Ecological Modelling 379 (2018) 39–53

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Modeling and uncertainty analysis of carbon and water fluxes in a broad-leaved Korean pine mixed forest based on model-data fusion

Xiaoli Ren[a], Honglin He[a,b,*], Li Zhang[a,b], Fan Li[c], Min Liu[d], Guirui Yu[a,b], Junhui Zhang[e]

[a] Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[b] College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100190, China
[c] China National Offshore Oil Corporation Research Institute, Beijing 100028, China
[d] Shanghai Key Laboratory for Urban Ecological Processes and Eco-Restoration, East China Normal University, Shanghai 200062, China
[e] Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110016, China

ARTICLE INFO

Keywords:
Sensitivity analysis
Uncertainty analysis
Morris global sensitivity analysis method
Markov Chain Monte Carlo (MCMC)
Sobol' variance decomposition method

ABSTRACT

Process-based ecosystem models are increasingly used to estimate the carbon and water exchanges between ecosystems and the atmosphere. These models inevitably suffer from deficiencies and uncertainties, which should be thoroughly examined to better understand the processes governing the ecosystem dynamics. In this paper, we systematically explored the uncertainties in model predictions of Changbaishan (CBS) broad-leaved Korean pine mixed forest using the SImplified PhotosyNthesis and Evapo-Transpiration (SIPNET) model and eddy flux and meteorological data from 2004 to 2009. We first screened out 21 key parameters from 42 model parameters using Morris global sensitivity analysis method, and then estimated their probability distributions through Markov Chain Monte Carlo technique. Two optimization set-ups, i.e. using observed net ecosystem exchange of CO_2 (NEE) only and using observed NEE and evapotranspiration (ET) simultaneously, were conducted to detect the different constraints of different observations on model parameters. Four parameters were well constrained using observed NEE only, including photosynthesis and respiration related parameters. While seven parameters were well constrained using measured NEE and ET simultaneously, four of which were water related parameters. Obviously, more information can be derived from the simultaneous optimization, since there was additional process information in water flux observation. The modeled ET of the NEE and ET optimization set-up had a much better fit to measured values than the NEE only optimization set-up (R^2 = 0.70 vs. R^2 = 0.30), although the modeled NEE from the two set-ups had a good fit to the observations (R^2 = 0.85 vs. R^2 = 0.83). This implied that assimilating carbon and water fluxes simultaneously can improve the parameterization and overall performance of the model. Then, we quantified the uncertainties in model predictions using Monte Carlo simulation, and trace them to specific parameter and parameter interactions through Sobol' variance decomposition method. The uncertainties of five outputs of interest in CBS site, NEE, gross primary productivity (GPP), ecosystem respiration (RE), ET and transpiration (T), were 50.82%, 22.35%, 21.25%, 9.98% and 19.54%, respectively. The uncertainty in predicted NEE was much larger since NEE is a small difference between two large fluxes, i.e. GPP and RE. The maximum net CO_2 assimilation rate (A_{max}) and carbon content of leaves (*SLW*) were classified as highly sensitive parameters for all outputs of interest in CBS site, contributing more than 70% of the uncertainties in all outputs except NEE. The importance of these two parameters holds for one subtropical evergreen coniferous plantation and one subtropical evergreen broad-leaved forest, too. Therefore, these two parameters and their underlying processes should be a focus of future model research, plant trait data collection and field measurement, at least for the sites in this study. This can help connect the model simulation research and field data collection, making them mutually informative.

1. Introduction

Forest ecosystems, functioning as a significant carbon sink, play an important role in the global terrestrial carbon cycle (Dixon et al., 1994; Fang et al., 2001; Pan et al., 2011). In the context of a changing climate, understanding the exchanges of mass and energy between forests and

* Corresponding author at: 11A, Datun Road, Chaoyang District, Beijing 100101, China.
E-mail address: hehl@igsnrr.ac.cn (H. He).

https://doi.org/10.1016/j.ecolmodel.2018.03.013
Received 17 March 2017; Received in revised form 28 March 2018; Accepted 29 March 2018
Available online 27 April 2018
0304-3800/

atmosphere is of paramount importance since the response of the terrestrial biosphere is still one of the largest sources of uncertainty in future climate projections (Ballantyne et al., 2017; Bonan, 2008; Dietze et al., 2013; Friedlingstein et al., 2014; Heimann and Reichstein, 2008; Sitch et al., 2008). With the rapid development of eddy flux networks, remote sensing techniques, and other observing systems, large amounts of data have been collected, covering different spatial and temporal scales (Raupach et al., 2005; Scholze et al., 2017). However, no one data source incorporates all the information about the underlying ecosystem processes (Dietze et al., 2013).

The great need for understanding the ecosystem processes contributed to the evolution of process-based ecosystem models, which are increasingly used to estimate the carbon and water fluxes of ecosystems (Aber and Federer, 1992; Bonan et al., 2002; Dufrene et al., 2005; Pappas et al., 2013; Sellers et al., 1986). However, different models have different degrees of complexity and abstraction, inherently uncertain in the case of model structure (Friedlingstein et al., 2014; Lin et al., 2011; Updegraff et al., 2010). In addition, uncertainty also stems from imprecisely known model parameters and driving variables. However, the observations of model parameters are greatly lacking or not available, or contain errors (Lin et al., 2011; Medlyn et al., 2005). Parameter uncertainty is usually a dominant source of uncertainty since process-based ecosystem models are complex and contain many parameters (Dietze et al., 2013).

