

Article

Modeling the Carbon Cycle of a Subtropical Chinese Fir Plantation Using a Multi-Source Data Fusion Approach

Longwei Hu ¹, Honglin He ^{2,3,4,*}, Yan Shen ¹, Xiaoli Ren ^{2,3}, Shao-kui Yan ⁵, Wenhua Xiang ^{1,6} , Rong Ge ^{2,3,7} , Zhongen Niu ^{2,3,7}, Qian Xu ^{2,3,7} and Xiaobo Zhu ^{2,3,8,9} 

¹ Central South University of Forestry & Technology, Changsha 410004, China; hulw1002@163.com (L.H.); 13873340123@163.com (Y.S.); xiangwh2005@163.com (W.X.)

² Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; renxl@igsrr.ac.cn (X.R.); ge7218@163.com (R.G.); niuze.16b@igsrr.ac.cn (Z.N.); 15221858038@163.com (Q.X.); wavelet@email.swu.edu.cn (X.Z.)

³ National Ecosystem Science Data Center, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

⁴ College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China

⁵ Huitong Experimental Station of Forest Ecology, State Key Laboratory of Forest and Soil Ecology, Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110164, China; yanshaokui@126.com

⁶ National Engineering Laboratory for Applied Technology of Forestry & Ecology in South China, Changsha 410004, China

⁷ University of Chinese Academy of Sciences, Beijing 100049, China

⁸ Chongqing Jinpo Mountain Field Scientific Observation and Research Station for Karst Ecosystem, Ministry of Education, Chongqing 400715, China

⁹ Chongqing Engineering Research Center for Remote Sensing Big Data Application, School of Geographical Sciences, Southwest University, Chongqing 400715, China

* Correspondence: hehl@igsrr.ac.cn; Tel.: +86-10-6488-9599

Received: 22 February 2020; Accepted: 21 March 2020; Published: 26 March 2020



Abstract: Process-based terrestrial ecosystem models are increasingly being used to predict carbon (C) cycling in forest ecosystems. Given the complexity of ecosystems, these models inevitably have certain deficiencies, and thus the model parameters and simulations can be highly uncertain. Through long-term direct observation of ecosystems, numerous different types of data have accumulated, providing valuable opportunities to determine which sources of data can most effectively reduce the uncertainty of simulation results, and thereby improve simulation accuracy. In this study, based on a long-term series of observations (biometric and flux data) of a subtropical Chinese fir plantation ecosystem, we use a model–data fusion framework to evaluate the effects of different constrained data on the parameter estimation and uncertainty of related variables, and systematically evaluate the uncertainty of parameters. We found that plant C pool observational data contributed to significant reductions in the uncertainty of parameter estimates and simulation, as these data provide information on C pool size. However, none of the data effectively constrained the foliage C pool, indicating that this pool should be a target for future observational activities. The assimilation of soil organic C observations was found to be important for reducing the uncertainty or bias in soil C pools. The key findings of this study are that the assimilation of multiple time scales and types of data stream are critical for model constraint and that the most accurate simulation results are obtained when all available biometric and flux data are used as constraints. Accordingly, our results highlight the importance of using multi-source data when seeking to constrain process-based terrestrial ecosystem models.

Keywords: process-based terrestrial ecosystem model; model–data fusion; multi-source data; carbon cycle; plantation

1. Introduction

Forest ecosystems are among the most important of terrestrial ecosystems, in that they store large amounts of carbon (C) and play vital roles in regulating the global C balance and mitigating climate change. In the context of climate change, it is essential to accurately estimate C cycling in forest ecosystems, as this can provide a basis for predicting climate change and controlling the greenhouse effect. In recent years, a large number of different types of observations related to changes in ecosystem C cycles have accumulated, which have mainly been derived from environmental control experiments and eddy covariance measurements [1–5]. However, deciding how most effectively to utilize the accumulated data to enhance our understanding of terrestrial ecosystems remains a challenge [6].

Process-based terrestrial ecosystem models are important tools for studying key processes of C cycles and the mechanisms underlying their control and have been widely used to estimate regional and/or global C budgets [7–11]. However, given the complexity of ecosystems, our current understanding of ecosystem-related key processes and control mechanisms is insufficiently comprehensive, as model parameters inevitably have associated uncertainties, and thus these models are still unable to accurately simulate and predict ecosystem processes and C source/sink distribution and changes [12–15]. In this regard, the model–data fusion technique (MDF) provides a powerful tool for reducing the uncertainty when simulating ecosystem C cycles by combining observations and models, which can contribute to improving model simulation accuracy [16–18].

MDF makes full use of existing observations and applies mathematical methods to optimize the parameters and/or state variables of the model to achieve the best combination between simulations and measurements, thereby enabling more accurate simulation of the changes in ecosystem state [19–21]. The accuracy of these simulations, in turn, depends on the appropriate acquisition of model parameters [2,12,22–24] and, given that models contain a large number of parameters that are difficult to estimate accurately, most of the research on MDF in the field of C cycles have tended to focus on parameter estimation [9,12,15,25,26].

At present, various time scales and types of observational data are commonly used in terrestrial process-based terrestrial ecosystem modeling, which may contain different types of information on ecosystem processes. For example, C flux data contains considerable amounts of information regarding how “fast” processes respond to environmental drivers (e.g., photosynthesis and respiration), whereas biometric data contains information relating to numerous “slow” processes (e.g., C pool size and turnover) [1,15,27,28]. Therefore, the data used as constraints require multiple data streams. In the MDF, due to differences in the circumstances of data collection, biometric data estimation parameters [26] and flux data [29–34] have been used to independently constrain parameters, as well as being used in combination [9,25,35]. However, although there have been numerous studies in this field, few have specifically quantified the impact of different data streams on simulation effects.

Plantations grow rapidly, with higher carbon sequestration rates and greater potential [36,37]. Accordingly, accurate simulation of the C cycle of planted forests is of considerable significance for reasonable evaluations of the C sequestration rates and potentials of forest ecosystems. In this study, we used a Data Assimilation Linked Ecosystem Carbon (DALEC) [38] model in conjunction with the Markov Chain-Monte Carlo (MCMC) method, in an inverse analysis using long-term data obtained from the Chinese Ecosystem Research Network (CERN, <http://www.cern.ac.cn/>) and China FLUX (<http://www.chinaflux.org/>) to model the C cycle of a subtropical coniferous plantation in Huitong, China. We constrained the model parameters using a range of different data streams, including field measurements of leaf area index (LAI), foliage, wood, and fine root biomass, annual litterfall, soil organic matter, and eddy covariance measurements of net ecosystem exchange (NEE), as well as

manual chamber measurements of soil respiration (Rs). Our objective is to use the MDF framework to determine how combinations of different constraint data influence parameter estimates and how these different parameter-sets influence simulation. On the basis of this information, we can identify which observational data are most likely to reduce model uncertainties.

2. Materials and Methods

2.1. Site Description

The Huitong subtropical Chinese fir plantation (HTF) observation station (26°47' N, 109°35' E) is located in Huitong County, Hunan Province, China. It lies in a transition zone that extends from the Yun-Guizhou Plateau to the Jiangnan Hills, which is characterized by low mountains and hills with elevations of between approximately 300 and 1000 m. The parent rock is dominated by slab and shale, and the soil is a red or yellow soil. It is a typical subtropical region, with an average annual temperature of 16.5 °C and annual precipitation of 1200–1400 mm. The Chinese fir plantation, which comprises pure stands of Chinese fir, was planted in 1983. The landform is hilly with a slope of 20°, which runs in a southeast direction [39]. The observation station of HTF has a comprehensive observation field (15,000 m²) of Chinese fir plantation, and long-term observation of water, soil, meteorology, and biology. There are permanent plots, and destructive plots are set in the comprehensive observation field. The area of permanent plots are about 5000 m², and a first-class plot (40 × 50 m) is set inside, which is divided into 25 secondary plots (8 × 10 m), five of them are surveyed every five years, with a cycle of 25 years planned.

2.2. Data

The collected data included a meteorological observation dataset (drive data) and eight sets of observational data (constraint data), including one flux dataset (NEE, Rs), three biomass datasets (foliage, fine root, and wood), and single datasets for litter, soil organic C, and canopy dynamics (LAI).

