Agricultural and Forest Meteorology 256–257 (2018) 445–457

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

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

Estimating bamboo forest aboveground biomass using EnKF-assimilated MODIS LAI spatiotemporal data and machine learning algorithms

Xuejian Li[a,b,c,1], Huaqiang Du[a,b,c,*,1], Fangjie Mao[a,b,c,1], Guomo Zhou[a,b,c,*], Liang Chen[a,b,c], Luqi Xing[a,b,c], Weiliang Fan[a,b,c], Xiaojun Xu[a,b,c], Yuli Liu[a,b,c], Lu Cui[a,b,c], Yangguang Li[a,b,c], Dien Zhu[a,b,c], Tengyan Liu[a,b,c]

[a] *State Key Laboratory of Subtropical Silviculture, Lin'an 311300, Zhejiang, China*
[b] *Key Laboratory of Carbon Cycling in Forest Ecosystems and Carbon Sequestration of Zhejiang Province, Zhejiang A & F University, Lin'an 311300, Zhejiang Province, China*
[c] *School of Environmental and Resources Science, Zhejiang A & F University, Lin'an 311300, Zhejiang Province, China*

ARTICLE INFO

Keywords:
MODIS LAI
EnKF
Data assimilation
Random forest
AGB
Bamboo forest

ABSTRACT

High-precision LAI (leaf area index) spatiotemporal data obtained from MODIS satellite remote sensing products are important for studying vegetation growth status, biomass carbon reserves, and the spatiotemporal dynamics of carbon cycling. LAI significantly influences biomass accumulation during the growth of bamboo forest in subtropical zones. Therefore, we applied the ensemble Kalman filter (EnKF) data assimilation algorithm to assimilate MODIS LAI products, and used assimilated LAI and the normalized difference vegetation index, enhanced vegetation index, simple ratio index as variables in the random forest model to estimate bamboo forest above ground biomass (AGB) in Zhejiang Province. Assimilated LAI spatiotemporal data using EnKF greatly improve the accuracy of MODIS LAI products, the R^2 between assimilated and observed LAI was 0.92, and the RMSE was 0.37. Variations in the assimilated LAI time series were consistent with the seasonal dynamics of bamboo forest growth and had a significant effect on AGB. Moreover, the random forest model had strong predictive capabilities. A comparison of training and testing results produced accuracy (R) values for the random forest model using the assimilated LAI time series of 0.71 and 0.73, respectively. Using the assimilated LAI achieved a more accurate AGB estimate than using MODIS LAI time series products, as the R values were 54.3% and 58.7% higher, and the RMSE values were 19.2% and 19.1% lower for training and testing results, respectively. The calculated spatial distribution of bamboo forest AGB in Zhejiang province was consistent with the observed values. By combining assimilation technology of the MODIS LAI time series with the random forest model to more accurately estimate bamboo forest AGB in Zhejiang province, this study provided a new method for estimating large scale forest AGB based on low-resolution time series data.

1. Introduction

The leaf area index (LAI) is defined as one-half of the total area of the plant (Chen and Black, 1992), and changes in LAI time series data can reflect the growth status of vegetation. Thus, it is considered an important parameter for research into carbon and water cycling and the energy exchange of terrestrial ecosystems (Dong et al., 2016; He et al., 2015; Jonckheere et al., 2004; Zhang et al., 2014).

Remote sensing data, such as the MODIS LAI time-series products, were widely used data sources for studying the dynamics of vegetation LAI and carbon cycles (Dong et al., 2016; Li et al., 2017; Xiao et al., 2011), providing the only viable option for obtaining spatiotemporal LAI information (Gray and Song, 2012). However, due to the impacts of cloud cover, aerosols, snow cover, and sensor failure, many existing satellite-based LAI products are characterized by high noise, low accuracy, and large fluctuations in the time series, which cannot adequately reflect plant growth continuity, thereby constraining the global application of LAI products (Fang et al., 2012; Heinsch et al., 2006; Li et al., 2017; Xiao et al., 2011). Data assimilation methods, which were widely used in remote sensing, meteorological, soil, ecological, and hydrological studies, constitute an important data processing step for improving data quality and data accuracy (Gu et al., 2009; Li et al., 2014; Mclaughlin, 2002; Moradkhani et al., 2005). High-precision LAI time series, which have made some progress in estimating vegetation

