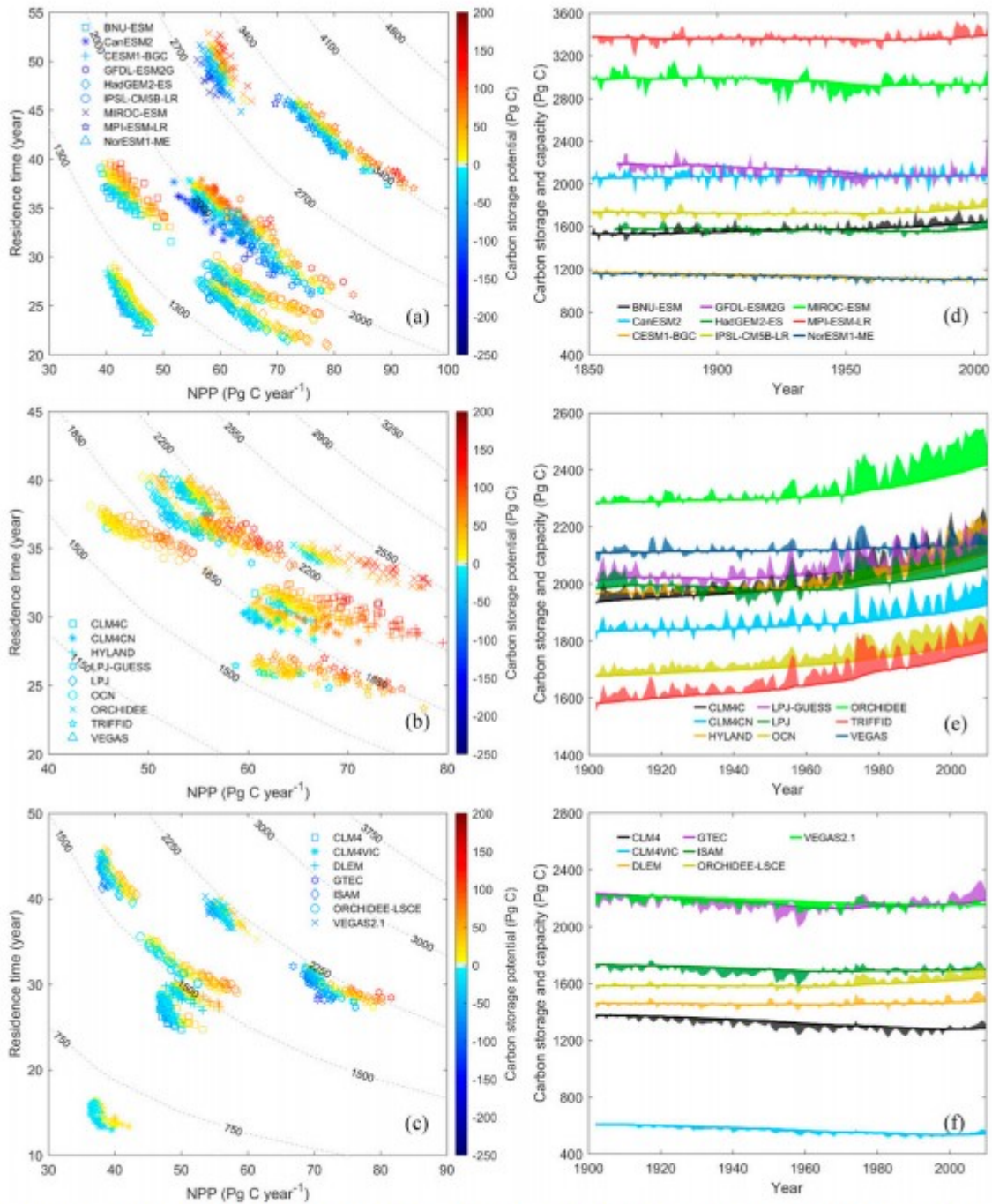


FIG. 2. The 3D model output space (carbon residence time, NPP, and carbon storage potential), and time series of annual carbon storage and carbon storage capacity (solid lines) with the shaded outlines indicating the year-to-year fluctuations due to changes in carbon storage capacity (PgC) for the models in (a),(d) CMIP5, (b),(e) TRENDY, and (c),(f) MsTMIP. The points in (a)–(c) represent the global annual values for the three variables. The contour lines in (a)–(c) represent the carbon storage capacity. The shades in (d)–(f) show the values of the carbon storage potential for the models (positive above the solid lines, and negative below the solid lines).

The time series of global annual carbon storage and carbon storage capacity for the models in CMIP5, TRENDY, and MsTMIP are shown in Figs. 2d, 2e, and 2f, respectively. Large diversity in the simulated carbon storage is found for the 25 models. The annual carbon storage varies

considerably from less than 600 PgC for CLM4VIC in MsTMIP to about 3200 PgC for MPI-ESM-LR in CMIP5. The large range of carbon storage is highly related to that of the carbon storage capacity. In responses to the external environmental changes, the carbon storage capacity changes quickly, and it drives the carbon storage change. The change rate of annual carbon storage is much slower than that of annual carbon storage capacity, as it is also regulated by the internal carbon cycle processes.

The difference between carbon storage capacity and terrestrial carbon storage is expressed as carbon storage potential in Figs. 2d-f. The interannual patterns of carbon storage in the three MIPs are mainly affected by the carbon storage potential, because the sign and value of the carbon storage potential determine the direction and rate of carbon storage change, respectively. The nine DGVMs in TRENDY present positive carbon storage potential, with larger values over the recent three decades than the first half of the twentieth century, resulting in an increasing direction of carbon storage change toward the carbon storage capacity. However, the models in CMIP5 and MsTMIP exhibit lower variation in annual carbon storage. The carbon storage potential fluctuates between positive and negative values, so the current carbon storage rises and falls frequently within a small range over the study period. Several models in CMIP5 [i.e., CESM1(BGC), GFDL-ESM2G, MIROC-ESM, and NorESM1-ME] and MsTMIP (i.e., CLM4, CLM4VIC, GTEC, ISAM, and VEGAS2.1) show negative carbon storage potential over most years, resulting in a long-term decline in carbon storage (Figs. 2d,f; and supplemental Table S2).

#### *b. Traceability analyses of carbon residence time and NPP*

The carbon residence time and NPP are traced to their baseline values and the environmental scalars. Figure 3 shows the environmental space consisting of air temperature and precipitation for the three MIPs. As the climate and carbon cycle form an intimately coupled system, the environmental space of annual temperature and precipitation is different across the nine ESMs in CMIP5. The simulated precipitation varies from about 680 to 1020 mm yr<sup>-1</sup>, and temperature from 10.7° to 15.5°C among the nine ESMs over the period 1850–2005 (Fig. 3). Generally, the environmental space is not widespread for each model. The ranges of global mean annual temperature and precipitation are less than 3°C and 110 mm yr<sup>-1</sup>, respectively. The environmental space is identical for the models in MsTMIP and TRENDY because the same climate forcing (CRU-NCEP) is used to drive the uncoupled models. The global mean annual precipitation and temperature over land (excluding Antarctica and Greenland) from CRU-NCEP range from 718 to 818 mm yr<sup>-1</sup> and from 12.8° to 14.3°C, respectively, over the period 1901–2010 (Fig. 3). Because the traceability analysis is performed at the global annual scale,

the environmental space shows small diversity and does not reflect the seasonal and spatial variability of temperature and precipitation, such as their large variations in semiarid ecosystems (Poulter et al. 2014).

