Vegetation biomass change in China in the 20th century: an assessment based on a combination of multi-model simulations and field observations

Song, X., Li, F. ORCID: https://orcid.org/0000-0002-3686-2257, Harrison, S. P., Luo, T., Arneth, A., Forrest, M. ORCID: https://orcid.org/0000-0003-1858-3489, Hantson, S., Lasslop, G., Mangeon, S., Ni, J., Yue, C., Hickler, T., Luo, Y., Sitch, S., Xu, X. and Zhu, Z. (2020) Vegetation biomass change in China in the 20th century: an assessment based on a combination of multi-model simulations and field observations. Environmental Research Letters, 15 (9). 094026. ISSN 1748-9326 doi: https://doi.org/10.1088/1748-9326/ab94e8 Available at http://centaur.reading.ac.uk/92582/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1088/1748-9326/ab94e8
All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading’s research outputs online
LETTER • OPEN ACCESS

Vegetation biomass change in China in the 20th century: an assessment based on a combination of multi-model simulations and field observations

To cite this article: Xiang Song et al 2020 Environ. Res. Lett. 15 094026

View the article online for updates and enhancements.
Environmental Research Letters

LETTER

Vegetation biomass change in China in the 20th century: an assessment based on a combination of multi-model simulations and field observations

Xiang Song1, Fang Li1,2, Sandy P Harrison3, Tianxiang Luo1, Almut Arneth1, Matthew Forrest4, Stijn Hantson1, Gitta Lasslop5, Stephane Mangeon6, Jian Ni7, Chao Yue1,2, Thomas Hickler8, Yiqi Luo1, Stephen Sitch9, Xin Xu10 and Zaichun Zhu11

1 International Center for Climate and Environment Sciences, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, People's Republic of China
2 School of Archaeology, Geography & Environmental Science, Reading University, Reading, United Kingdom
3 Department of Earth System Science, Tsinghua University, Beijing, People's Republic of China
4 Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, People's Republic of China
5 Karlsruhe Institute of Technology, Institute of Meteorology and Climate research, Atmospheric Environmental Research, Garmisch-Partenkirchen, Germany
6 Senckenberg Biodiversity and Climate Research Centre (BiK-F), Senckenberganlage, Germany
7 Geospatial Data Solutions Center, University of California, Irvine, CA, United States of America
8 Department of Physics, Imperial College London, London, United Kingdom
9 Now at CSIRO, Data61, Brisbane, QLD, Australia
10 Zhejiang Normal University, Jinhua, Zhejiang, People's Republic of China
11 State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Northwest A&F University, Yangling, Shanxi, People's Republic of China
12 Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, France
13 Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, United States of America
14 College of Life and Environmental Sciences, University of Exeter, Exeter, United Kingdom
15 School of Urban Planning and Design, Peking University Shenzhen Graduate School, Shenzhen, People's Republic of China
16 Author to whom any correspondence should be addressed.

E-mail: lifang@mail.iap.ac.cn

Keywords: vegetation biomass, China, 20th century changes, land use and land cover change, climate change, CO₂ concentration, DGVM simulations

Abstract

Vegetation biomass is a key and active component of the carbon cycle. Though China's vegetation biomass in recent decades has been widely investigated, only two studies have quantitatively assessed its century-scale changes so far and reported totally opposite trends. This study provided the first multi-model estimates of China's vegetation biomass change for the 20th century and its responses to historical changes in environmental and anthropogenic factors, based on simulations evaluated with the field observations from 3757 inventory plots in China and bias-corrected using machine learning (Gaussian process regression). A significant decline in vegetation biomass over the 20th century was shown by bias-corrected simulations from the six Dynamic Global Vegetation models (DGVMs) with trends ranging from −32.48 to −11.10 Tg C yr⁻¹ and a mean trend of −17.74 Tg C yr⁻¹. Land use and land cover change (LULCC) was primarily responsible for the simulated downward trend (−50.71 to −24.28 Tg C yr⁻¹), while increasing atmospheric CO₂ concentration lead to increased vegetation biomass (+9.27 to +13.37 Tg C yr⁻¹). Climate change had limited impacts on the long-term trend (−3.75 to +5.06 Tg C yr⁻¹). This study highlights the importance of LULCC for historical reconstruction and future projection of vegetation biomass over China. It also suggests that the incorrect change in China’s forest area for 1980–2000 in the LULCC dataset used as model input data of many existing and ongoing model intercomparison projects (MIPs) has likely led to inaccurate estimations of historical vegetation biomass changes in China.