2.2.1. Flux Data

The NEE data were obtained from ChinaFLUX and were aggregated to the daily time step from 30 min CO₂ flux data measured using the eddy covariance technique and processed via quality control and gap-filling [40]. The flux observation system consists of an ultrasonic anemometer (CSAT3, Campbell, Camden, NJ, USA) and a CO₂/H₂O infrared gas analyzer (Li7500, Li-cor, Lincoln, NE, USA). The height of the observation tower is 32.5 m, and the installation height of the flux observation system is 32.5 m; flux observation time was May 2007. A negative NEE value indicates that the ecosystem absorbs CO₂ from the atmosphere, whereas a positive value indicates that the ecosystem releases CO₂ to the atmosphere. Rs values were measured using static chamber-gas chromatography techniques [41]. A total of four to six replicate measurements were obtained two to three times per month, with sampling times between 09:00 and 11:00. The Rs values thus obtained were converted to monthly averages to reduce the effects of wind, rain, and other weather.

2.2.2. Biometric Data

LAI was measured optically using an LAI-2000 plant canopy analyzer (LI-COR, Lincoln, NE, USA) at least quarterly every year and corrected using a foliage clumping index, which was set for plant functional type (PFT)-specific empirical values [42]. In the MDF analysis, the seasonal variation in LAI combined with the leaf C mass per leaf area (LCMA) parameter constrained the dynamic trajectory of the foliage C.

Data obtained for biomass observations made during the study period (2005–2015) were multiplied by 50% for conversion to C values. The biomass of leaves, stems, and roots was calculated using an allometric method according to measured diameters at breast height and tree height. The ratio of the biomass of fine roots and the biomass of entire roots in typical Chinese forests was obtained from

Zhang and Wu (2001) [43] to facilitate the division of biomass into fine and coarse root components, and the biomass of coarse roots was combined with the biomass of branches and stems to provide estimates of woody biomass. The data for leaf, fine root, and woody biomass were used to constrain the corresponding C pools in the inverse analysis.

The aboveground litterfall biomass (foliage and fine twigs) was measured by 10 replicates of baskets (100 × 100 cm). Litter was collected monthly during the growing season and once during the non-growing season. All collected litter was dried at 70 °C for 24 h and weighed. We used the annual litter biomass data (2005, 2006, 2007, 2010, 2011) for inversion analysis to avoid the influence of wind on the measurement of litterfall biomass within an individual month. The results of long-term annual measurements indicate that the proportion of C content to dry matter in litter is essentially maintained at approximately 50%, and thus the average dry weight of litter can be reasonably accurately converted to litter C.

2.2.3. Soil Data

Soil organic matter (SOM) and soil bulk density were measured using the potassium dichromate oxidation titrimetric method and the cutting ring method, respectively. Soil sampling was performed in five layers (at depths of 0–10, 10–20, 20–40, 40–60, and 60–100 cm) at 5-year intervals (2005 and 2010), with three replicates per sampling. We calculated the soil organic C (SOC) using Equation (1) [44]:

$$SOC = \sum_{i=1}^n 0.58 \times H_i \times B_i \times O_i \times 100, \quad (1)$$

where SOC is the soil organic C density (g C/m²) of all n layers, H_i is soil thickness (cm), B_i is soil bulk density (g/cm³), and O_i is the SOM content of the i -th layer (%).

2.2.4. Meteorological Driving Data

Data for meteorological drivers were obtained from the CERN Science and Technology Resource Service System, including daily meteorological data from 2005 to 2015 (air temperature (T, °C), photosynthetically active radiation (PAR, mol m⁻²/day), relative humidity (RH, %), and saturated vapor pressure difference (VPD, hPa)). The data were processed by standardized quality control and gap-filling [40,45].

2.3. DALEC Model

The Data Assimilation Linked Ecosystem Carbon (DALEC) model is a simple ecosystem process model developed by Williams and colleagues [38], which integrates assimilated data and has been widely used in MDF analyses [5,9,27,38,46]. The DALEC model connects the C pools by flux. In the present study, we used the evergreen version of DALEC [27] (Figure 1), which represents the C pools of foliage (C_f), wood (C_w), fine roots (C_r), litter (C_{lit}), and soil (C_{soil}).

In general, the C cycle of forest ecosystems can be characterized in detail based on four properties, which are outlined as follows. (1) CO₂ in the atmosphere enters the ecosystem via photosynthesis, and the C cycle typically commences with canopy C [8]. In this regard, we estimated gross primary productivity (GPP) in the present study using a canopy photosynthesis model, which is a function of LAI, T, PAR, VPD, and RH [47]. (2) A proportion of the GPP is consumed as autotrophic respiration (R_a), whereas the remainder (i.e., net primary production, NPP) is allocated to the different plant C pools (i.e., foliage, wood, and fine root pools). The C from the plant pool subsequently enters the dead organic matter pool with temperature-dependent loss (heterotrophic respiration, R_h). (3) Subsequent C transfer is dominated by the upstream C pools (i.e., the litter is decomposed (D) and is incorporated into the soil). (4) The C released from C reservoirs is based on a first-order differential equation.

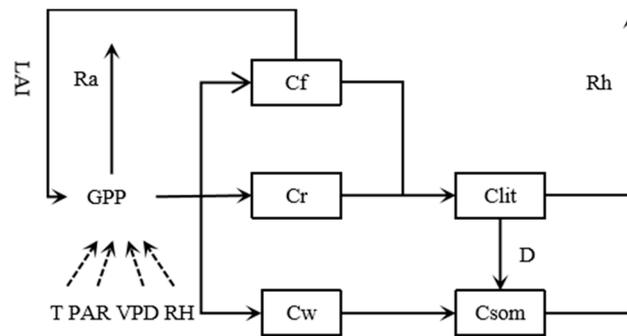


Figure 1. Schematic diagram of the structure of the Data Assimilation Linked Ecosystem Carbon (DALEC) model. Air temperature (T), photosynthetically active radiation (PAR), saturated vapor pressure difference (VPD), relative humidity (RH), leaf area index (LAI), gross primary productivity (GPP), decomposition (D), autotrophic respiration (Ra), heterotrophic respiration (Rh). C pools: foliage (Cf), wood (Cw), fine roots (Cr), litter (Clit), soil (Csom).

2.4. Parameterization of DALEC

In order to parameterize DALEC, we used the Markov Chain–Monte Carlo (MCMC) method, which is a non-sequential global search method based on Bayesian theory [2]. The MCMC algorithm treats the solved parameters as random variables that meet a certain prior probability distribution. According to the Bayesian formula, the MCMC sampling sequence is constructed through a large number of random samples [21]. The prior distribution of the parameters is used to determine their posterior distribution, and then infer the unknown parameters based on the posterior distribution. The Bayes formula is as follows (Equation (2)):

$$\rho(\theta|x) = \frac{\rho(x|\theta)\rho(\theta)}{\rho(x)}, \quad (2)$$

where $\rho(\theta|x)$ and $\rho(\theta)$ are the posterior and prior probability density distributions of the parameters, $\rho(x)$ is the probability of the observation data, and $\rho(x|\theta)$ is the prior parameter of the observations under the value, that is, the likelihood function of θ . In this study, we assumed that the random error follows a normal distribution, and its likelihood function is represented as follows:

$$L = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \eta_i)^2}{2\sigma^2}}, \quad (3)$$

where n is the number of data points in the observations, x_i and η_i are the i -th observations and simulations, respectively, and σ is the standard error of the observation.

We designed five optimization experiments (Runs 1 to 5; Table 1) to examine the effects of different data streams on parameters and simulation results. Using LAI observations as the base dataset (Run 1), we added other biometric data and C flux data in a step-by-step manner. All experiments assimilated the LAI time series, which can be used to constrain leaf C mass per unit area (LCMA) and indirectly constrains foliage carbon. Run 2 assimilated the biometric data (foliage, fine root, and wood biomass) reflecting plant C. We designed two experiments to constrain the associated C pool for the soil C simulation results, which were highly uncertain, with Run 3 assimilating the soil C information reflecting the dynamics of soil C storage and Run 4 assimilating the observational data for litter [48]. Run 5 assimilated all the observational data used, including C flux data (NEE, Rs) reflecting the carbon exchange process of the ecosystem. The root mean square error (RMSE) and mean absolute error (MAE) for simulated and observed values were used to quantify the error.