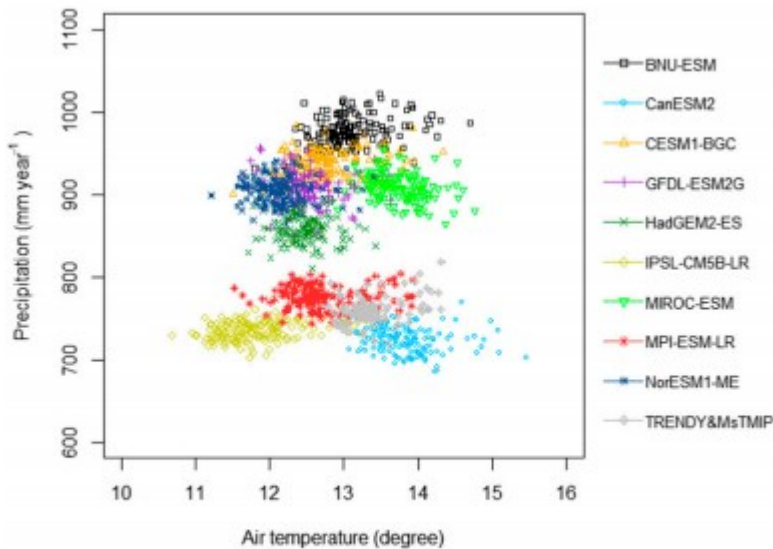


FIG. 3. Distribution of global mean annual air temperature and precipitation over land (excluding Antarctica and Greenland) for the nine ESMs in CMIP5 and the models in TRENDY and MsTMIP. The temperature and precipitation are the same for the models in TRENDY and MsTMIP.

Figure 4 shows the dependence of carbon residence time and NPP on their baseline values and the environmental scalars in the three MIPs. The difference in carbon residence time (or NPP) results from the baseline carbon residence time (or baseline NPP) and the environmental scalar across different models. There is a one- to threefold variation in the baseline carbon residence time and baseline NPP among the models in the three MIPs. The baseline carbon residence time ranges from 21 to 42 years in CMIP5, from 23 to 35 years in TRENDY, and from 12 to 37 years in MsTMIP (Table 2; Figs. 4a-c). And the baseline NPP varies from 49 to 91  $\text{PgC yr}^{-1}$  in CMIP5, from 58 to 82  $\text{PgC yr}^{-1}$  in TRENDY, and from 42 to 85  $\text{PgC yr}^{-1}$  in MsTMIP (Table 2; Figs. 4d-f). However, the distributions of the environmental scalars are much closer across different models, ranging from about 0.7 to 1, both for the carbon residence time and NPP. Thus, the large ranges in carbon residence time and NPP in Figs. 2a-c are mainly attributed to the baseline carbon residence time and baseline NPP among the models in the three MIPs.



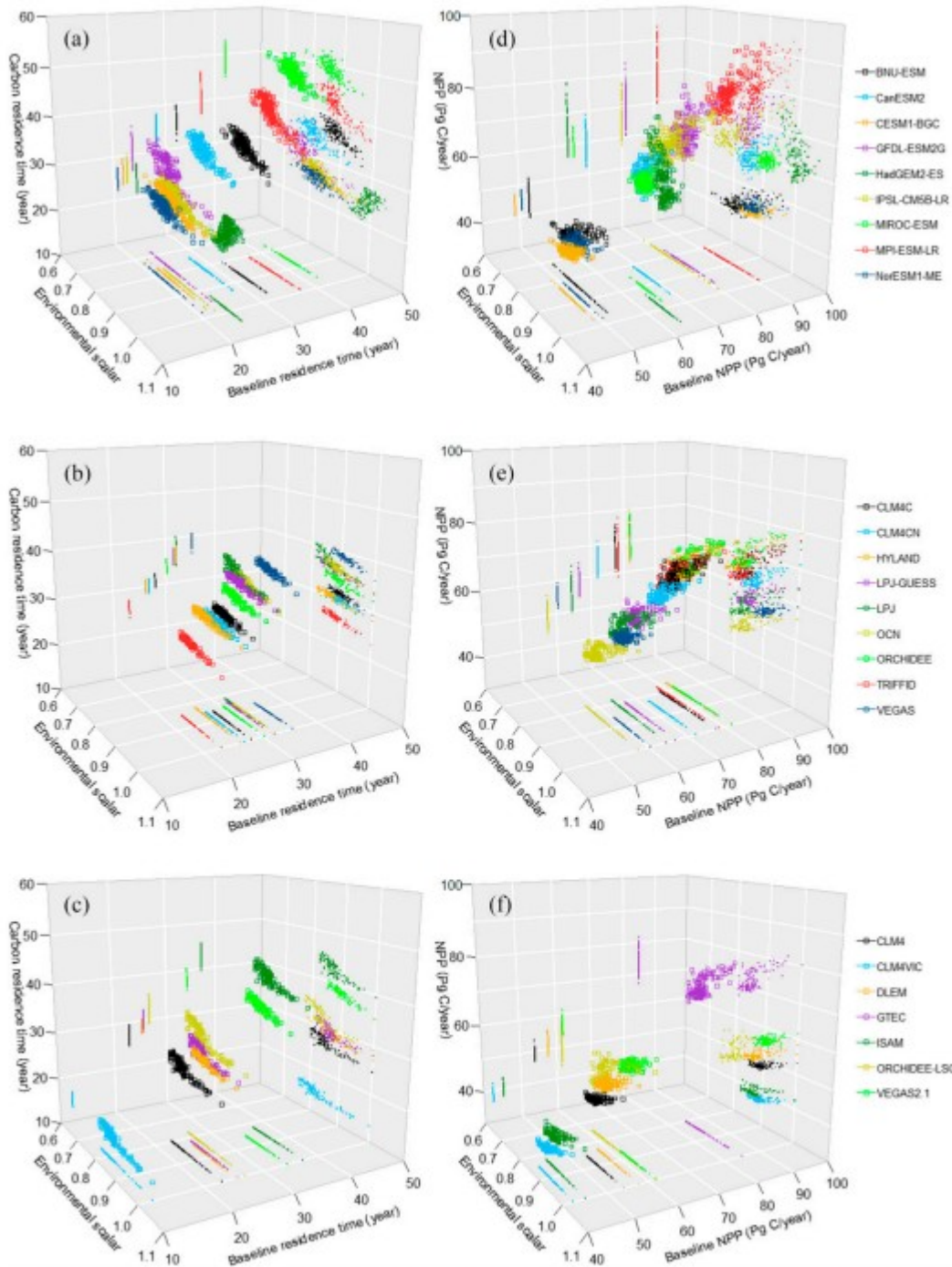


FIG. 4. Decomposition of the carbon residence time into the baseline carbon residence time and the environmental scalar and decomposition of annual NPP into the baseline NPP and the environmental scalar for (a),(d) CMIP5, (b),(e) TRENDY, and (c),(f) MsTMIP. The environmental scalar is a product of the temperature and water scalars, which convert the baseline carbon residence time and baseline NPP into their actual values.

TABLE 2. Calibrated parameters and the performance of the optimization method for the carbon residence time and NPP. The quantities  $T_0$  and  $W_0$  are the reference values of temperature and precipitation.