© 2020 The Author(s). Published by IOP Publishing Ltd
1. Introduction

China is the most populous country and the largest carbon emitter in the world, and has undergone distinct land-use and climate changes during the 20th century (Fang et al. 2018). Assessment of vegetation biomass and its change over China is essential for understanding the regional, and even the global, carbon cycle (Kondo et al. 2013, Fang et al. 2018).

Many studies have estimated China's vegetation biomass in recent decades using various approaches. Estimates of present-day vegetation biomass in China range from 13.34 to 35.23 Pg C based on numerical models (Li et al. 2004, Ni 2001, and references therein), and from 13.09 to 34.68 Pg C based on field and satellite observations (Carvalhais et al. 2014, Li et al. 2015, Xu et al. 2018, Fang et al. 2018; and references therein). Fang et al. (2018) and Tang et al. (2018) estimated that forests, grasslands, shrublands, and croplands contributed to 80%, 10%, 5%, and 4% of current vegetation biomass, respectively, using extrapolation from site-level field observations based on a land cover map estimate. Xu et al. (2018) investigated the spatial changes of 2004–2014 vegetation biomass by collecting data based on literature review, and concluded that mean annual precipitation is an important driver of the spatial changes. Several studies reported that vegetation biomass increased after ~1980 probably as a result of forest expansion and regrowth (Fang et al. 2001, 2007, Tang et al. 2018).

However, until now, only Wang et al. (2007) and Mao et al. (2009) have investigated the long-term change of China’s vegetation biomass over the 20th century. Both studies were based on a single model, specifically, the Integrated Terrestrial Ecosystem C-budget model (InTEC) in Wang et al. (2007) and a modified version of the Sheffield Dynamic Global Vegetation Model (SDGVM) in Mao et al. (2009). The two studies reported opposite trends: China’s vegetation biomass for the 20th century was decreased by 2.15 Pg C from Wang et al. (2007) while increased by 2.07 Pg C from Mao et al. (2009).

Historical changes in vegetation biomass are regulated by climate, atmospheric CO\(_2\) concentration, and land use and land cover change (LULCC) (Kim et al. 2017, Fang et al. 2018, Xin et al. 2018). Dynamic Global Vegetation Models (DGVMs) simulate the response of global terrestrial ecosystems to these drivers, and integrate biogeochemistry, biogeophysics, and vegetation dynamics at the land surface within a physically and chemically consistent modelling framework (e.g., Foley et al. 1996; Sitch et al. 2003). They can be used to isolate how each driver affects vegetation biomass, providing that the models produce (or can be induced to produce via bias correction) realistic simulations of the modern state (Piao et al. 2011, Ichii et al. 2013, Chen et al. 2016). However, up to now, there has been no attempt to evaluate vegetation biomass simulated by DGVMs or correct their simulation bias before applying them to investigate the historical changes and drivers of global and regional vegetation biomass. The latter is important given that DGVMs aim to model the global simulations with low spatial resolution and generally do not conduct regional optimization. Furthermore, although Mao et al. (2009) found that atmospheric CO\(_2\) concentration change was more important than changes in temperature and precipitation in explaining vegetation biomass change of the 20th century, there has been no study to investigate the potential impact of other drivers on the historical change.

In this study, we used simulations by six state-of-the-art, fire-enabled DGVMs which were run following the same protocol and the same forcing datasets. We evaluated these simulations using site-specific biomass data from 3757 inventory plots in China. Biases in each model simulation were statistically corrected using machine learning (Gaussian process regression, GPR). The bias-corrected simulations provided the first multi-model estimates of vegetation biomass for the 20th century. We then used three sensitivity experiments in which potential drivers of 20th century vegetation changes, specifically, climate, CO\(_2\) concentration and land-use, were constant, respectively, throughout the historical period to analyze the causes of the simulated changes in vegetation biomass.