Table 1. Multi-source data optimization experiments based on DALEC (Data Assimilation Linked Ecosystem Carbon).

Run ID	Constraining Data Streams Included in Parameter Optimization
1	LAI only
2	Run 1 + foliage + fine root + wood
3	Run 2 + SOC
4	Run 3 + litter
5	Run 4 + NEE + Rs

Notes: soil organic C, SOC; net ecosystem exchange, NEE; soil respiration, Rs.

3. Results

3.1. Constraint Effect of Multi-Source Data on Model Parameters

3.1.1. Parameter Posterior Distribution

In this study, we used biometric data and C flux data to constrain the parameters of the DALEC model. The MCMC method was used to estimate the posterior distribution of the 10 parameters of the DALEC model (Figure 2). According to the shape of posterior distributions, the optimized parameters could be classified into three groups: well-constrained, poorly constrained, and edge hitting, and thus we were able to determine the constraint effect of different data streams on model parameters.

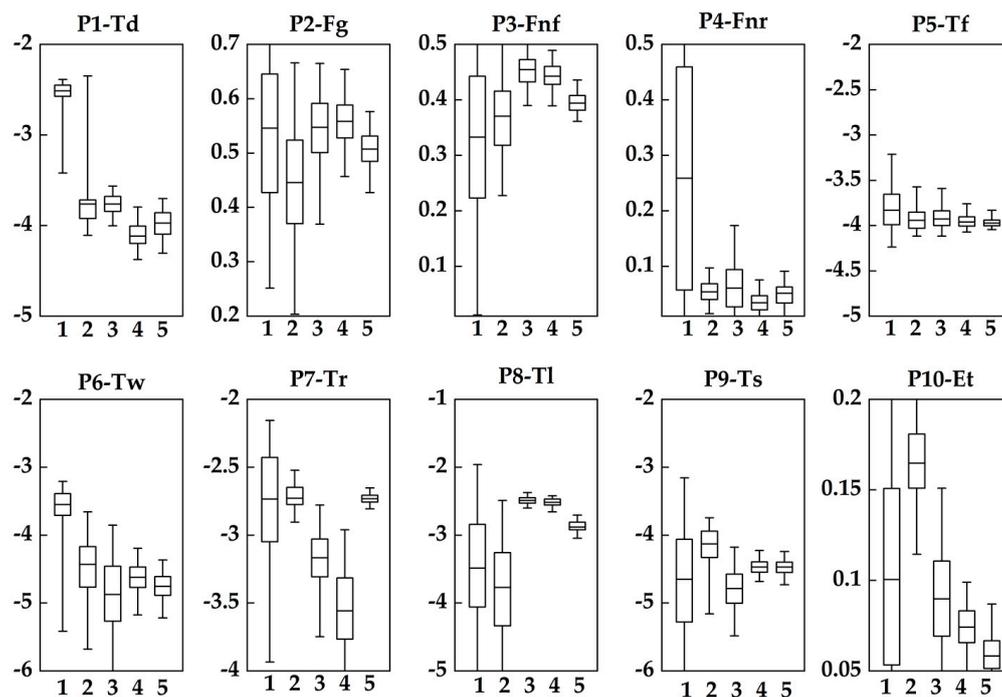


Figure 2. Posterior distributions of parameters (listed in Table 2) estimated for the DALEC model using a variety of different data constraints (Runs 1–5, x-axis, see Table 1). All y-axis represent priori range. Whiskers indicate 95% confidence interval; box indicate interquartile range).

Table 2. DALEC model parameters.

Parameters	Abbreviation	Description	Prior Range (Low, High)
P1	Td	Decomposition rate (per day)	$\text{Log}_{10}(-6, -2)$
P2	Fg	Fraction of GPP respired	(0.2, 0.7)
P3	Fnf	Fraction of NPP allocated to foliage	(0.01, 0.5)
P4	Fnr	Fraction of NPP allocated to roots	(0.01, 0.5)
P5	Tf	Turnover rate of foliage (per day)	$\text{Log}_{10}(-5, -2)$
P6	Tw	Turnover rate of wood (per day)	$\text{Log}_{10}(-6, -2)$
P7	Tr	Turnover rate of roots (per day)	$\text{Log}_{10}(-4, -2)$
P8	Tl	Mineralization rate of litter (per day)	$\text{Log}_{10}(-4, -1)$
P9	Ts	Mineralization rate of SOM (per day)	$\text{Log}_{10}(-6, -2)$
P10	Et	Parameter in exponential term of temperature-dependent rate parameter	(0.05, 0.2)

The parameters were considered to be well constrained if their posterior distribution was normally distributed (or approximated to a normal distribution) [49]. The posterior distribution of the poorly constrained parameters was uniformly distributed, with the exception of the uniform distribution of Fg (fraction of GPP respired) in Run 1, this type of distribution was rarely observed in other experiments. The posterior distribution of the edge-hitting parameters was often concentrated near the extreme value of the parameter value range. For example, Td (decomposition rate) and Tw (rate of wood turnover) in Run 1. Although edge hitting was rarely detected in the posterior distribution with increases in constraint data, Tf (rate of foliage turnover) was the only parameter that was observed in all five experiments. Edge hitting may be partly ascribed to the insignificant constraining effect of the data in model parameters and the limit of the threshold of model parameters. Our analysis indicated that only two parameters (Fg and Tw) could be constrained using only LAI data (Run 1). Six parameters—Fg, Fnf (fraction of NPP allocated to foliage), Fnr (fraction of NPP allocated to roots), Tw, Tr (rate of root wood turnover), Et (parameter in exponential term of temperature-dependent rate parameter)—could be well constrained by plant C stock data (Run 2). These parameters are mainly related to the C pools of foliage, fine roots, and wood. Moreover, when the data for the soil C pool were included, Tl and Ts (representing the mineralization rates of litter and SOM, respectively) conformed to a normal distribution. With the addition of litter data (Run 4), Fg, Fnf, Tw, Tr, and Ts were all normally distributed. When all data streams were used in parameter optimization (Run 5), with the exception of Tf and Et, most parameters could be well constrained. In general, increasing the amounts of constraint data was found to be conducive to increasing the number of parameters that can be constrained. In particular, for certain parameters (e.g., Fnf and Td), a multiple constraint approach may be required in order to achieve well-constrained posterior distribution.

3.1.2. Parameter Uncertainties

The degree to which the posterior distribution of the parameters was constrained and the improvement in the prior value of the uniform distribution (Figure 2) depend on both the data used to constrain the calibration (Table 1) and the parameters involved (Table 2). We found that the parameter posterior distributions generally included the upper and lower limits of the prior value range. However, the parameters Td, Tf, Tw, Tr, Tl, and Ts were noteworthy in all instances, for which the posterior interquartile ranges were significantly reduced (with average reductions of 81%, 87%, 74%, 67%, 75%, and 76%, respectively) compared with the parameter prior interquartile range (= half the width of the prior range, given a uniform prior).

Whether the parameters could be constrained by the data was dependent on the information contained in the data and the expression of this information in the model. When only LAI was used to constrain parameters (Run 1), the 95% confidence interval of the posterior distribution of six of the 10 parameters was outside the range of the prior interquartile, whereas for four of the 10 parameters, the posterior interquartile range of the posterior distribution was not substantially

reduced compared with the prior value interquartile (a less than 20% reduction). Although increasing the number of constraint data sets used could generally reduce the uncertainty of the parameters, which parameters were affected was highly dependent on the data being assimilated. For example, soil respiration data (Run 5) were required to constrain the estimates of F_g , fine root data could strongly constrain F_{nr} , and litter (Run 4) data were needed to effectively prevent the posterior distribution of T_f from widening. Unsurprising, when all data streams were used (Run 5), we observed the tightest parameter distribution.