The uncertainties in simulated carbon and water fluxes have been acknowledged being as important as the predicted values themselves (Larocque et al., 2008). In order to reduce these uncertainties, the first step would be to quantify them and determine the contribution of each model parameter, thereby helping us better understand which error source should get priority and be constrained in the future. There have been large quantities of literature to quantify model uncertainty, but the studies that systematically quantify the uncertainties in model outputs and trace to specific parameters and their interactions are relatively few (Dietze et al., 2013; Larocque et al., 2008; Lin et al., 2011; MacBean et al., 2016; Ren et al., 2013a; Smith and Heath, 2001). Monte Carlo (MC) method was applied most often to quantify the output uncertainty, and it can also rank parameters for uncertainty contribution when combined with multiple linear regressions (Larocque et al., 2008; Verbeeck et al., 2006; Zhang et al., 2012). One limitation of MC method is that it cannot quantify the uncertainty contributions of model parameter interactions. While Sobol' variance decomposition method can apportion the total output variance to components that stem from each parameter and the interactions between them (Pianosi et al., 2017; Saltelli, 2002; Sobol, 1993; Tarantola et al., 2006; Safta et al., 2015).

The combination of MC and Sobol' methods will effectively accomplish the uncertainty quantification and partitioning of the model predictions. However, the bottleneck in implementation is the specification of uncertainty of model parameters, expressed as probability density functions (PDFs), where information are limited (Dietze et al., 2013; Keenan et al., 2013; Larocque et al., 2008; Radtke et al., 2001). Model-data fusion technique (MDF), which combines the prior information about parameters and observations such as carbon and water fluxes, provides a way to directly inform the model and derive posterior probability distributions of model parameters (MacBean et al., 2016; Raupach et al., 2005; Ricciuto et al., 2011; Williams et al., 2005; Zobitz et al., 2011). Since process-based models have large number of parameters, it would be helpful to identify key parameters controlling model behavior through parameter screening using sensitivity analysis approach at first (Quillet et al., 2013; Tang et al., 2007a).

We have proposed a methodological framework for systematic uncertainty analysis of process-based ecosystem models in a previous paper, combining the aforementioned four procedures, i.e. sensitivity analysis method (one-at-a-time, OAT), MDF technique (Markov Chain Monte Carlo, MCMC), MC uncertainty quantification method and Sobol' variance decomposition technique (Ren et al., 2013a). The framework provides a paradigm for systematic and quantitative uncertainty analysis of ecosystem models and has been successfully applied in a subtropical coniferous plantation of China (Ren et al., 2013a). However, it has not been tested and employed in other types of ecosystems, and still has several limitations needing further research. One of the limitations is that the OAT method is a local sensitivity analysis method, which is very sensitive to the initial values of model parameters. In this research, we replace the OAT method with Morris global sensitivity analysis technique, which is more appropriate when the model is non-linear and non-monotonic, to overcome this limitation, and then use the improved framework to conduct systematic uncertainty analysis with a simple process-based model for a broad-leaved Korean pine mixed forest.

The overall objectives of this paper are threefold: (1) determine the key model parameters of the simplified photosynthesis and evapotranspiration (SIPNET) model in Changbaishan (CBS) broad-leaved Korean pine mixed forest in China through Morris global sensitivity analysis; (2) perform two optimization set-ups, i.e. carbon flux only, and both carbon and water fluxes assimilated, to compare the estimated posterior distributions of key parameters and the informed model performance; (3) quantify the uncertainty in predicted carbon and water fluxes based on the better informed model, and determine how much each key parameter and their interactions contribute to the prediction uncertainty, thereby helping identify future research priorities, at least for this specific site.

2. Material and methods

2.1. Study site and observational data

The Changbaishan forest site (CBS, 42°24′N, 128°05′E, 738 m) of ChinaFLUX is a broad-leaved Korean pine mixed forest located in Jilin province, which is the dominant vegetation type of northeastern China (Yu et al., 2006; Zhang et al., 2006b). The forest has a temperate continental climate controlled by monsoon (Zhang et al., 2012). The annual mean temperature is 3.6 °C, and annual precipitation is 695 mm (Guan et al., 2005). The forest is dominated by Korean pine (*Pinus koriaensis*), Tuan linden (*Tilia amurensis*), Mono Maple (*Acer mimo*), Manchurian ash (*Fraxinus mandshurica*), Mongolian oak (*Quercus monogolica*), elm (*Ulmus glabra*), with an average forest age of more than 200-year-old and average canopy height of about 26 m (Guan et al., 2005; Zhang et al., 2006b). There is 40% coverage of understory consisting of multi-species broad-leaved shrub, and the soil is dark brown forest soil (Zhang et al., 2006a).

The data used in this paper is half-hourly meteorological and eddy covariance flux data of CBS site from January 2004 to December 2009. The data of the first three years are applied to inform the model, i.e. estimate the posterior probability distributions of model parameters, and the remaining three years are used to validate the informed model. The data of the year 2004 are selected to perform the parameter sensitivity analysis and uncertainty analysis of model predictions. The specific data employed are air temperature, soil temperature, precipitation, photosynthetic active radiation (PAR), water vapor pressure, wind speed, net ecosystem exchange of CO_2 (NEE), evapotranspiration (ET), gross primary productivity (GPP) and ecosystem respiration (RE). The ChinaFLUX data processing system (Li et al., 2008; Liu et al., 2009, 2012) was used to conduct quality control, gap-filling and separation of the flux data. The gaps in carbon flux data were filled using the non-linear regression method (Michaelis-Menten and Lloyd&Tayor equations), and the separation was conducted in the meantime. The latent heat flux was gap-filled using look-up table method, and divided by the product of the Latent Heat of Vaporization and the water density to derive ET.

Because the model is run at half-daily time step (day and night), we aggregated the half-hourly meteorological and flux data to half-daily scale. The length of each day or night time step is calculated based on the determination of sunrise and sunset times using the latitude of CBS