Models	$T_0$ (°C)	$W_0$ (mm yr <sup>-1</sup> )	The carbon residence time				NPP				
			$\tau'_E$ (yr)	$Q_{10}$	$R^2$	RMSE	NPP (PgC yr <sup>-1</sup> )	$Q_{10}$	$R^2$	RMSE	
CMIP5	BNU-ESM	14.05	1022.75	32.31	1.64	0.53	1.06	52.69	2.51	0.60	1.76
	CanESM2	15.45	769.63	28.34	2.45	0.46	1.21	69.27	1.75	0.31	2.69
	CESM1(BGC)	14.34	984.43	21.84	2.41	0.61	0.85	48.77	1.68	0.61	0.95
	GFDL-ESM2G	13.78	956.57	23.69	4.81	0.35	1.91	81.24	2.39	0.31	3.97
	HadGEM2-ES	13.43	893.66	24.37	0.74	0.34	2.12	64.01	0.67	0.25	5.86
	IPSL-CM5B-LR	13.30	761.16	22.63	2.52	0.62	0.92	80.07	3.33	0.64	2.89
	MIROC-ESM	14.77	956.47	42.38	2.59	0.45	1.26	65.61	1.51	0.23	1.44
	MPI-ESM-LR	13.94	803.91	37.31	2.07	0.62	1.36	91.11	2.22	0.59	3.21
	NorESM1-ME	13.41	940.70	20.83	3.42	0.36	1.01	50.63	2.11	0.35	1.22
	TRENDY	CLM4C	14.31	817.81	28.02	1.36	0.46	0.61	77.66	2.58	0.67
CLM4CN		14.31	817.81	26.83	1.43	0.41	0.64	72.28	1.93	0.75	1.24
HYLAND		14.31	817.81	26.20	1.97	0.58	0.73	81.93	3.49	0.76	2.01
LPJ_GUESS		14.31	817.81	31.70	1.82	0.69	0.73	66.84	2.32	0.62	1.64
LPJ		14.31	817.81	32.17	2.12	0.63	0.91	64.55	2.82	0.69	1.82
OCN		14.31	817.81	32.25	1.57	0.70	0.60	57.53	2.74	0.71	1.44
ORCHIDEE		14.31	817.81	30.31	1.53	0.69	0.56	82.19	2.76	0.69	2.07
TRIFFID		14.31	817.81	22.99	1.61	0.65	0.47	78.43	3.11	0.73	2.11
VEGAS		14.31	817.81	35.37	1.22	0.69	0.53	60.48	1.29	0.63	0.89
MsTMIP		CLM4	14.31	817.81	22.96	2.46	0.50	0.88	54.01	1.49	0.76
	CLM4VIC	14.31	817.81	12.20	3.50	0.49	0.66	42.39	1.54	0.76	0.57
	DLEM	14.31	817.81	25.36	1.72	0.68	0.54	57.74	1.79	0.71	1.00
	GTEC	14.31	817.81	25.69	2.06	0.47	0.92	84.84	2.34	0.63	2.16
	ISAM	14.31	817.81	37.22	1.99	0.61	0.99	45.16	1.91	0.56	1.04
	ORCHIDEE-LSCE	14.31	817.81	26.83	2.99	0.54	1.24	61.57	3.95	0.61	2.38
	VEGAS2.1	14.31	817.81	34.29	1.47	0.74	0.56	61.92	1.26	0.61	0.92

It should be noted that the 3D points in Fig. 4 are scattered. In the traceability analysis, we use the optimization method to decompose the carbon residence time (or NPP) into the baseline carbon residence time (or baseline NPP) and the environmental scalar. As a consequence, the product of them cannot fully explain the variation in the carbon residence time (or NPP) (Table 2 and Fig. S2 in the online supplemental material). The product of the baseline value and the environmental scalar explains  $55 \pm 12\%$  (mean  $\pm 1$  standard deviation) of the variation in the carbon residence time, and  $59 \pm 16\%$  of the variation in NPP, for the three MIPs. The optimization method performs better for the models in TRENDY ( $R^2 = 0.61 \pm 0.10$  for carbon residence time and  $R^2 = 0.69 \pm 0.04$  for NPP) and MsTMIP ( $R^2 = 0.58 \pm 0.10$  for carbon residence time and  $R^2 = 0.66 \pm 0.07$  for NPP). The variations in the carbon residence time and NPP are difficult to capture using the optimization method for several models in CMIP5, such as HadGEM2-ES, probably due to low or even opposite sensitivities of carbon residence time (and NPP) to temperature and precipitation over different regions. In addition, other environmental factors, such as atmospheric CO<sub>2</sub>, land-use change, and nitrogen availability, also influence the interannual variability of carbon residence time and NPP, which calls for an expanded parameterization that incorporates more controlling factors to improve the traceability analysis.

### c. Variance decomposition of the simulated carbon storage

The variation in the carbon storage is decomposed into several components for the three MIPs using the variance decomposition method (Fig. 5). The carbon storage variation is dominated by the carbon residence time and NPP, and the absolute contribution of the carbon storage potential is less than 1%. The baseline carbon residence time and baseline NPP contribute more than 90% to the variation in carbon residence time and NPP, respectively, for each MIP and all three MIPs combined. Specifically, the contribution of the baseline carbon residence time to the carbon storage variation is 45% for CMIP5, 46% for TRENDY, 68% for MsTMIP, and 44% for the three MIPs, and that of the baseline NPP is 50% for CMIP5, 48% for TRENDY, 34% for MsTMIP, and 55% for the three MIPs. However, the temperature and water scalars contribute no more than 5% of variations in the carbon residence time and NPP, respectively. As a consequence, the variation in carbon storage is dominated by the baseline carbon residence time and baseline NPP. These results are consistent with the large ranges of the baseline carbon residence time and baseline NPP and the close distributions of the environmental scalars across different models (Fig. 4).