3.2. The Influence of Multi-Source Data on Predictions

3.2.1. C Pools

When all the observed data were used to constrain state variables (Run 5), the goodness-of-fit statistics (Table 3) and simulated C pool of the DALEC model (Figure 3) indicated that the model was capable of capturing the observational data. When the simulations and observations of C pools (foliage (C_f), wood (C_w), fine root (C_r), litter (C_{lit}), and soil (C_{som})) were compared with Run 1, we observed that the RMSE decreased by 42%, 94%, 75%, 97%, and 93%, respectively, and that the MAE decreased by 60%, 94%, 75%, 96%, and 94%, respectively. In contrast, the RMSE and MAE of the simulated LAI error increased slightly (by less than 4%).

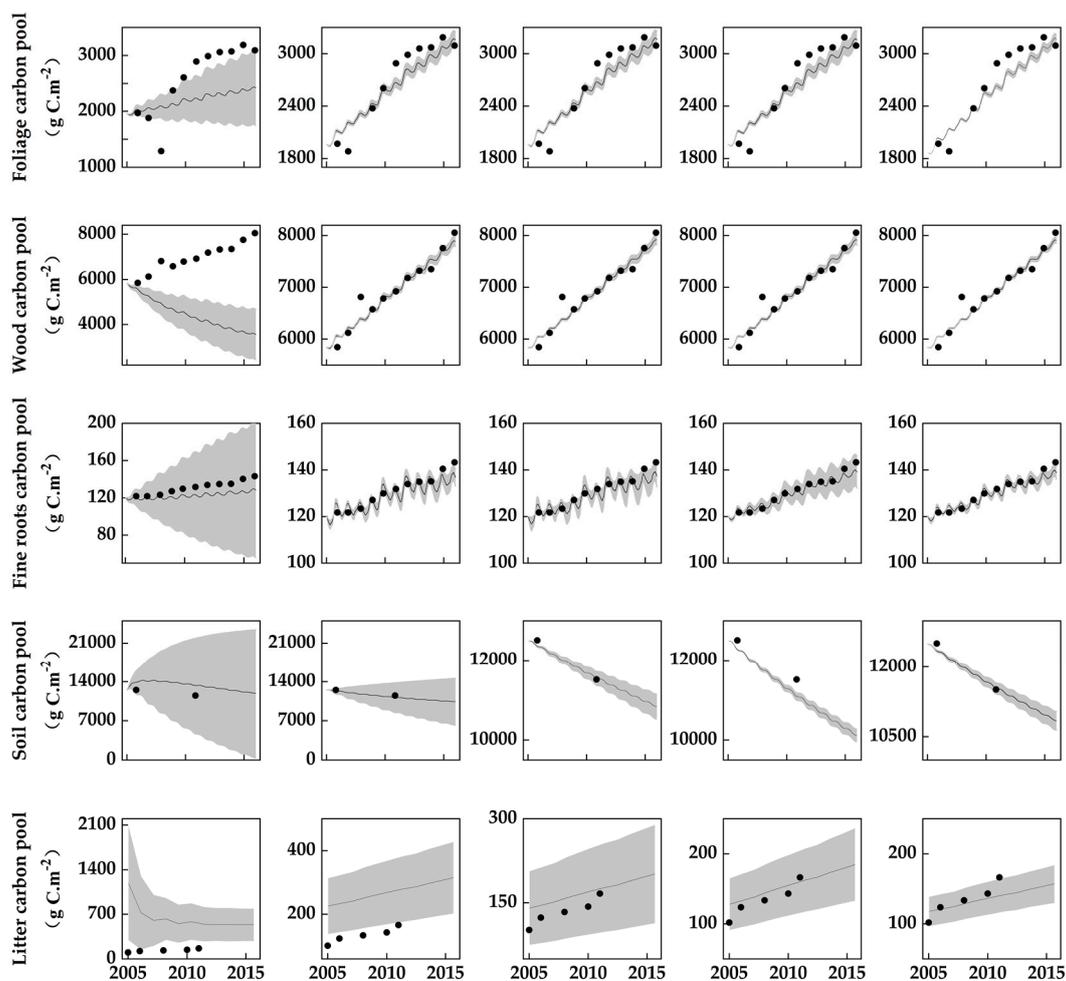


Figure 3. Comparison between modeled and observed values of carbon pools (Run 1 to 5 from left to right (Table 1). The solid line in the figure represents the average value of modeled results, the gray area represents the 95% confidence interval, and the blank circle represents the observed value).

Table 3. Fit between the modeled and observed values for six variables at the Huitong subtropical Chinese fir plantation.

		Run 1	Run 2	Run 3	Run 4	Run 5
Foliage carbon Pool (g C·m ⁻²)	R ²	0.79	0.78	0.78	0.78	0.78
	RMSE	594.33	347.74	348.07	347.72	346.82
	MAE	536.59	217.58	216.20	218.57	215.71
Wood carbon Pool (g C·m ⁻²)	R ²	0.93	0.93	0.93	0.93	0.93
	RMSE	2847.25	167.00	163.93	165.11	164.21
	MAE	2564.48	119.28	109.58	114.85	111.73
Fine roots carbon Pool (g C·m ⁻²)	R ²	0.94	0.88	0.91	0.96	0.95
	RMSE	8.35	2.91	2.68	1.89	2.10
	MAE	7.22	2.37	2.21	1.5953	1.79
Soil carbon Pool (g C·m ⁻²)	R ²	-	-	-	-	-
	RMSE	1569.98	326.74	108.99	343.15	113.20
	MAE	1554.83	320.31	90.08	329.92	90.68
Litter carbon Pool (g C·m ⁻²)	R ²	0.77	0.90	0.91	0.91	0.90
	RMSE	651.56	119.52	26.87	15.00	13.52
	MAE	600.34	119.33	25.31	13.03	10.50
LAI (m m ⁻²)	R ²	0.64	0.66	0.66	0.66	0.66
	RMSE	0.03	0.06	0.07	0.08	0.12
	MAE	0.03	0.05	0.06	0.07	0.12

Simulating foliage C pools represents a more challenging exercise. Although all the experiments captured the interannual variability of foliage C, the magnitude of this pool was significantly underestimated (Figure 3, Table 3). Assimilation of LAI time series data (Run 1), which reflect canopy dynamics, failed to adequately constrain foliage C. When biometric data (Run 2) reflecting the dynamics of vegetation C stocks were assimilated, RMSE and MAE were significantly reduced by 41% and 59%, respectively, compared with Run 1. Although the 95% confidence interval was reduced by 86% and the uncertainty was reduced, the RMSE and MAE remained at 347.74 and 217.55 g C m⁻², respectively. With the addition of soil C and litterfall information, the simulations of Cf were still markedly different from the observations. The smallest confidence interval for Cf was obtained in Run 5; however, the simulation results were still highly uncertain (RMSE > 300 g C m⁻², MAE > 200 g C m⁻²). These observations thus indicated that the fast-turnover C pool was not sufficiently constrained, which may indicate the need for additional types of data and/or larger amounts of data to constrain this C pool. Alternatively, the insufficient constraint could be attributable to uncertainties in the observations.

We found that the assimilation of information on plant C pools (Run 2) strongly constrained the wood and fine roots C pools (Figure 3, Table 3). When only LAI was assimilated (Run 1), the simulation of Cw was significantly underestimated, with RMSE and MAE values of 2847.26 and 2564.48 g C m⁻², respectively. Subsequent to the assimilation of wood information (Run 2), compared with Run 1, the RMSE decreased by 17-fold to 167.01 g C m⁻² (Table 3), the MAE decreased by 21-fold to 119.28 g C m⁻², and the 95% confidence interval decreased by 91%. However, the incorporation of additional observation data (Run 3 to 5) had little effect on the simulation results of Cw (R² > 0.93), with the increases in RMSE and MAE being less than 3 and 10 g C m⁻², respectively (Table 3). Similar to Cw, we detected the largest confidence interval when information relating to Cr was not assimilated (i.e., Run 1). With the exception of Run 1, Cr (R² > 0.88) was well simulated in Run 2 to 5, and compared with Run 1, the RMSE and MAE decreased by an average of 71% and 72%, respectively, and the confidence interval decrease of more than 90%.