2. Data and methods

2.1. Model simulations

We used vegetation simulations performed as part of the Fire Model Intercomparison Project (FireMIP, Hantson et al. 2016, Rabin et al. 2017) from six state-of-the-art fire-enabled DGVMs: CLM4.5 with an updated fire module for CLM5 (CLM4.5), JULES-INFERNO (JULES), JSBACH-SPITFIRE (JSBACH), LPJ-GUESS-SPITFIRE (LPJ-GS), LPJ-GUESS-SIMFIRE-BLAZE (LPJ-GSB), and ORCHIDEE-SPITFIRE (ORCHIDEE). The distribution of natural vegetation was simulated in JULES, LPJ-GS, and LPJ-GSB, but prescribed in the others. Only CLM4.5 and LPJ-GSB simulated an interactive nitrogen cycle (table 1). All the DGVMs are not stochastic except for LPJ-GUESS. LPJ-GUESS includes stochastic representation of individual establishment, mortality, and disturbance impacts. This does not affect the interpretability of the results because the simulations used here feature sufficient (50 for LPJ-GS and 25 for LPJ-GSB) replications so that the effect of the stochastic variability is small and the effect of drivers is dominant.

All the DGVM simulations started from a spin-up simulation for 1700 conditions to get the initial fields for subsequent 1701–2013 simulations (figure 1, see Rabin et al. 2017 for details). For the spin-up runs, models were recycled until carbon values in the slowest soil carbon pool varied by less than 1%
between consecutive 50 year periods for every grid cell. The atmospheric reanalysis data were from CRU-NCEP v5.3.2. Annual global atmospheric CO₂ concentration was derived from ice core and NOAA monitoring station data (Le Quéré et al 2014). LULCC data were from LUH1 (Hurtt et al 2011) in which changes in fractional coverage of croplands and pastures were from History Database of the Global Environment (HYDE) version 3.1 (Klein Goldewijk et al 2011, Klein Goldewijk and Verburg 2013). China’s land use and land cover change in LUH1’s LULCC data was characterized by expansion of croplands and a decrease in forest area in the 20th century (Liu and Tian 2010).

Sensitivity simulations were made to help understand the drivers of simulated changes in terrestrial ecosystems, including: (1) no climate change, (2) constant atmospheric CO₂ concentration, and (3) no LULCC (table 2; see Rabin et al 2017 for details). In the no climate change simulation, the climate throughout the run was the recycled time-varying 1901–1920 atmospheric forcing, but all other inputs were allowed to vary as in the simulation. In the constant CO₂ simulation, the atmospheric CO₂ concentration was held constant at 277.33 ppm for the whole of the simulation. In the no LULCC simulation, land cover in 1700 was used throughout the simulation.

All model outputs were converted to 1º spatial resolution, based on bilinear interpolation from low to high resolution for CLM4.5 (∼1.9º latitude × 2.5º longitude), JSBACH (1.875º), and JULES (∼1.2º latitude × 1.9º longitude), and on area-weighted averaging for those models with an original spatial resolution of 0.5º.

Not all of the FireMIP models were used in this study. LPJ-GUESS-GlobFIRM and LPJ-GUESS-SIMFIRE-BLAZE had the same vegetation model but used different fire models. Evaluation by Andela et al (2017) and Li et al (2019) showed that the fire model GlobFIRM simulated the global burned area and fire carbon emissions poorly. MC2 was developed for regional applications and applied globally without adequate calibration (Hantson et al 2019). An old version of CLASS-CTEM outputs did not complete all the sensitivity experiments (Li et al 2019), and
some bugs were identified in temporal variability of carbon fluxes and pools in its new version. Therefore, the simulations of the three models (LPJ-GUESS, GlobFIRM, MC2 and CLASS-CTEM) were excluded from our analyses.