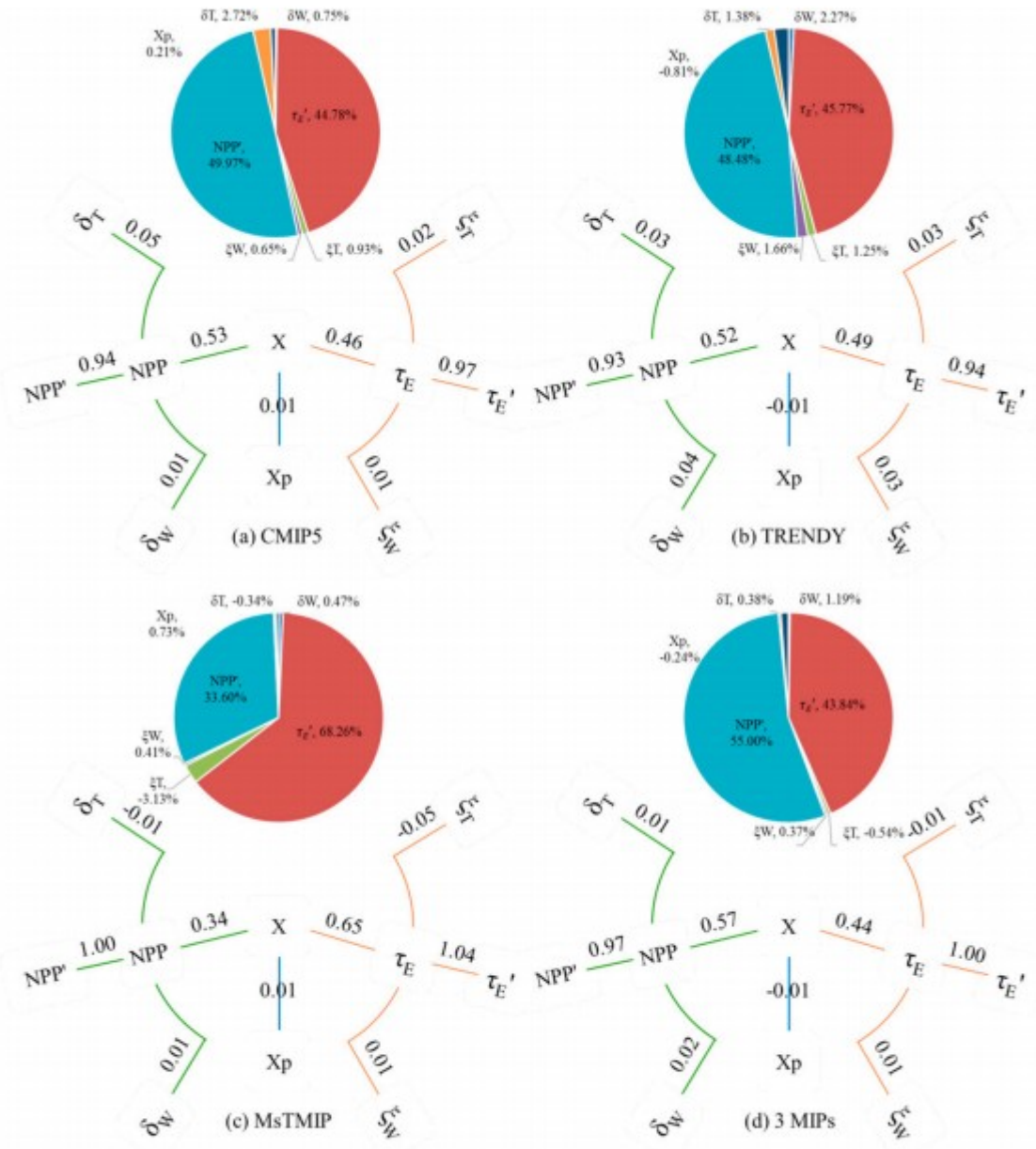


FIG. 5. Variance decomposition of the carbon storage based on global annual data from models in the three MIPs. First, the variation of the carbon storage  $X$  is decomposed into that of the carbon residence time  $\tau_E$ , NPP, and the carbon storage potential  $X_p$ . Second, variations of the carbon residence time and NPP are decomposed into their baseline values ( $\tau_E'$  and NPP') and the temperature ( $\xi_T$  and  $\delta_T$ ) and water ( $\xi_W$  and  $\delta_W$ ) scalars. Positive/negative values mean positive/negative contributions of the variables to the variation of carbon storage.

Although carbon storage variation is mainly attributed to baseline carbon residence and baseline NPP among different models, it is determined by the environmental conditions for individual models when the baseline values are constant. Figure 6 shows the distributions of the air temperature contributions to the variations in carbon residence time and NPP for all three MIPs. Within each model, the total contribution of the temperature and water scalars to the variations in carbon residence time and NPP equals 100%, according to the transient traceability framework. The contributions of

precipitation are shown in Fig. S3 in the online supplemental material. In CMIP5, air temperature explains most of the variations in the carbon residence time ( $74 \pm 20\%$ ) and NPP ( $63 \pm 21\%$ ) for the nine ESMs. The contributions of air temperature in TRENDY ( $40 \pm 13\%$  to carbon residence time and  $63 \pm 15\%$  to NPP) and MsTMIP ( $59 \pm 13\%$  to carbon residence time and  $49 \pm 18\%$  to NPP) are smaller than those in CMIP5. For the 25 models in the three MIPs, the mean contribution of air temperature is more than precipitation, both for the carbon residence time ( $58 \pm 22\%$ ) and NPP ( $59 \pm 19\%$ ).

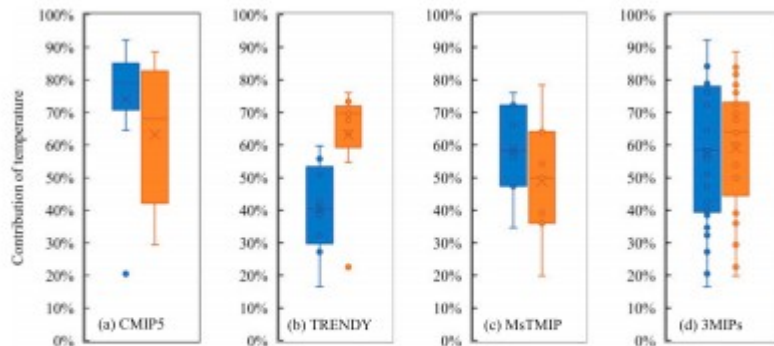


FIG. 6. Distribution of the temperature contributions to the variations in the carbon residence time (blue box) and NPP (orange box) for (a) CMIP5, (b) TRENDY, (c) MsTMIP, and (d) the three MIPs combined.

## 4. Discussion

### *a. Model differences in the baseline carbon residence time and baseline NPP*

All the models in the three MIPs can simulate the processes of photosynthetic carbon input, carbon allocation and transformation, and carbon loss through respiration. Most terrestrial carbon cycle models broadly share a similar structure for carbon cycle simulation (i.e., a pool-and-flux structure) (Luo et al. 2015). In these models, the processes of carbon flow through different pools from its entrance via photosynthesis to its release via respiration are simulated based on a set of carbon balance equations. As indicated by Luo et al. (2015), the internal carbon cycle processes can be characterized by five fundamental properties for all terrestrial ecosystems: compartmentalization, photosynthesis as the dominant carbon input, partitioning among pools, donor pool-dominant transfers, and first-order decay. These five properties have been incorporated into terrestrial carbon cycle models using the carbon balance equations, which can be further summarized as a matrix equation. For a given carbon cycle model, its structure can be represented by the matrix equation, with a given number of carbon balance equations (Luo et al. 2017; Huang et al. 2018). Despite the fact that model structures show a high degree of underlying similarity, across-model variation in carbon cycle parameters results in large differences in NPP and carbon residence time.

Photosynthetic carbon assimilation is the major pathway of carbon flow in terrestrial ecosystems, and it is usually simulated based on the Farquhar model. In the Farquhar model, assimilation rate is jointly controlled by the rubisco limitation of carboxylation and the electron transport rate (Farquhar et al. 1980). Leaf photosynthetic capacity, as determined by rubisco and electron transport capacities, plays an important role in the simulation of ecosystem carbon input. In addition, ecosystem carbon input (e.g., NPP) is also affected by leaf area index, and regulated by environmental factors, such as temperature, radiation, and water availability (Boisvenue and Running 2006; Nemani et al. 2003; Schloss et al. 1999). In this study, NPP variation is traced to the baseline NPP and two environmental scalars. Baseline NPP is related to the maximum carbon input at optimal environmental conditions, and the environmental scalars represent the environmental limitations.

Once assimilated, photosynthetic carbon is allocated into different plant pools (e.g., leaves, stems, and roots) for plant biomass growth and plant respiration. After death, plant organs are transferred to litter pools, which will be decomposed by microorganisms through heterotrophic respiration or transferred to the soil pool in the form of soil organic matter. Soil organic carbon can be stored for hundreds to thousands of years before it is released back into the atmosphere through microbial respiration (Luo and Zhou 2006). The carbon decomposition rate (i.e., the inverse of carbon residence time) in terrestrial ecosystems is greatly affected by environmental conditions, especially temperature and soil moisture (Davidson and Janssens 2006; Sierra et al. 2015). The maximum rate of carbon decomposition at optimal temperature and moisture conditions corresponds to the baseline carbon residence time (i.e., the shortest carbon residence time).

In this study, the processes of carbon input and output simulated in the terrestrial carbon cycle models are summarized as ecosystem NPP and carbon residence time, respectively. The carbon cycle parameters, especially those that determine carbon uptake and release at optimal environmental conditions, are important for simulations of NPP and carbon residence time as they alter their corresponding baseline values. Thus, the differences in the carbon cycle parameters can be measured by the variations of the baseline NPP and baseline carbon residence time. By comparing the output of the Australian Community Atmosphere Biosphere Land Exchange (CABLE) model and Community Land Model, version 3.5 (CLM3.5 or CLM-CASA'), Rafique et al. (2016) indicated that the parameter setting related to NPP and the baseline carbon residence time leads to the eventual model differences. Friend et al. (2014) also identified variations in the carbon residence time as the key difference among models to explain their diverging projections. Our study is consistent in showing that the baseline NPP and baseline carbon residence time contribute to more than 90% of the carbon storage variation, much more than the external environmental scalars or forcing. This result confirms the important role of model

parameters that affect the baseline carbon residence time and baseline NPP in determining the output of terrestrial carbon cycle models. For future model improvement, modelers should pay more attention to the carbon cycle processes and parameters related to the baseline carbon residence time and baseline NPP.