Soil C is a key component of process-based terrestrial ecosystem models and one of the largest sources of uncertainty, and in this regard, soil C time series data represent key data for C research, which can minimize the uncertainty of simulation results [48]. We obtained the smallest average RMSE and MAE values (108.99 and 90.08 g C m⁻², respectively) when assimilating soil C data. Compared with the experiments in which soil C information was not assimilated (Run 1 and 2), RMSE and MAE

decreased by an average of 80% and 83%, respectively, and the average confidence interval decreased by an average of 95% (Figure 3, Table 3). However, after assimilating information for litterfall, the RMSE and MAE values increased by 215% and 266%, although the confidence interval decreased by 48%. With the assimilation of the C flux, the bias decreased, with RMSE and MAE values reaching 113.20 and 90.68 g C m⁻², respectively.

The litter C pool is notably sensitive to the assimilation biometric data (Figure 3, Table 3), and we found that when assimilating only LAI, the simulated values overestimated Clit, with RMSE and MAE values of 651.56 and 600.34 g C m⁻², respectively. With the assimilation of information on foliage, fine roots, and wood, the bias between the simulated and observed value was appreciably reduced ($R^2 = 0.90$), and RMSE and MAE decreased by 82% and 80%, respectively. With the addition of soil C data, the bias showed further decreases. After assimilating litterfall data, the RMSE and MAE further decreased to 15.00 and 13.05 g C m⁻², respectively. Compared with the experiments without assimilated litter (Run 1 to 3), RMSE and MAE decreased by an average of 76% and 78%, respectively, and the confidence interval decreased by an average of 58%.

When all observations with the exception of C_{som} were used to constrain the DALEC model, the confidence interval of the other four carbon pools were the narrowest (C_{som} was the narrowest at Run 4), with smaller bias, indicating that multi-source data is effective in constraining the DALEC model. It should be pointed out, however, that MDF entails a balance between variables [50]. For example, the simulation results of Run 3 and 4 revealed that better simulation results for Clit were gained at the expense of increasing the bias of other state variables (such as C_{som}).

3.2.2. C Fluxes

Compared with the observational data ($R^2 = 0.003$ and 0.001 , respectively), assimilation of LAI and plant C pool information (Run 1 and 2) yielded mean RMSE and MAE values of 2.81 and 2.20 g C m⁻², respectively, thereby indicating that these C sources do not reflect the true state of Chinese fir plantations (Figure 4, Table 4) and that assimilation of LAI and plant C data alone cannot constrain NEE. When the soil C data were assimilated (Run 3), the simulation results were C sink ($R^2 = 0.094$), and the average values of RMSE and MAE were 1.17 and 1.47 g C m⁻², respectively. After the addition of litterfall data (Run 4), RMSE and MAE increased slightly ($R^2 = 0.098$). Furthermore, with the assimilation of C flux (Run 5), the simulated total annual NEE in 2008 was -140.92 g C m⁻² (-0.39 ± 1.22 g C m⁻²), the RMSE and MAE were 1.44 and 1.17 g C m⁻² ($R^2 = 0.099$), respectively, and the simulation corresponded well with the observation (-168.90 g C m⁻²) (Table 4, Figure 4).

Table 4. Statistics of C flux simulation results for the Huitong subtropical Chinese fir plantation.

Run ID	NEE (g C·m ⁻² day ⁻¹)				Rs (g C·m ⁻² m ⁻¹)			
	RMSE	MAE	Mean	Standard deviation	RMSE	MAE	Mean	Standard deviation
1	2.63	2.15	1.48	1.61	5.41	4.21	5.41	3.53
2	2.99	2.24	1.42	1.90	7.13	5.21	6.38	4.67
3	1.41	1.14	-0.40	1.18	1.19	0.99	2.20	1.03
4	1.52	1.21	0.20	1.15	0.91	0.79	2.50	1.20
5	1.44	1.17	-0.39	1.22	0.66	0.55	1.78	0.80

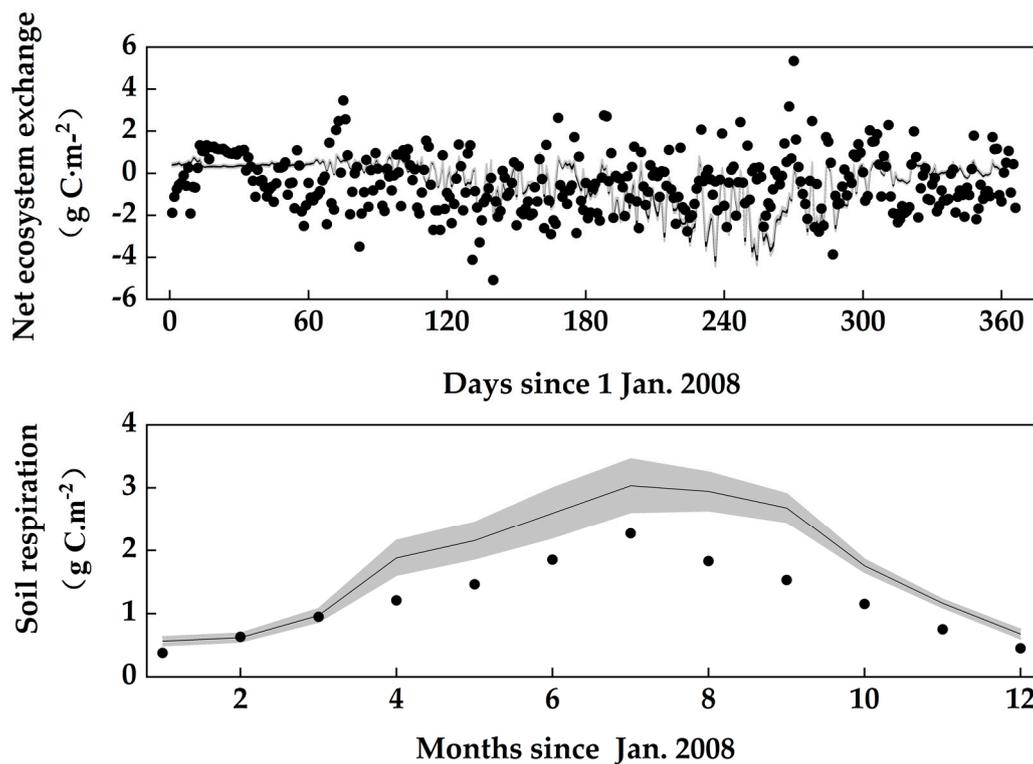


Figure 4. Comparison between modeled and observed values of net ecosystem carbon flux (NEE) and soil respiration (Rs) in 2008, constrained with a variety of different data (Run 5 in Table 1) (the solid line in the figure represents the average value of modeled results, the gray area represents the 95% confidence interval, and the blank circle represents the observed value).

Soil respiration was also found to be sensitive to information assimilation (Figure 4, Table 4), with the bias of the simulation results generally becoming smaller, with an increase in assimilated constraint data. The assimilation of LAI had virtually no constraining effect on Rs ($R^2 = 0.97$), whereas, in response to the addition of information on foliage, fine roots, and wood (Run 2), the bias of the simulation results increased ($R^2 = 0.96$), whereas in contrast, the addition of soil C (Run 3), resulted in a significant decrease in bias ($R^2 = 0.85$), and compared with Run 2, RMSE and MAE decreased by 83% and 81%, respectively. When information on litterfall was added (Run 4), the bias was further decreased ($R^2 = 0.91$). Not surprisingly, the bias reached a minimum, with RMSE and MAE values of 0.66 and 0.55 g C m^{-2} , respectively, when C flux data were assimilated (Run 5), although the simulated value for soil respiration ($1.75 \pm 0.80 \text{ g C m}^{-2}$) was slightly different from the observed value ($1.21 \pm 0.60 \text{ g C m}^{-2}$) ($R^2 = 0.93$), which can be attributed mainly to an overestimation of summer soil respiration (Figure 4, Table 4).