2.2. Field observations

The field observations of vegetation biomass used here were collated and derived from Chinese forest survey inventories of Luo (1996) and Pan et al (2004) and from literature review (supplement B.xls and supplement B reference.docx). The vegetation biomass is derived as the sum of leaf, stem and branches, and root biomass or the sum of above-ground and below-ground biomass.

The dataset from Luo (1996) and Pan et al (2004) focused on forest plots, and provided information on location, forest type, and leaf, branch, stem, and root biomass for trees for all plots. It included 1266 forest plots, sampling tropical and monsoon forests, sub-tropical evergreen broadleaf and coniferous forests, temperate deciduous broadleaf forests, boreal evergreen/deciduous coniferous forests where the measurements were made on forests aged from 3 to 350 years during 1956–1993. Most of the organ biomass (leaf, branch and stem, root) was obtained by the biomass allometric growth model method. Missing values of tree roots were calculated from the ratio of above-ground to below-ground biomass of the same species in nearby stands (Luo 1996).

Furthermore, we compiled an additional dataset of above-/below- ground vegetation biomass for 1112 forest plots and 1379 grassland plots measured during 2000–2013 from published literature. We selected plots with observations of both above- and below-ground biomass and a clear description of the functional types or species. Observations in the literature were mainly obtained by the clear-cutting method, average standard tree method, and/or biomass allometric growth method. A coefficient of 0.45 was used to convert vegetation biomass density to C density (kg C m⁻²) (e.g. Xu et al 2018).

2.3. Bias correction for vegetation biomass

To derive bias-corrected outputs from each of the DGVMs, we built statistical models between the default (uncorrected) vegetation biomass simulations and observations of vegetation biomass (total vegetation biomass) as follows. First, according to the vegetation-type or species records in the field observations and the plant functional type (PFT) information in the model outputs, we sorted the observations and simulations into evergreen broad-leaf, deciduous broad-leaf, evergreen needle-leaf, and deciduous needle-leaf tree PFTs and grass (not separating C3 and C4) PFT. Second, simulations were temporally averaged to match the representative periods of observations. Third, we built statistical models based on simulation-observation pairs that belonged to the same PFT over all grid cells.

Three methods for bias correction were tested: first order linear regression:

\[ B_{ij,obs} = a_{i,j,0} + a_{i,j,1} \times B_{ij,uncorrected} \]

(1)

cubic regression:

\[ B_{ij,obs} = a_{i,j,0} + a_{i,j,1} \times B_{ij,uncorrected} + a_{i,j,2} \times B_{ij,uncorrected}^2 + a_{i,j,3} \times B_{ij,uncorrected}^3 \]

(2)

and machine learning (Gaussian process regression, GPR). Here, \( a_{i,j,n} (n = 0, \ldots, 3) \) was regression coefficient for the \( i \)th functional type of the \( j \)th model, \( B_{ij,obs} \) and \( B_{ij,uncorrected} \) were the spatial arrays of field observations and the default (uncorrected) model outputs, respectively. For the machine learning GPR, we used the GPR package of Matlab version 2018b. We carried out batch training based on the GPR model to optimize the parameters of kernel function (ard squared exponential function). Hyperparameters were then determined by automatic hyperparameter optimization based on minimizing five-fold cross-validation loss.

We adopted the commonly-used leave-one-out cross-validation framework (LOOCV, Wilks 1995) to build and evaluate the statistical models based on the three bias-correction methods for preventing over-fitting. In LOOCV, one observation-simulation pair was left out as test set, and the statistical model was built by using remaining observation-simulation pairs \((n - 1)\) as training set. Then, the bias-corrected simulation value was predicted with the statistical model and the independent left-out simulation (uncorrected) value. The process was repeated \(n\) times to obtain \(n\) bias-corrected simulation values. The correlation coefficient \((R)\) and root mean square error (RMSE) between simulations and field

| Exp. name | Climate | CO₂ | LULCC |
|-----------|---------|-----|-------|
| CTRL | Transient | Transient | Transient |
| S-CLIMATE | 1901–1920 cycled | Transient | Transient |
| S-CO₂ | Transient | 277.33 ppm | Transient |
| S-LULCC | Transient | Fixed at 1700 values |