*b. Model intercomparison across the three MIPs*

Different from previous model intercomparison studies (Keenan et al. 2012; Schwalm et al. 2010; Sitch et al. 2008; Zaehle et al. 2014), which aim to compare and improve model performance through model-data analysis, this study attributes the large variations in annual carbon storage to the variations in carbon residence time, NPP, and carbon storage potential to understand intermodel differences.

The 25 models in the three MIPs are different in terms of the three decomposed variables, resulting in great spreads in annual carbon storage. The CMIP5 and MsTMIP present large variations in the simulated carbon storage among different models than TRENDY. According to the variance decomposition method, the widespread baseline carbon residence time and baseline NPP result in large carbon storage variations in CMIP5 and MsTMIP. The large variations in the carbon residence time and NPP among different models may be related to land-use change (Erb et al. 2016), since the simulations in CMIP5 and MsTMIP employed time-variant historical land-use change with different representations of vegetation types, while TRENDY models utilized a constant land-use mask. This is reflected by the nine DGVMs in TRENDY, which consistently show an increasing trend of carbon storage (capacity) over the period 1901–2010, especially the recent three decades. On the contrary, there are no obvious trends in CMIP5 and MsTMIP simulations, where the CO<sub>2</sub> fertilization effect on carbon storage (capacity) may be attenuated by land-use change and the related nitrogen decline in soils and aboveground biomass (Yang et al. 2010; Zhou et al. 2017). The increasing trend of carbon storage in TRENDY is implied by the large and mostly positive carbon storage potential, which determines the direction and rate of carbon storage change (see also Fig. 2b). However, the effects of land-use change on carbon residence time and NPP were not incorporated in the transient traceability analysis in this study. Nevertheless, land-use change and other disturbances influence the carbon cycle by (i) either depleting or adding carbon in pools, (ii) either decreasing or increasing canopy photosynthesis, and (iii) altering carbon residence time via changes in respiration and decomposition (Luo and Weng 2011). All those influences induced by land-use change can be represented by the three dynamics properties (i.e., carbon input, residence time, and the carbon storage potential) and thus analyzed by those techniques developed in this study.

The carbon cycle is coupled with the climate dynamics for the nine ESMs in CMIP5 but not for the models in TRENDY and MsTMIP. The nine ESMs allow for feedback between carbon cycling and climate change, while the climate

forcing is prescribed in TRENDY and MStMIP. The traceability analysis shows that the dependence of the carbon residence time and NPP on the environmental scalars is stronger for the uncoupled models than that for the coupled models (Fig. 4). The relationships between the carbon residence time (or NPP) and the environmental variables are more complex in the ESMs, given the feedbacks of carbon cycle and climate change (Heimann and Reichstein 2008; Sokolov et al. 2008). Thus, variations in the carbon residence time and NPP are not easy to be captured by the temperature and water scalars.

Although the environmental space varies with the simulated climate for the models in CMIP5, and is identical in TRENDY and MStMIP, the total contribution of the environmental scalars to the carbon storage variation in CMIP5 is no more than that in TRENDY and MStMIP. Indeed, both the carbon residence time and NPP are climate dependent. By comparing the “coupled” and “uncoupled” simulations of 11 coupled climate-carbon cycle models, Friedlingstein et al. (2006) show that the impact of climate change on land carbon storage is significant in all models. However, the climate impact on the carbon storage variation across different models is weak, both for the coupled and uncoupled models, because they strongly differ in model parameters of carbon cycle processes. Our results indicate that the difference in the environmental scalars is much smaller than that in the baseline carbon residence time and baseline NPP among the models in all of the three MIPs.

Compared with precipitation, air temperature contributes more to the variations in carbon residence time and NPP for the 25 models. Many studies have shown that carbon cycle processes, including carbon accumulation and decomposition, are sensitive to climate warming (Lu et al. 2013; Xia et al. 2014). However, the change in precipitation varies greatly over different regions, resulting in inconsistent effects on the carbon cycle across the globe, so the precipitation impact on terrestrial carbon sequestration is rather weak at the global scale (Sokolov et al. 2008). The models in CMIP5 exhibit larger mean temperature contributions than those in TRENDY and MStMIP. The large temperature contribution for the nine ESMs may be related to the strong positive feedback between carbon cycle and climate warming (Luo 2007; Zeng 2004), which further enhances the role of air temperature in the carbon cycle processes.

It should be noted that differences in model behaviors are also related to whether a nitrogen cycle is included in the model. Since the productivity of many terrestrial ecosystems is limited by lack of reactive nitrogen (Norby et al. 2010; Zhang et al. 2014), and the CO<sub>2</sub> fertilization effect is strongly down-regulated by nitrogen limitation (Rastetter et al. 1992; Hungate et al. 2003; Walker et al. 2015), NPP is generally lower in the models with nitrogen limitation, resulting in smaller carbon storage capacity and carbon storage, than those that do not include nitrogen limitation. We found the carbon storage potential is mostly negative for the models that include a nitrogen



limitation, which indicates a decreasing trend of carbon storage over the study period, while the other models exhibit positive carbon storage potential over most years. It was reported that nitrogen limitation will reduce the CO<sub>2</sub> fertilization effect, and even cause a reduced NPP for some ecosystems (McMurtrie et al. 2008; Norby et al. 2010; Thornton et al. 2007). The differing behavior between models with and without a nitrogen cycle indicates that the carbon-nitrogen feedback should be considered when assessing model differences.

### *c. Understanding the variation in carbon storage among different models*

In this study, we developed a suite of new techniques for tracing predominant model parameters that govern the simulated global carbon budget in a multimodel setting. We applied these methods to compare the carbon storage dynamics simulated by 25 models in three MIPs. These new techniques include a 3D model output space, a transient traceability framework, and a variance decomposition method, which allow us to elucidate the main source of variability in the historical simulations of carbon storage across different models. In addition, these new techniques can be also applied to future projections to understand large divergence in model predictions (Friedlingstein et al. 2014).

The 3D model output space can measure the difference in the projected carbon storage dynamics in terms of NPP, carbon residence time, and carbon storage potential (Fig. 2; Luo et al. 2017). Despite differences in model structure and carbon cycle parameters, the differences in photosynthetic processes and parameter values can be summarized by that in NPP. Similarly, the differences in processes after photosynthesis and relevant parameters are revealed by the variation in the carbon residence time across different models. The product of NPP and carbon residence time measures the difference in carbon storage capacity at steady state, and the carbon storage potential can capture the transient dynamics of the terrestrial carbon cycle in response to changes in environmental conditions. The 3D model output space can clearly illustrate how and how much the model output differs. Thus, we can perform model evaluations by comparing model output with observations in terms of the three variables, and improve model projections by adjusting the parameters related to NPP and carbon residence time according to their differences with observed values.

The transient traceability framework can decompose a complex carbon cycle model into traceable components by simulating biogeochemical processes. It extended the traceability framework developed by Xia et al. (2013) in two aspects. First, this new framework can directly analyze the transient dynamics of terrestrial carbon storage simulated by the models through involving the third dimension: the carbon storage potential. Second, the modeled NPP is decomposed into the baseline NPP and environmental scalars for temperature and precipitation. Thus, we can attribute the model

differences in NPP and carbon residence time to the variability in model parameters and environmental forcing.