4. Discussion

In this study, we used multi-source data to constrain the parameters of a DALEC model for a subtropical coniferous plantation in Huitong, China. After assimilating all collected observations, we observed that the posterior interquartile range of parameters Td, Tf, Tw, Tr, Tl, and Ts were reduced by at least 60%, and the bias of simulations was markedly reduced, which enhanced the accuracy of the model. When constraint is based on a single source of data, the effect on the model is generally limited [6,9,12,25,51,52]. For example, when using only flux data or biometric data, only a fraction of the parameters in the model are typically well constrained. However, with the incorporation of additional observations, the estimates of many (although not all) model parameters are significantly tightened, and the uncertainty associated with the simulation results is generally substantially reduced (Tables 3 and 4; Figures 3 and 4) [9,25]. When all the relevant data were used to constrain the DALEC

model, we found that only two of the 10 assessed parameters (Tr and Et) were poorly constrained (Figure 2), thereby indicating that other types of data constraint may be required, although this does not necessarily imply that larger quantities of constraint data will have a better constraining effect on the model [6,20]. In this regard, Keenan et al. (2013) [6] found that in terms of constraining model performance, much of the data information is redundant, as these authors observed that only five of the 17 data sources they evaluated were required for model constraint.

Whether the model parameters can be well constrained by the data is dependent on the data information content and how this information is represented in the model [1,20]. With respect to forest C cycling, for example, the data per se contains only limited information on some aspects of the C flux, such as NEE and Rs, and does not contain sufficient information to constrain parameters related to C pools, whereas C pool data includes the allocation and transfer between different C pools, which is relevant to internal system dynamic information. These data are extremely effective for constraining parameters and variables, and thus multiple data streams are needed to constrain the model. Plant C pool observations are effective in constraining the relevant parameters of leaf, fine root, and wood C pools, whereas C flux (NEE, Rs) has a strong constraining effect on NPP allocation and related parameters of soil C pools [53]. This indicates that observational data for foliage, fine root, and woody biomass provide information on C transfer from plant C to litter and soil C, and C flux data provide information on NPP allocation and carbon transfer between the litter and soil C pools.

Simulating foliage C remains a challenge necessitating further attention. Even with the assimilation of litterfall data (Run 4), the simulation of foliage C pools was not further constrained. This may be because fir canopies contain multiple years of foliage (i.e., foliage longevity is two or more years). In addition, the dynamics of foliar C pool has a limit point in the evergreen version of DALEC model [54], which depends on the values of Fnf and Tf, there may also be a need for further constraints on these two parameters. In the present study, we found that even when multiple types of observational data were assimilated, the simulation results for foliage C pools continued to show a large bias. Therefore, we speculate that other types of data are necessary for foliage C simulation (for example, soil moisture [55] and chlorophyll content [56] that are closely related to plant growth), suggesting that these should be targets for future measurement efforts. In addition, the simulated annual summed values of NEE were in good agreement with observed value (Run 5). However, it appears that the DALEC model significantly overpredicts Rs, thus should underpredict NEE for the year. To better simulate C flux, more amounts and smaller scales of data may be required. It is worth noting that the RMSE, MAE, and confidence intervals mainly reflect the impact of the differences in parameter values on the simulation results (uncertainty), and we did not consider the influence of the driving data, observational data, or model structure on the simulation results. Data errors (including system errors) are particularly important factors with regards to the estimation of parameters and state variables, and in future research, it will be necessary to take into account data errors and model structure, and the quantitative errors of the system.

The uncertainty associated with the simulation results for C pools increases with time, whereas the uncertainty for C flux is generally relatively small (Figures 3 and 4) [57,58]. This is because the size of C pools is a cumulative quantity, and simulation errors typically show a gradual increase with an increase in time, whereas the C flux represents a change within a given time period, and thus the error of the simulation results is relatively independent. In addition, in the present study, we applied a parameter optimization method for the batch method, that is, we assimilated all our observational data to optimize the parameters and state variables of the model, whereas the sequential method [such as the Kalman filter method [59,60] can make parameters and state variables with the appearance of the new observation update (calibration), if the flux data, simulation is carried out in real-time. This may be very important (for example, in numerical weather prediction), as it may effectively reduce the uncertainty of the simulation results.

The uncertainty of model parameters and model prediction results generally decrease with the addition of new data to the inversion analysis, which enables us to gain a more comprehensive

understanding of the information contained in these data [25,61]. C flux data also provides valuable constraining information, and since the measurement of these fluxes is often performed at many sites, these data are increasingly being used in MDF analysis. Our analysis indicates that incorporating multiple constraints in the inversion analysis can contribute to enhancing model simulation results by reducing the prediction bias and uncertainty during the forward run of the model.

5. Conclusions

Biometric data obtained from site observations and NEE data derived using the eddy covariance method provide rich sources of information for the parameters and state variables of constraint models. There are still challenges in simulating the foliar C pool, which needs to be paid attention to in future research. In this study, we found that the degrees to which model parameters are constrained and model performance is enhanced are dependent on the different constraint data employed. With an increase in assimilated data, we observed a gradual decrease in the uncertainty of simulations, and when all evaluated observations were used for model constraint, parameter uncertainties were minimized and the simulated values were in good agreement with the observed values. Plant carbon data were found to be the most effective in terms of significantly reducing the uncertainty of the DALEC model carbon pool, and when all observation data are used to constrain the model, the number of well-constrained model parameters increases, thereby enhancing the accuracy of model simulations. Furthermore, we observed that incorporating the additional constraints of soil organic carbon observations is a key factor with respect to reducing the uncertainty or bias associated with the soil carbon pool.

Author Contributions: Conceptualization, H.H. and L.H.; Data curation, L.H. and R.G.; Formal analysis, H.H. and L.H.; Funding acquisition, H.H.; Methodology, X.R., R.G. and Q.X.; Resources, Y.S., S.-k.Y., W.X., Z.N., and X.Z.; Validation, L.H. and R.G.; Writing—original draft, L.H.; Writing—review and editing, H.H. and L.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was jointly supported by the Strategic Priority Research Program of Chinese Academy of Science (grant number: XDA19020301), the National Natural Science Foundation of China (grant number: 41571424).

Acknowledgments: We would like to thank the staff of ChinaFLUX and CERN for their dedication to observation and data processing.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References

1. Niu, S.; Luo, Y.; Dietze, M.C.; Keenan, T.F.; Shi, Z.; Li, J.; Iii, F.S.C. The role of data assimilation in predictive ecology. *Ecosphere* **2014**, *5*, 1–16. [[CrossRef](#)]
2. Peng, C.; Guiot, J.; Wu, H.; Jiang, H.; Luo, Y. Integrating models with data in ecology and palaeoecology: Advances towards a model-data fusion approach. *Ecol. Lett.* **2011**, *14*, 522–536. [[CrossRef](#)] [[PubMed](#)]
3. López-Blanco, E.; Exbrayat, J.-F.; Lund, M.; Christensen, T.R.; Tamstorf, M.P.; Slevin, D.; Hugelius, G.; Bloom, A.A.; Williams, M. Evaluation of terrestrial pan-Arctic carbon cycling using a data-assimilation system. *Earth Syst. Dyn.* **2018**, *10*, 233–255. [[CrossRef](#)]
4. Exbrayat, J.-F.; Bloom, A.A.; Falloon, P.; Ito, A.; Smallman, T.L.; Williams, M. Reliability ensemble averaging of 21st century projections of terrestrial net primary productivity reduces global and regional uncertainties. *Earth Syst. Dyn.* **2018**, *9*, 153–165. [[CrossRef](#)]
5. Bloom, A.A.; Exbrayat, J.-F.; van der Velde, I.R.; Feng, L.; Williams, M. The decadal state of the terrestrial carbon cycle: Global retrievals of terrestrial carbon allocation, pools, and residence times. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 1285–1290. [[CrossRef](#)]
6. Keenan, T.F.; Davidson, E.A.; Munger, J.W.; Richardson, A.D. Rate my data: Quantifying the value of ecological data for the development of models of the terrestrial carbon cycle. *Ecol. Appl.* **2013**, *23*, 273–286. [[CrossRef](#)]
7. Li, W.; Ciais, P.; Wang, Y.; Peng, S.; Broquet, G.; Ballantyne, A.P.; Canadell, J.G.; Cooper, L.; Friedlingstein, P.; Quere, C.L.; et al. Reducing uncertainties in decadal variability of the global carbon budget with multiple datasets. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 13104–13108. [[CrossRef](#)]