*Repeatedly cycling 1901–1920 atmospheric forcing (precipitation, temperature, specific humidity, surface pressure, wind speed, and solar radiation).*

Table 2. Experiment design of control (CTRL) and three sensitivity simulations.
observations were used as skill scores (skill is higher with higher R and lower RMSE). We selected the method with highest cross-validation skill as the correction method for all models and functional types which was denoted as correction function $f$. Analyses of cross-validation skill (table S1 (available online at stacks.iop.org/ERL/15/094026/mmedia)) showed that machine learning GPR outperformed the first order linear and cubic regressions for all functional types and models except deciduous forest in JULES.

We derived the bias-corrected biomass simulations ($B_{i,j,k,corrected}$) for functional type $i$ of model $j$ in year $k$ during 1950–2013 as:

$$B_{i,j,k,corrected} = f(B_{i,j,k,uncorrected})$$

Then, the total vegetation biomass ratio between with and without bias-correction averaged during 1950–2000 was used as bias-correct coefficient to modify the mean total vegetation biomass in 2001–2010, as well as trend in vegetation biomass change in the 20th century. This method assumes that the statistical relationship between the multi-year average of DGVM simulations and field observations is constant.

3. Results

3.1. Evaluation and bias correction

The models simulated present-day vegetation biomass in China ranged from $14.65 \pm 0.18$ to $30.88 \pm 0.25$ Pg C, with a multi-model ensemble mean (MMEM) of $21.85 \pm 0.11$ Pg C (table 3). The simulated biomass from CLM4.5 ($26.96 \pm 0.63$ Pg C), JSBACH ($30.88 \pm 0.25$ Pg C) and LPJ-GS ($25.90 \pm 0.28$ Pg C), and the MMEM ($21.85 \pm 0.11$ Pg C) were up to more than twice the latest estimates based on extensive literature review ($14.60 \pm 3.24$ Pg C) (Xu et al 2018) and on an observation campaign ($13.09$ Pg C) (Fang et al 2018) (table 3).

Before bias correction, all the DGVMs showed a systematic bias in simulated vegetation biomass for tree, and grass PFTs compared to observations (figures S1–2). Specifically, CLM4.5, JSBACH, and JULES overestimated the vegetation biomass for evergreen tree PFTs, while ORCHIDEE underestimated it. LPJ-GS and ORCHIDEE overestimated the vegetation biomass for deciduous tree PFTs (figure S1). All models tended to underestimate grass vegetation biomass (figure S2). Moreover, the spatial heterogeneity of vegetation biomass was considerably underestimated by JULES and LPJ-GS for tree PFTs (figure S1) and by all models except for CLM4.5 for grass PFTs (figure S2).

The bias-corrected results showed a significant improvement compared to the default simulations (figure 2). The correlation coefficients between corrected simulations and observations were much higher than uncorrected ones, and the RMSE was also lower in the bias-corrected simulations. All correlation coefficients for bias-corrected simulations passed the Student’s t-test at the 0.001 significance level, indicating that the bias-corrected simulations can reproduce China’s vegetation biomass reasonably well. In addition, the bias-corrected simulations provided an estimate of China’s vegetation biomass averaged for the period 2001–2010 ranging from $10.57 \pm 0.05$ to $15.47 \pm 0.13$ Pg C with a MMEM of $12.48 \pm 0.06$ Pg C (table 3), close to the latest estimates of $13.09$ Pg C from Fang et al (2018) and $14.6 \pm 3.24$ Pg C in Xu et al (2018). The uncertainty (inter-model spread) in simulated vegetation biomass was reduced through bias correction: standard deviation of the multi-model simulations reduced from 6.89 to 1.70 Pg C.