The variance decomposition method can separate the relative contributions of NPP, carbon residence time, and carbon storage potential to variations in terrestrial carbon storage. Thus, the variation in the simulated carbon storage among different models can be quantitatively attributed to the three variables and hence the source factors. In addition, our decomposition method is also applicable to assess the responses of carbon storage to the changing environment (e.g., climate warming, rising atmosphere CO<sub>2</sub>, and other disturbances) and quantify the contributions of the decomposed components to the projected change in carbon storage. This quantitative method can help us investigate the response mechanisms of the terrestrial carbon storage to the environmental changes and therefore better predict terrestrial carbon sequestration response under future climate change.

#### *d. Limitations and recommendations*

In this study, we compared the simulated global mean carbon storage dynamics in the three MIPs based on the 3D model output space, identified the sources of carbon storage variation following the traceability framework, and quantified the relative contributions of the source factors. The three MIPs show a large spread in the simulated carbon storage dynamics, which is effectively revealed by the 3D model output space. Specifically, our study shows that the baseline NPP and baseline carbon residence time are major sources of intermodel variations. Future modeling research needs to better constrain parameters related to these two variables with observations of almost all carbon-related variables, including plant allocation, decomposition, and microbial carbon use efficiency, especially under favorable environmental conditions in order to improve the model projections.

Our study is the first to perform model intercomparison based on the 3D model output space and the transient traceability framework in a multi-MIP setting. This is a post-simulation model evaluation. We took the terrestrial ecosystem as one pool to estimate the carbon residence time and carbon storage potential while the original models have complex structures and variable parameters. Using this simple yet effective method, we are able to obtain the three variables and environmental scalars of the models to develop the 3D model output space, perform traceability analysis, and conduct variance decomposition for understanding model variations. These post-MIP analyses demonstrate that the three techniques developed in this study can be used as an important means to track the origins of model differences from a completed MIP. The main limitation of this study is that the model intercomparison analyses were performed at the global annual scale. The motivation of this is to get an overview on intermodel differences in simulating land carbon storage. More studies should be done in the future to gain understanding of the subannual and grid-scale variability of the three variables, and the difference among biomes as well. Another limitation is

that the traceability analysis of the carbon residence time and NPP was done by only considering the effect of temperature and precipitation. Although the optimization method explained most of the variations in carbon residence time and NPP for the majority of the models, we need more information on regulations of carbon cycle processes by various factors and processes from original models to fully understand variations in model performance. Should we have all carbon balance equations, response functions, and their parameters, the transient traceability analysis can account for almost all variations among models (Luo et al. 2017). For future MIPs, we recommend a matrix approach to reorganize all carbon balance equations in any original model into one matrix equation as for CLM4.5 (Huang et al. 2018) and the Terrestrial Ecosystem Model (TECO) (Jiang et al. 2018). The matrix approach is applicable to almost all land carbon cycle models. Once all the models that are involved in one MIP are converted to matrix equations, we can analyze model uncertainty in a unified diagnostic system. That is, all the land carbon cycle models are represented with one unified formula, model outputs are evaluated in the 3D space, and uncertainty among models can be traced to various components (e.g., carbon input, plant allocation, decomposition rates, and environmental scalars) with the traceability framework (Fig. 1; also see Jiang et al. 2018). We expect that this diagnostic system can greatly improve our understanding of uncertainty sources of land carbon modeling. So, those techniques developed in this study, which are parts of the diagnostic system, can effectively identify the sources of model differences and guide directions for future model improvement.

Future research is needed to perform the analysis at the grid or regional scale. The global analysis cannot fully reveal the origins of differences in model output and may introduce some biases (e.g., due to compensatory effects in time and/or space) in the relationship between the carbon residence time (or NPP) and the environmental scalars. At the regional scale, the decomposed traceable components can be compared with observations to illustrate the deviations of model output from real-world values. In addition, we can determine the key processes or parameters that explain the differences among models as well as between models and observations over different regions. For regional analysis, more environmental factors should be considered in the traceability analysis, such as solar radiation, atmospheric CO<sub>2</sub>, land-use change, and nitrogen availability, to capture the temporal and spatial variability of the carbon residence time and NPP. The carbon storage potential should also be decomposed into its traceable components to further enhance our understanding. As indicated in Eq. (2), the carbon storage potential can be decomposed into the chasing time and the rate of carbon storage change. The chasing time is closely related to the carbon residence time, and they are identical when we use the one-pool model. The rate of carbon storage change is affected by various factors, both internal processes and external forcing. So, it is critical to perform decomposition analysis to identify key processes governing the rate of

carbon storage change. Through analyzing the carbon storage potential, our transient traceability framework can better evaluate the transient terrestrial carbon cycle responses to external forcing and internal processes. Upon careful consideration of the carbon cycling processes, their responses to environmental drivers, and model parameters, the transient traceability framework can elucidate how various processes and parameter settings influence ecosystem carbon storage through the simulated changes in NPP, carbon residence time, and carbon storage potential. Thus, we can efficiently improve model performance toward more realistic projections by adjusting the highlighted carbon cycle processes and parameters in future studies.

## Acknowledgments

We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in supplemental Table S1) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. Funding for the Multi-Scale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP; <http://nacp.ornl.gov/MsTMIP.shtm>) activity was provided through NASA ROSES Grant NNX10AG01A. Data management support for preparing, documenting, and distributing model driver and output data was performed by the Modeling and Synthesis Thematic Data Center at Oak Ridge National Laboratory (ORNL; <http://nacp.ornl.gov>), with funding thorough NASA ROSES Grant NNH10AN681. Finalized MsTMIP data products are archived at the ORNL DAAC (<http://daac.ornl.gov>). We acknowledge the TRENDY-v1 modelers, including Benjamin Poulter from Montana State University, for contributing model output used in this work. This paper is financially supported by the Research and Development Special Fund for Public Welfare Industry of the Ministry of Water Research in China (201501028). JBF and CRS were supported in part by NASA's Carbon Cycle Science program. JBF was also supported in part by NASA's Terrestrial Ecology and Carbon Monitoring System programs. JT acknowledges RCN funded project EVA (229771) and BCCR-BIGCHANGE.

## REFERENCES

Ahlström, A., G. Schurgers, A. Arneeth, and B. Smith, 2012: Robustness and uncertainty in terrestrial ecosystem carbon response to CMIP5 climate change projections. *Environ. Res. Lett.*, 7, 044008, <https://doi.org/10.1088/1748-9326/7/4/044008>.

Anav, A., G. Murray-Tortarolo, P. Friedlingstein, S. Sitch, S. Piao, and Z. Zhu, 2013: Evaluation of land surface models in reproducing satellite derived leaf area index over the high-latitude Northern Hemisphere. Part II: Earth system models. *Remote Sens.*, 5, 3637-3661, <https://doi.org/10.3390/rs5083637>.

Boisvenue, C., and S. W. Running, 2006: Impacts of climate change on natural forest productivity—evidence since the middle of the 20th century. *Global Change Biol.*, 12, 862–882, <https://doi.org/10.1111/j.1365-2486.2006.01134.x>.

Cramer, W., and Coauthors, 1999: Comparing global models of terrestrial net primary productivity (NPP): Overview and key results. *Global Change Biol.*, 5, 1–15, <https://doi.org/10.1046/j.1365-2486.1999.00009.x>.

Davidson, E. A., and I. A. Janssens, 2006: Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature*, 440, 165–173, <https://doi.org/10.1038/nature04514>.

De Kauwe, M. G., and Coauthors, 2014: Where does the carbon go? A model-data intercomparison of vegetation carbon allocation and turnover processes at two temperate forest free-air CO<sub>2</sub> enrichment sites. *New Phytol.*, 203, 883–899, <https://doi.org/10.1111/nph.12847>.