8. Xia, J.; Luo, Y.; Wang, Y.; Hararuk, O. Traceable components of terrestrial carbon storage capacity in biogeochemical models. *Glob. Chang. Biol.* **2013**, *19*, 2104–2116. [[CrossRef](#)]
9. Smallman, T.L.; Exbrayat, J.F.; Mencuccini, M.; Bloom, A.A.; Williams, M. Assimilation of repeated woody biomass observations constrains decadal ecosystem carbon cycle uncertainty in aggrading forests. *J. Geophys. Res. Biogeosci.* **2017**, *122*, 528–545. [[CrossRef](#)]
10. Keenan, T.F.; Davidson, E.; Moffat, A.M.; Munger, W.; Richardson, A.D. Using model-data fusion to interpret past trends, and quantify uncertainties in future projections, of terrestrial ecosystem carbon cycling. *Glob. Chang. Biol.* **2012**, *18*, 2555–2569. [[CrossRef](#)]
11. Wu, D.; Piao, S.; Liu, Y. Evaluation of CMIP5 Earth System Models for the Spatial Patterns of Biomass and Soil Carbon Turnover Times and Their Linkage with Climate. *J. Clim.* **2018**, *31*, 5947–5960. [[CrossRef](#)]
12. Ren, X.; He, H.; Zhang, L.; Li, F.; Liu, M.; Yu, G.; Zhang, J. Modeling and uncertainty analysis of carbon and water fluxes in a broad-leaved Korean pine mixed forest based on model-data fusion. *Ecol. Model.* **2018**, *379*, 39–53. [[CrossRef](#)]
13. Wang, Y.P.; Trudinger, C.M.; Enting, I.G. A Review of Applications of Model-Data Fusion to Studies of Terrestrial Carbon Fluxes at Different Scales. *Agric. For. Meteorol.* **2009**, *149*, 1829–1842. [[CrossRef](#)]
14. Luo, Y.; Weng, E.; Wu, X.; Gao, C.; Zhou, X.; Zhang, L. Parameter identifiability, constraint, and equifinality in data assimilation with ecosystem models. *Ecol. Appl.* **2009**, *19*, 571–574. [[CrossRef](#)]
15. Liu, M.; He, H.; Ren, X.; Sun, X.; Yu, G.; Han, S.; Han, S.; Wang, H.; Zhou, G. The effects of constraining variables on parameter optimization in carbon and water flux modeling over different forest ecosystems. *Ecol. Model.* **2015**, *303*, 30–41. [[CrossRef](#)]
16. Scholz, K.; Hammerle, A.; Hiltbrunner, E.; Wohlfahrt, G. Analyzing the effects of growing season length on the net ecosystem production of an alpine grassland using model–data fusion. *Ecosystems* **2018**, *21*, 982–999. [[CrossRef](#)]
17. Zhang, T.; Sun, R.; Peng, C.; Peng, C.; Zhou, G.; Wang, C.; Zhu, Q.; Yang, Y. Integrating a model with remote sensing observations by a data assimilation approach to improve the model simulation accuracy of carbon flux and evapotranspiration at two flux sites. *Sci. China Earth Sci.* **2016**, *59*, 337–348. [[CrossRef](#)]
18. Zhou, T.; Shi, P.; Jia, G.; Luo, Y. Nonsteady state carbon sequestration in forest ecosystems of china estimated by data assimilation. *J. Geophys. Res. Biogeosci.* **2013**, *118*, 1369–1384. [[CrossRef](#)]
19. MacBean, N.; Peylin, P.; Chevallier, F.; Scholze, M.; Schurmann, G. Consistent assimilation of multiple data streams in a carbon cycle data assimilation system. *Geosci. Model. Dev.* **2016**, *9*, 3569–3588. [[CrossRef](#)]
20. Ricciuto, D.M.; King, A.W.; Dragoni, D.; Post, W.M. Parameter and prediction uncertainty in an optimized terrestrial carbon cycle model: Effects of constraining variables and data record length. *J. Geophys. Res.* **2011**, *116*, 1–17. [[CrossRef](#)]
21. Zobitz, J.M.; Desai, A.R.; Moore, D.J.P.; Chadwick, M.A. A primer for data assimilation with ecological models using Markov Chain Monte Carlo (MCMC). *Oecologia* **2011**, *167*, 599–611. [[CrossRef](#)] [[PubMed](#)]
22. Lin, J.C.; Pejam, M.R.; Chan, E.; Wofsy, S.C.; Gottlieb, E.W.; Margolis, H.A.; McCaughey, J.H. Attributing uncertainties in simulated biospheric carbon fluxes to different error sources. *Glob. Biogeochem. Cycles* **2011**, *25*, 1–17. [[CrossRef](#)]
23. Keenan, T.F.; Carbone, M.S.; Richardson, R.A.D. The model-data fusion pitfall: Assuming certainty in an uncertain world. *Oecologia* **2011**, *167*, 587–597. [[CrossRef](#)] [[PubMed](#)]
24. Thum, T.; Zaehle, S.; Köhler, P.; Aalto, T.; Aurela, M.; Guanter, L.; Kolari, P.; Laurila, T.; Lohila, A.; Magnani, F.; et al. Modelling sun-induced fluorescence and photosynthesis with a land surface model at local and regional scales in northern Europe. *Biogeosciences* **2017**, *14*, 1969–1987. [[CrossRef](#)]
25. Richardson, A.D.; Williams, M.; Hollinger, D.Y.; Moore, D.J.; Dail, D.B.; Davidson, E.A.; Scott, N.A.; Evans, R.S.; Hughes, H.; Lee, J.T.; et al. Estimating parameters of a forest ecosystem C model with measurements of stocks and fluxes as joint constraints. *Oecologia* **2010**, *164*, 25–40. [[CrossRef](#)]
26. Bloom, A.A.; Williams, M. Constraining ecosystem carbon dynamics in a data-limited world: Integrating ecological “common sense” in a model–data fusion framework. *Biogeosciences* **2015**, *12*, 1299–1315. [[CrossRef](#)]
27. Fox, A.; Williams, M.; Richardson, A.D.; Cameron, D.; Gove, J.H.; Quaipe, T.; Ricciuto, D.; Reichstein, M.; Tomelleri, E.; Trudinger, C.M.; et al. The REFLEX project: Comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data. *Agric. For. Meteorol.* **2009**, *149*, 1597–1615. [[CrossRef](#)]