Spatially, all the bias-corrected DGVM simulations and the MMEM consistently showed that the vegetation biomass increases from northwest to southeast, in agreement with field observations and Carvalhais et al (2014) (figure 3). Grass or shrub was simulated as the dominant vegetation with low biomass in northwestern China because of cold or dry conditions, while warm and wet conditions in southeastern China resulted in more forest coverage and high biomass, in accordance with Köppen-Geiger classifications. When compared with uncorrected simulations (figure S3), bias correction decreased vegetation biomass estimates mostly in southern China and Northeast China.

3.2. Historical changes over the 20th century

All the six DGVMs exhibited an overall decline in vegetation biomass over the 20th century for China
Figure 2. Comparison of correlation coefficient (R) and root mean square error (RMSE) between 1950–2000 and/or 2000–2013 averaged PFT vegetation biomass (kg C (m² PFT area)⁻¹) from DGVM simulations with/without correction and field observations for evergreen and deciduous forests and grass in China. Simulations with higher R and lower RMSE is more skillful. R passed the Student’s t-test at the 0.001 (*) or 0.1(†) significance level.

Figure 3. Spatial patterns of bias-corrected vegetation biomass (kg C m⁻²) estimated by (a) MMEM of six DGVMs, (b)–(g) six DGVMs, (h) vegetation biomass observations over forest sites from Luo (1996) and literature review, and (i) estimates from Carvalhais et al (2014).

(significant at the 0.05 level using the Mann-Kendall test). The trends ranged from −32.48 Tg C yr⁻¹ (LPJ-GS) to −11.10 Tg C yr⁻¹ (JULES) (figure 4(a)). Only the ORCHIDEE model showed a different trajectory: biomass declined initially before the 1960s and increased thereafter, with a trend of −17.51 Tg C yr⁻¹
for the 20th century. The different change of ORCHIDEE after the 1960s was caused by the stronger response to rising [CO$_2$] (figure 4(b)) partly due to no modeling of nitrogen limitation (table 1) and the weaker response to LULCC since the 1980s (figure 4(c)) likely due to weaker decline in forest area (figure 6(a)) than other models. The MMEM of all DGVMs also exhibited a significant downward trend (–17.74 Tg C yr$^{-1}$).

All the DGVMs and MMEM agreed that rising [CO$_2$] between 1901 and 2000 tended to increase vegetation biomass. The trend ranged from +9.27 Tg C yr$^{-1}$ to +13.37 Tg C yr$^{-1}$, with a trend of around +10.04 Tg C yr$^{-1}$ (figure 4(b)). Rising [CO$_2$] increased vegetation biomass over most regions of China (figure S4) probably because the models were constructed to respond to changes in [CO$_2$] by changing light absorption and light-use efficiency (Norby et al 2005, Mao et al 2009) and also changing forest water-use efficiency (Keenan et al 2013).

The model simulations for all the DGVMs and MMEM showed that LULCC decreased vegetation biomass over the 20th century (figure 4(c)). The effect of LULCC ranged from a strong decrease of −50.71 Tg C yr$^{-1}$ for LPJ-GS to a moderate decrease of −24.28 Tg C yr$^{-1}$ for JULES with an estimate of −39.47 Tg C yr$^{-1}$ for the MMEM. LULCC impacts were the dominant driver for long-term changes in vegetation biomass over China for MMEM and all models (figure 4). Spatially, the largest impacts of LULCC on vegetation biomass were in the northeastern and southwestern forested regions (figure S5).

The impacts of climate change on the long-term trends of vegetation biomass were generally limited (figure 4(d)), with trends ranging from −3.75 Tg C yr$^{-1}$ (CLM4.5) to +5.06 Tg C yr$^{-1}$ (JSBACH). There was no significant trend (at the 0.05 level) in the MMEM (−0.02 Tg C yr$^{-1}$). This was probably because the impacts of climate change over the 20th century are generally weak and have large spatial heterogeneity (figure S6).