Dhaene, J., A. Tsanakas, E. A. Valdez, and S. Vanduffel, 2012: Optimal capital allocation principles. *J. Risk Insur.*, 79, 1–28, <https://doi.org/10.1111/j.1539-6975.2011.01408.x>.

Drud, A., 1985: CONOPT: A GRG code for large sparse dynamic nonlinear optimization problems. *Math. Program.*, 31, 153–191, <https://doi.org/10.1007/BF02591747>.

Erb, K. H., and Coauthors, 2016: Biomass turnover time in terrestrial ecosystems halved by land use. *Nat. Geosci.*, 9, 674–678, <https://doi.org/10.1038/ngeo2782>.

Exbrayat, J. F., A. J. Pitman, and G. Abramowitz, 2014: Response of microbial decomposition to spin-up explains CMIP5 soil carbon range until 2100. *Geosci. Model Dev.*, 7, 2683–2692, <https://doi.org/10.5194/gmd-7-2683-2014>.

Farquhar, G. D., S. von Caemmerer, and J. Berry, 1980: A biochemical model of photosynthetic CO<sub>2</sub> assimilation in leaves of C<sub>3</sub> species. *Planta*, 149, 78–90, <https://doi.org/10.1007/BF00386231>.

Fisher, J. B., D. N. Huntzinger, C. R. Schwalm, and S. Sitch, 2014: Modeling the terrestrial biosphere. *Annu. Rev. Environ. Resour.*, 39, 91–123, <https://doi.org/10.1146/annurev-environ-012913-093456>.

Friedlingstein, P., and Coauthors, 2006: Climate–carbon cycle feedback analysis: Results from the C<sup>4</sup>MIP model intercomparison. *J. Climate*, 19, 3337–3353, <https://doi.org/10.1175/JCLI3800.1>.

Friedlingstein, P., M. Meinshausen, V. Arora, C. Jones, A. Anav, S. Liddicoat, and R. Knutti, 2014: Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *J. Climate*, 27, 511–526, <https://doi.org/10.1175/JCLI-D-12-00579.1>.

- Friend, A. D., A. K. Stevens, R. G. Knox, and M. G. R. Cannell, 1997: A process-based, terrestrial biosphere model of ecosystem dynamics (Hybrid v3. 0). *Ecol. Modell.*, 95, 249–287, [https://doi.org/10.1016/S0304-3800\(96\)00034-8](https://doi.org/10.1016/S0304-3800(96)00034-8).
- Friend, A. D., and Coauthors, 2014: Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO<sub>2</sub>. *Proc. Natl. Acad. Sci. USA*, 111, 3280–3285, <https://doi.org/10.1073/pnas.1222477110>.
- Harris, I., P. D. Jones, T. J. Osborn, and D. H. Lister, 2014: Updated high-resolution grids of monthly climatic observations: The CRU TS3.10 dataset. *Int. J. Climatol.*, 34, 623–642, <https://doi.org/10.1002/joc.3711>.
- Heimann, M., and M. Reichstein, 2008: Terrestrial ecosystem carbon dynamics and climate feedbacks. *Nature*, 451, 289–292, <https://doi.org/10.1038/nature06591>.
- Huang, Y., and Coauthors, 2018: Matrix approach to land carbon cycle modeling: A case study with Community Land Model. *Global Change Biol.*, 24, 1394–1404, <https://doi.org/10.1111/gcb.13948>.
- Hungate, B. A., J. S. Dukes, M. R. Shaw, Y. Luo, and C. B. Field, 2003: Atmospheric science: Nitrogen and climate change. *Science*, 302, 1512–1513, <https://doi.org/10.1126/science.1091390>.
- Huntzinger, D. N., and Coauthors, 2013: The North American Carbon Program Multi-Scale Synthesis and Terrestrial Model Intercomparison Project—Part 1: Overview and experimental design. *Geosci. Model Dev.*, 6, 2121–2133, <https://doi.org/10.5194/gmd-6-2121-2013>.
- Hurttt, G. C., and Coauthors, 2011: Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Climatic Change*, 109, 117–161, <https://doi.org/10.1007/s10584-011-0153-2>.
- Ichii, K., and Coauthors, 2010: Multi-model analysis of terrestrial carbon cycles in Japan: Limitations and implications of model calibration using eddy flux observations. *Biogeosciences*, 7, 2061–2080, <https://doi.org/10.5194/bg-7-2061-2010>.
- Jiang, L., Z. Shi, J. Xia, J. Liang, X. Lu, Y. Wang, and Y. Luo, 2018: Transient traceability analysis of land carbon storage dynamics: Procedures and its application to two forest ecosystems. *J. Adv. Model. Earth. Syst.*, 9, 2822–2835, <https://doi.org/10.1002/2017MS001004>.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, 77, 437–471, [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2).

Keenan, T. F., and Coauthors, 2012: Terrestrial biosphere model performance for inter-annual variability of land-atmosphere CO<sub>2</sub> exchange. *Global Change Biol.*, 18, 1971–1987, <https://doi.org/10.1111/j.1365-2486.2012.02678.x>.

Knutti, R., and J. Sedláček, 2013: Robustness and uncertainties in the new CMIP5 climate model projections. *Nat. Climate Change*, 3, 369–373, <https://doi.org/10.1038/nclimate1716>.

Lu, M., and Coauthors, 2013: Responses of ecosystem carbon cycle to experimental warming: A meta analysis. *Ecology*, 94, 726–738, <https://doi.org/10.1890/12-0279.1>.

Luo, Y., 2007: Terrestrial carbon-cycle feedback to climate warming. *Annu. Rev. Ecol. Evol. Syst.*, 38, 683–712, <https://doi.org/10.1146/annurev.ecolsys.38.091206.095808>.

Luo, Y., and X. Zhou, 2006: *Soil Respiration and the Environment*. Academic Press, 328 pp.

Luo, Y., and E. Weng, 2011: Dynamic disequilibrium of the terrestrial carbon cycle under global change. *Trends Ecol. Evol.*, 26, 96–104, <https://doi.org/10.1016/j.tree.2010.11.003>.

Luo, Y., and Coauthors, 2003: Sustainability of terrestrial carbon sequestration: A case study in Duke Forest with inversion approach. *Global Biogeochem. Cycles*, 17, 1021, <https://doi.org/10.1029/2002GB001923>.

Luo, Y., T. F. Keenan, and M. Smith, 2015: Predictability of the terrestrial carbon cycle. *Global Change Biol.*, 21, 1737–1751, <https://doi.org/10.1111/gcb.12766>.

Luo, Y., and Coauthors, 2017: Transient dynamics of terrestrial carbon storage: Mathematical foundation and its applications. *Biogeosciences*, 14, 145–161, <https://doi.org/10.5194/bg-14-145-2017>.

McMurtrie, R. E., R. J. Norby, B. E. Medlyn, R. C. Dewar, D. A. Pepper, P. B. Reich, and C. V. Barton, 2008: Why is plant-growth response to elevated CO<sub>2</sub> amplified when water is limiting, but reduced when nitrogen is limiting? A growth-optimisation hypothesis. *Funct. Plant Biol.*, 35, 521–534, <https://doi.org/10.1071/FP08128>.

Meehl, G. A., C. Covey, B. McAvaney, M. Latif, and R. J. Stouffer, 2005: Overview of the Coupled Model Intercomparison Project. *Bull. Amer. Meteor. Soc.*, 86, 89–93, <https://doi.org/10.1175/BAMS-86-1-89>.

Nemani, R.

R., C. Keeling, H. Hashimoto, W. Jolly, S. Piper, C. Tucker, R. Myneni, and S. Running, 2003: Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science*, 300, 1560–1563, <https://doi.org/10.1126/science.1082750>.