28. Friend, A.D.; Arneth, A.; Kiang, N.Y.; Lomas, M.; Ogeer, J.; RÖdenbeck, C.; Running, S.W.; Santaren, J.-D.; Sitch, S.; Viovy, N.; et al. FLUXNET and modelling the global carbon cycle. *Glob. Chang. Biol.* **2007**, *13*, 610–633. [[CrossRef](#)]
29. Dutta, D.; Schimel, D.S.; Sun, Y.; Van der Tol, C.; Frankenberg, C. Optimal inverse estimation of ecosystem parameters from observations of carbon and energy fluxes. *Biogeosciences* **2019**, *16*, 77–103. [[CrossRef](#)]
30. Li, Y.; Yang, H.; Wang, T.; Macbean, N.; Bacour, C.; Ciais, P.; Zhang, Y.; Zhou, G.; Piao, S. Reducing the uncertainty of parameters controlling seasonal carbon and water fluxes in Chinese forests and its implication for simulated climate sensitivities. *Glob. Biogeochem. Cycles* **2017**, *31*, 1344–1366. [[CrossRef](#)]
31. Safta, C.; Ricciuto, D.M.; Sargsyan, K.; Debusschere, B.; Najm, H.N.; Williams, M.; Thornton, P. Global sensitivity analysis, probabilistic calibration, and predictive assessment for the data assimilation linked ecosystem carbon model. *Geosci. Model Dev.* **2015**, *8*, 1899–1918. [[CrossRef](#)]
32. Hill, T.C.; Williams, M.; Moncrieff, G.B. Constraining ecosystem processes from tower fluxes and atmospheric profiles. *Ecol. Appl.* **2011**, *21*, 1474–1489. [[CrossRef](#)] [[PubMed](#)]
33. Williams, M.; Richardson, A.D.; Reichstein, M.; Stoy, P.C.; Peylin, P.; Verbeeck, H.; Carvalhais, N.; Jung, M.; Hollinger, D.Y.; Kattge, K.; et al. Improving land surface models with fluxnet data. *Biogeosciences* **2009**, *6*, 1341–1359. [[CrossRef](#)]
34. Kuppel, S.; Peylin, P.; Chevallier, F.; Bacour, C.; Maignan, F.; Richardson, A.D. Constraining a global ecosystem model with multi-site eddy-covariance data. *Biogeosci. Discuss.* **2012**, *9*, 3317–3380. [[CrossRef](#)]
35. Kondo, M.; Ichii, K.; Ueyama, M.; Mizoguchi, Y.; Hirata, R.; Saigusa, N. The role of carbon flux and biometric observations in constraining a terrestrial ecosystem model: A case study in disturbed forests in East Asia. *Ecol. Res.* **2013**, *28*, 893–905. [[CrossRef](#)]
36. Yu, G.; Chen, Z.; Piao, S.; Peng, C.; Ciais, P.; Wang, Q.; Li, X.; Zhu, X. High carbon dioxide uptake by subtropical forest ecosystems in the East Asian monsoon region. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 4910–4915. [[CrossRef](#)]
37. Fang, J.; Yu, G.; Liu, L.; Hu, S.; Chapin, F.S. Climate change, human impacts, and carbon sequestration in China. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 4015–4020. [[CrossRef](#)]
38. Williams, M.; Schwarz, P.A.; Law, B.E.; Irvine, J.; Kurpius, M.R. An improved analysis of forest carbon dynamics using data assimilation. *Glob. Chang. Biol.* **2005**, *11*, 89–105. [[CrossRef](#)]
39. Li, J.; Bai, Y.F.; Peng, Y.; Jiang, C.Q.; Xu, R. Carbon accounting of Chinese fir plantation in Huitong, Hunan province. *For. Res.* **2017**, *30*, 436–443.
40. Li, C.; He, H.; Liu, M.; Su, W.; Fu, Y.; Zhang, L.; Wen, X.; Yu, G. The design and application of CO₂ Flux Data Processing System at ChinaFLUX. *Geo-Inf. Sci.* **2008**, *10*, 557–565.
41. Zheng, Z.M.; Yu, G.R.; Sun, X.M.; Li, S.G.; Wang, Y.S.; Wang, Y.H.; Wang, Y.H.; Fu, Y.L.; Wang, Q.F. Spatio-temporal variability of soil respiration of forest ecosystems in China: Influencing factors and evaluation model. *Environ. Manag.* **2010**, *46*, 633–642. [[CrossRef](#)] [[PubMed](#)]
42. Zhu, G.; Ju, W.; Chen, J.M.; Gong, P.; Zhu, J. Foliage clumping index over China's landmass retrieved from the MODIS BRDF parameters product. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 2122–2137. [[CrossRef](#)]
43. Zhang, X.; Wu, K. Fine-root production and turnover for forest ecosystems. *Sci. Silvae Sin.* **2001**, *37*, 126–138.
44. Post, W.; Pastor, J.; Zinke, P.J.; Stangenberger, A.G. Global patterns of soil nitrogen storage. *Nature* **1985**, *317*, 613–616. [[CrossRef](#)]
45. Liu, H.; Tang, L.; Hu, B.; Liu, G.; Wang, Y.; Bai, F.; Chen, Y.; Cheng, Y.; Du, J.; Du, M.; et al. Photosynthetically active radiation dataset in Chinese Ecosystem Research Network (2005–2015). *China Sci. Data* **2017**, *2*, 1–10.
46. Ge, R.; He, H.; Ren, X.; Zhang, L.; Yu, G.; Smallman, T.L.; Zhou, T.; Yu, S.Y.; Luo, Y.; Xie, Z.; et al. Underestimated ecosystem carbon turnover time and sequestration under the steady state assumption: A perspective from long-term data assimilation. *Glob. Chang. Biol.* **2019**, *25*, 935. [[CrossRef](#)]
47. Ji, J. A climate-vegetation interaction model: Simulating physical and biological processes at the surface. *J. Biogeogr.* **1995**, *22*, 445. [[CrossRef](#)]
48. Luo, Y.; Ahlström, A.; Allison, S.D.; Batjes, N.H.; Brovkin, V.; Carvalhais, N.; Chappell, A.; Ciais, P.; Davidson, E.A.; Finzi, A.; et al. Toward more realistic projections of soil carbon dynamics by Earth system models. *Glob. Biogeochem. Cycles* **2016**, *30*, 40–56. [[CrossRef](#)]
49. Braswell, B.H.; Sacks, W.J.; Linder, E.; Schimel, D.S. Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations. *Glob. Chang. Biol.* **2005**, *11*, 335–355. [[CrossRef](#)]

50. Wutzler, T.; Carvalhais, N. Balancing multiple constraints in model-data integration: Weights and the parameter block approach. *J. Geophys. Res. Biogeosci.* **2014**, *119*, 2112–2129. [[CrossRef](#)]
51. Wu, X.; Luo, Y.; Weng, E.; White, L.; Ma, Y.; Zhou, X. Conditional inversion to estimate parameters from eddy-flux observations. *J. Plant Ecol.* **2009**, *2*, 55–68. [[CrossRef](#)]
52. Wang, Y.P.; Baldocchi, D.; Leuning, R.; Falge, E.; Vesal, T. Estimating parameters in a land-surface model by applying nonlinear inversion to eddy covariance flux measurements from eight FLUXNET sites. *Glob. Chang. Biol.* **2007**, *13*, 652–670. [[CrossRef](#)]
53. Peylin, P.; Bacour, C.; MacBean, N.; Leonard, S.; Rayner, P.; Kuppel, S.; Koffi, E.; Kane, A.; Maignan, F.; Chevallier, F.; et al. A new stepwise carbon cycle data assimilation system using multiple data streams to constrain the simulated land surface carbon cycle. *Geosci. Model Dev.* **2016**, *9*, 3321–3346. [[CrossRef](#)]
54. Chuter, A.M.; Aston, P.J.; Skeldon, A.C.; Roulstone, I. A Dynamical Systems Analysis of the Data Assimilation Linked Ecosystem Carbon (DALEC) Models. *Chaos* **2015**, *25*, 1–43. [[CrossRef](#)] [[PubMed](#)]
55. Trugman, T.A.; Medvigy, D.; Mankin, J.S.; Anderegg, W.R.L. Soil moisture stress as a major driver of carbon cycle uncertainty. *Geophys. Res. Lett.* **2018**, *45*, 6195–6502. [[CrossRef](#)]
56. Ubierna, N.; Cernusak, L.A.; Holloway-Phillips, M.; Busch, F.A.; Cousins, A.B.; Farquhar, G.D. Critical review: Incorporating the arrangement of mitochondria and chloroplasts into models of photosynthesis and carbon isotope discrimination. *Photosynth. Res.* **2019**, *141*, 5–31. [[CrossRef](#)]
57. Zhang, L.; Yu, G.R.; Luo, Y.; He, H.L.; Zhang, L.M. Based on the model of data fusion in broadleaved Korean pine forest carbon cycling modeling. *Chin. J. Plant Ecol.* **2009**, *33*, 1044–1055.
58. Gao, C.; Wang, H.; Weng, E.; Lakshmivarahan, S.; Zhang, Y.; Luo, Y. Assimilation of multiple data sets with the ensemble kalman filter to improve forecasts of forest carbon dynamics. *Ecol. Appl.* **2011**, *21*, 1461–1473. [[CrossRef](#)]
59. Zhu, P.; Shi, L.; Zhu, Y.; Zhang, Q.; Huang, K.; Williams, M. Data assimilation of soil water flow via ensemble Kalman filter: Infusing soil moisture data at different scales. *J. Hydrol.* **2017**, *555*, 912–925. [[CrossRef](#)]
60. Hill, T.C.; Ryan, E.; Williams, M. The use of CO₂ flux time series for parameter and carbon stock estimation in carbon cycle research. *Glob. Chang. Biol.* **2011**, *18*, 179–193. [[CrossRef](#)]
61. Enting, I.G. Assessing the information content in environmental modelling: A carbon cycle perspective. *Entropy* **2008**, *10*, 556–575. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).