The DGVM simulations without bias-correction exhibited similar trends in vegetation biomass and similar responses to the various drivers, but the downward trend and responses were stronger than those in the bias-corrected simulations (Figs. S7–11), partly because the magnitude of vegetation biomass was overestimated in the uncorrected simulations (table 1, figure S3). For example, the decline in the vegetation biomass over the 20th century in the uncorrected simulations ranged from −71.70 to −14.36 Tg C yr$^{-1}$ (−31.81 Tg C yr$^{-1}$ for MMEM) (figure S7), compared to trends of −32.48 to −11.10 Tg C yr$^{-1}$ for bias-corrected simulations and −17.74 Tg C yr$^{-1}$ for the MMEM (figure 4(a)).

The bias-corrected MMEM showed a significant decline in vegetation biomass over most of China for the 20th century, especially in the southwest (figure 5(a)). Only limited areas in northern China showed a slight increase in vegetation biomass. LULCC was the main reason for changes in vegetation biomass over 72.42% of China’s land area (figure 5(b)), and largely driven by a decline in forest area. From 1901
to 2000, forest area in the DGVMs declined from $-1.38\%$ in JULES to $-9.43\%$ in ORCHIDEE (figure 5(a)). The areas dominated by the impact of rising [CO$_2$] were small, while climate change was the dominant driver for the increase in vegetation biomass only in part of northwestern China (figure 5(b)). Again, these conclusions were still robust without bias-correction, where LULCC was dominant over approximately 74.94\% of China (figure S11).

4. Discussion and conclusions

Because long-term vegetation biomass observations over large regions are unavailable, DGVMs are a good choice for quantifying the trends in regional and global vegetation biomass and isolating the impacts of each driver on the scale of a century (Piao et al 2011). Before trend and attribution analyses, evaluation should be done and statistical or dynamical
post-processing are often needed for the regional use of DGVM simulations because many parameters are global constants and some schemes are not optimized to fit the region of focus (Oleson et al. 2013). Our evaluation results showed that the DGVMs often overestimated the amount of vegetation biomass in China and underestimated the spatial heterogeneity of vegetation biomass for tree and grass PFTs. The overestimation in China’s vegetation biomass is partly due to underestimated turnover rate and overestimated photosynthesis for the DGVMs (table S3, supplement C) as well as overestimated forest area by 12.2% for LPJ-GS. The underestimation in spatial heterogeneity of vegetation biomass for a PFT is likely because the DGVMs generally do not consider the diversity in plant traits and individual growth within a PFT (Scheiter et al. 2013).

To reduce the bias in DGVM simulations, machine learning GPR was used as bias-correction method. The bias-corrected DGVM simulations were much more skillful, and could simulate China’s vegetation biomass reasonably well and reduced the uncertainty in DGVM simulations (smaller inter-model spread). This shows how statistical bias-correction can be used as the post-processing of model outputs to improve simulation skill and increase the reliability of subsequent quantification of temporal and spatial changes in terrestrial ecosystems. This method of bias-correction could also be applied to simulations of future changes in the vegetation biomass of China. Providing there are sufficient observations, it would be possible to use GPR bias-correction in other regions or for other simulated variables.

Our results also showed that China’s vegetation biomass decreased over the 20th century with significant trend of $-17.74$ ($-32.48$ to $-11.10$) Tg C yr$^{-1}$. This is consistent with the results from Wang et al. (2007) based on the InTEC model, but disagrees with the simulated increase of 0.02 Pg C yr$^{-1}$ by Mao et al. (2009) based on a version of SDGVM. The discrepancy is primarily because Mao et al. (2009) did not account for the impact of LULCC, which we have shown is the main driver of the decrease in vegetation biomass over the century.

LULCC datasets showed that China’s forest area had declined overall during the 20th century (Houghton and Hackler 2003, Hurtt et al. 2011, Lawrence et al. 2016). Fang et al. (2018) estimated that vegetation carbon density for forests is > 5 times of other functional types in China based on field observation campaigns. Together, they supported our conclusions that vegetation biomass had decreased over the 20th century and LULCC was the main driver for this decrease.