- Nishina, K., and Coauthors, 2014: Quantifying uncertainties in soil carbon responses to changes in global mean temperature and precipitation. *Earth Syst. Dyn.*, 5, 197–209, <https://doi.org/10.5194/esd-5-197-2014>.
- Nishina, K., and Coauthors, 2015: Decomposing uncertainties in the future terrestrial carbon budget associated with emission scenarios, climate projections, and ecosystem simulations using the ISI-MIP results. *Earth Syst. Dyn.*, 6, 435–445, <https://doi.org/10.5194/esd-6-435-2015>.
- Norby, R. J., J. M. Warren, C. M. Iversen, B. E. Medlyn, and R. E. McMurtrie, 2010: CO<sub>2</sub> enhancement of forest productivity constrained by limited nitrogen availability. *Proc. Natl. Acad. Sci. USA*, 107, 19 368–19 373, <https://doi.org/10.1073/pnas.1006463107>.
- Oleson, K., and Coauthors, 2013: Technical description of version 4.5 of the Community Land Model (CLM). NCAR Tech. Note NCAR/TN-503+STR, 420 pp., <https://doi.org/10.5065/D6RR1W7M>.
- Parton, W. J., 1996: The CENTURY model. *Evaluation of Soil Organic Matter Models*, D. Powlson, P. Smith, and J. Smith, Eds., Springer, 283–291.
- Potter, C. S., J. T. Randerson, C. B. Field, P. A. Matson, P. M. Vitousek, H. A. Mooney, and S. A. Klooster, 1993: Terrestrial ecosystem production: A process model based on global satellite and surface data. *Global Biogeochem. Cycles*, 7, 811–841, <https://doi.org/10.1029/93GB02725>.
- Poulter, B., and Coauthors, 2014: Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle. *Nature*, 509, 600–603, <https://doi.org/doi:10.1038/nature13376>.
- Rafique, R., J. Xia, O. Hararuk, G. R. Asrar, G. Leng, Y. Wang, and Y. Luo, 2016: Divergent predictions of carbon storage between two global land models: Attribution of the causes through traceability analysis. *Earth Syst. Dyn.*, 7, 649–658, <https://doi.org/10.5194/esd-7-649-2016>.
- Rastetter, E., R. McKane, G. Shaver, and J. Melillo, 1992: Changes in C storage by terrestrial ecosystems: How C-N interactions restrict responses to CO<sub>2</sub> and temperature. *Water Air Soil Pollut.*, 64, 327–344, <https://doi.org/10.1007/BF00477109>.
- Running, S. W., and E. R. Hunt, 1993: Generalization of a forest ecosystem process model for other biomes, BIOME-BCG, and an application for global-scale models. *Scaling Physiological Processes: Leaf to Globe*, J. R. Ehleringer and C. B. Field, Eds., Academic Press, 141–158.
- Schloss, A. L., and Coauthors, 1999: Comparing global models of terrestrial net primary productivity (NPP): Comparison of NPP to climate and the normalized difference vegetation index (NDVI). *Global Change Biol.*, 5, 25–34, <https://doi.org/10.1046/j.1365-2486.1999.00004.x>.



- Schwalm, C. R., and Coauthors, 2010: A model-data intercomparison of CO<sub>2</sub> exchange across North America: Results from the North American Carbon Program site synthesis. *J. Geophys. Res.*, 115, G00H05, <https://doi.org/10.1029/2009JG001229>.
- Sierra, C. A., S. E. Trumbore, E. A. Davidson, S. Vicca, and I. Janssens, 2015: Sensitivity of decomposition rates of soil organic matter with respect to simultaneous changes in temperature and moisture. *J. Adv. Model. Earth Syst.*, 7, 335–356, <https://doi.org/10.1002/2014MS000358>.
- Sitch, S., and Coauthors, 2008: Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs). *Global Change Biol.*, 14, 2015–2039, <https://doi.org/10.1111/j.1365-2486.2008.01626.x>.
- Sitch, S., and Coauthors, 2015: Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences*, 12, 653–679, <https://doi.org/10.5194/bg-12-653-2015>.
- Sokolov, A. P., D. W. Kicklighter, J. M. Melillo, B. S. Felzer, C. A. Schlosser, and T. W. Cronin, 2008: Consequences of considering carbon-nitrogen interactions on the feedbacks between climate and the terrestrial carbon cycle. *J. Climate*, 21, 3776–3796, <https://doi.org/10.1175/2008JCLI2038.1>.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, 93, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>.
- Thornton, P. E., J.-F. Lamarque, N. A. Rosenbloom, and N. M. Mahowald, 2007: Influence of carbon-nitrogen cycle coupling on land model response to CO<sub>2</sub> fertilization and climate variability. *Global Biogeochem. Cycles*, 21, GB4018, <https://doi.org/10.1029/2006GB002868>.
- Todd-Brown, K. E. O., J. T. Randerson, W. M. Post, F. M. Hoffman, C. Tarnocai, E. A. G. Schuur, and S. D. Allison, 2013: Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations. *Biogeosciences*, 10, 1717–1736, <https://doi.org/10.5194/bg-10-1717-2013>.
- Walker, A. P., and Coauthors, 2015: Predicting long-term carbon sequestration in response to CO<sub>2</sub> enrichment: How and why do current ecosystem models differ? *Global Biogeochem. Cycles*, 29, 476–495, <https://doi.org/10.1002/2014GB004995>.
- Wei, Y., and Coauthors, 2014: The North American Carbon Program Multi-Scale Synthesis and Terrestrial Model Intercomparison Project—Part 2: Environmental driver data. *Geosci. Model Dev.*, 7, 2875–2893, <https://doi.org/10.5194/gmd-7-2875-2014>.
- Xia, J., Y. Luo, Y. P. Wang, and O. Hararuk, 2013: Traceable components of terrestrial carbon storage capacity in biogeochemical models. *Global Change Biol.*, 19, 2104–2116, <https://doi.org/10.1111/gcb.12172>.

- Xia, J., J. Q. Chen, S. L. Piao, P. Ciais, Y. Q. Luo, and S. Q. Wan, 2014: Terrestrial carbon cycle affected by non-uniform climate warming. *Nat. Geosci.*, 7, 173–180, <https://doi.org/10.1038/ngeo2093>.
- Yang, X., T. Richardson, and A. Jain, 2010: Contributions of secondary forest and nitrogen dynamics to terrestrial carbon uptake. *Biogeosciences*, 7, 3041–3050, <https://doi.org/10.5194/bg-7-3041-2010>.
- Zaehle, S., and Coauthors, 2014: Evaluation of 11 terrestrial carbon–nitrogen cycle models against observations from two temperate free-air CO<sub>2</sub> enrichment studies. *New Phytol.*, 202, 803–822, <https://doi.org/10.1111/nph.12697>.
- Zeng, N., 2004: How strong is carbon cycle–climate feedback under global warming? *Geophys. Res. Lett.*, 31, L20203, <https://doi.org/10.1029/2004GL020904>.
- Zhang, Q., Y. P. Wang, R. J. Matear, A. J. Pitman, and Y. J. Dai, 2014: Nitrogen and phosphorous limitations significantly reduce future allowable CO<sub>2</sub> emissions. *Geophys. Res. Lett.*, 41, 632–637, <https://doi.org/10.1002/2013GL058352>.
- Zhang, Y., and Coauthors, 2016: Consistency between sun-induced chlorophyll fluorescence and gross primary production of vegetation in North America. *Remote Sens. Environ.*, 183, 154–169, <https://doi.org/10.1016/j.rse.2016.05.015>.
- Zhou, S., and Coauthors, 2017: Response of water use efficiency to global environmental change based on output from terrestrial biosphere models. *Global Biogeochem. Cycles*, 31, 1639–1655, <https://doi.org/10.1002/2017GB005733>.