* Corresponding authors at: State Key Laboratory of Subtropical Silviculture, Lin'an 311300, Zhejiang, China.
E-mail addresses: dhqrs@126.com (H. Du), zhougm@zafu.edu.cn (G. Zhou).
[1] These authors contributed equally to this study and shared first authorship.

https://doi.org/10.1016/j.agrformet.2018.04.002
Received 21 November 2017; Received in revised form 19 February 2018; Accepted 5 April 2018
0168-1923/

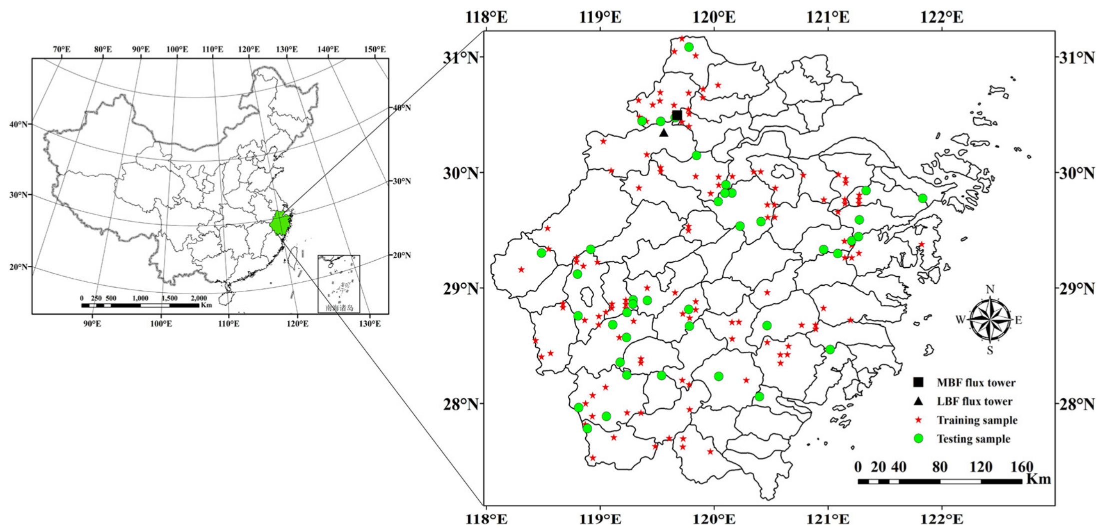

Fig. 1. Location of the study area and bamboo forest sample plots.

biomass and simulating the carbon cycle, are obtained through data assimilation techniques. For instance, He et al. (2015) improved the precision of aboveground dry biomass estimation for grass by coupling the crop growth model and grassland LAI using a 4-D variation data assimilation algorithm. Dong et al. (2016) used the ensemble Kalman filter (EnKF) to assimilate LAI into the simple algorithm for yield estimation model, which improved the estimation accuracy of winter wheat biomass. Li et al. (2016) assimilated MODIS LAI time series data using EnKF techniques coupled with the PROSAIL model, and then used in simulating of the carbon fluxes of bamboo forest through BEPS model, and reported assimilated LAI can improve the simulation of carbon flux in bamboo forests. Li et al. (2017) assimilated MODIS LAI time series data using EnKF techniques coupled with the PROSAIL model, and reported the results had high R^2 about 0.87. Mao et al. (2017b) used Dual EnKF and PROSAIL model to improve the quality of MODIS LAI time-series data, and reported the results had R about 0.93. Therefore, the use of time series remote sensing data has great potential for improving the accuracy of large-scale above ground biomass (AGB) estimation (Blackard et al., 2008; Dong et al., 2003; Dymond et al., 2002).

Estimating forest biomass and carbon storage by remote sensing is a popular domestic and international research topic (Dong et al., 2016; Lu, 2006; Pham and Brabyn, 2017; Yan et al., 2015). Bamboo forest resources have a large carbon sequestration capacity and play a significant role in maintaining the regional ecological environment and global carbon balance (Mao et al., 2016; Zhou and Jiang, 2004). Estimating bamboo forest biomass and carbon storage based on remote sensing data has progressed substantially in recent years (Du et al., 2012; Han et al., 2013; Xu et al., 2012; Zhou et al., 2011) through the application of different parametric and nonparametric models (Du et al., 2012; Shang et al., 2013; Xu et al., 2011; Zhou et al., 2011). Nonparametric models typically have better prediction accuracy because they do not make assumptions about the distribution of data (Du et al., 2012; Tian et al., 2014; Xu et al., 2011). Random Forest (RF) is the most important nonparametric machine learning model; the advantages of RF include high accuracy of prediction outcomes, high computational efficiency, robustness to outliers and noise, and the ability to estimate the importance of predictor variables. In recent years, RF has been widely used to estimate forest biomass (Feng et al., 2017; Greaves et al., 2016; Pham and Brabyn, 2017; Zhu and Liu, 2015).

Bamboo, such as Moso bamboo, is a gramineous plant with fast growth characteristics. It takes just about 50 days for bamboo to grow from new shoots to full height, after which leaf expanding begins and LAI gradually increases, leading to the rapid accumulation of biomass and carbon storage (Li et al., 2017; Lu et al., 2012). Previously, single-phase satellite remote sensing data was used to build bamboo models for estimating forest AGB, which made it difficult to reflect the variation characteristics of rapid bamboo growth. Du et al. (2012) used Landsat images to estimate moso bamboo biomass by multiple-linear regression and reported an R of 0.36. Shang et al. (2013) used Landsat and MODIS images to estimate Moso bamboo forest aboveground carbon storage based on multi-source remotely sensed images, and reported an R of 0.70. LAI time series data not only can reflect the changes in a bamboo forest as a time series, but also can express the dynamic accumulation of biomass during rapid bamboo growth. Estimating bamboo forest AGB with high-precision, spatiotemporal LAI data will make the estimation result more biological.

This study uses MODIS LAI products from Zhejiang province in 2014 as the remote sensing data source, and adopts the EnKF data assimilation algorithm to assimilate the MODIS LAI products and obtain high-precision, spatiotemporal LAI data of bamboo forest resources in Zhejiang province. The normalized difference vegetation index (NDVI), enhanced vegetation index (EVI) and simple ratio index (SR) and LAI data was then used as variables in the random forest model to estimate bamboo forest AGB, and model results were evaluated using observed AGB. Our results illustrate a new method for estimating large-scale forest AGB using remote sensing data.

2. Materials and methods

2.1. Study area

Zhejiang Province (118°01′–123°10′E, 27°06′–31°11′N) is located on the southeast coast of China, bordered to the east by the east China sea, to the south by Fujian Province, to the west by Anhui and Jiangxi Provinces, and to the north by Shanghai, Jiangsu and Anhui Provinces (Fig. 1). Zhejiang Province is located in the central subtropical zone, which has a typical subtropical monsoon climate with abundant light, mild climate and rainfall, and distinct seasons. The annual average temperature is 15 °C –18 °C, the annual sunshine hours are 1100–2200 h, and the average annual precipitation is 1100–2000 mm. Due to the influence of the ocean, the temperature and humidity conditions are more suitable for bamboo growth than inland monsoon areas at the same latitude. Currently, the area of bamboo forest in Zhejiang province is 0.83 million ha, which is one of the largest bamboo distribution province in China. Moso bamboo forest occupies approximately 87% of the total area of bamboo forest, and other bamboo