Sensitivity experiments in which single factors were held constant, as here, are a commonly-used approach to isolate the impact of a forcing (e.g. Gedney et al. 2006, Andela et al. 2017, Erb et al. 2018, Li et al. 2019) but do not allow potential interactions between different drivers to be identified. Four additional sensitivity experiments, including no change in all drivers, no CO$_2$ and land cover change, no climate and CO$_2$ change, no CO$_2$ and climate change, would facilitate a more comprehensive attribution analysis in the next phase of FireMIP (i.e. total $n^2-1$ simulations, $n$ is the number of drivers).

However, our bias correction still has limitations. First, the machine learning GPR is a statistical method and has the common problem that assumes the statistical relationship between dependent variables and independent variables (multi-year average of DGVM simulations and field observations in this study) to be constant. Second, although the field data that we collected and used had been quality controlled, they still included uncertainty and possibly errors due to sampling design, measurement techniques, and
observer bias. Such uncertainty and errors were not taken into account in our bias-correction, and may affect our estimation of trends in China’s vegetation biomass. In addition, the field data used in this study represented multi-year averages and did not provide information on temporal change, so the coefficients of bias-correction models were based on multi-year averages. That is, this study just corrected the multi-year amounts and spatial pattern, which would affect our trend estimates. Long-term field measures which could provide information on temporal change are needed to improve bias-correction further and our quantitative understanding of large-scale and long-term changes in vegetation carbon pool and their drivers.

Our estimated decline in vegetation biomass after ~1980 is likely incorrect. It is not consistent with earlier studies based on Chinese forest inventory data (Fang et al 2001, 2007) which reported an increase in vegetation biomass from 1981 to 2000. This opposite trend is mainly because the opposite change in China’s forest area used by these studies. The Chinese forest inventory data, also reported in Food and Agriculture Organization of the United Nations (FAO) national-level forest resource assessment, shows a sharp rise in China’s forest area due to reforestation policy since ~1980 (figure 6(b)) (Li and Li 1996, Ma et al 1997, State Administration of Forestry and Grassland 2017). However, in this study, the change in China’s forest area in the DGVMs, except for ORCHIDEE whose forest area change was constrained by Houghton et al. (2012), was characterized by a decrease after 1980 (figure 6(a)). The FireMIP models used the cropland and/or pasture area changes from Hurtt et al. (2011), and prescribed or modeled the changes in tree, grass, and or shrub PFTs areas which shared the remaining grid area after assigning areas for water/ice/urban (constant) and cropland/pasture (prescribed), an approach widely used in international MIPs, e.g. TRENDY (Sitch et al 2015), MsTMIP (Ito et al 2016), CMIP5 (Taylor et al 2012). The Hurtt et al. (2011) land use harmonization (LUH) model used inputs from HYDE 3.1 which showed an evident increase in China’s cropland area in recent decades that was incorrect (Liu and Tian 2010, Zhao and Zhang 2013). With the incorrect cropland area change, China’s forest area is decreased in the LULCC of Hurtt et al. (2011) and thus models in a variety of MIPs. A sharp decrease for the 1980 s and overall constant for the 1990 s in China’s forest area are also shown in LUH2_v2_h (figure 6(b)) (Lawrence et al 2016), that is the LULCC input data for many ongoing MIPs, e.g., CMIP6 as a basis for the next Intergovernmental Panel on Climate Change Assessment Report (IPCC AR6) (Eyring et al 2016). Given the importance of forest area change in simulated vegetation biomass change verified in the present study, the use of incorrect LULCC data will result in inaccurate simulation of the carbon sink and source in China for the last two decades of the 20th century in these MIPs. Therefore, the Chinese forest inventory data, and satellite-based land cover change data from Song et al. (2018) are recommended to LUH group for improving the accuracy of LULCC and thus simulations of various MIPs and IPCC reports.

Data availability

All DGVM simulations (both uncorrected and corrected) and field observations used in this study are publicly available at http://doi.org/10.5281/zenodo.3678535 (doi: 10.5281/zenodo.3678535) and Supplement B, respectively.

Acknowledgments

This study is co-supported by the National Key Research and Development Program of China (2017YFA0604804 and 2017YFA0604302) and National Natural Science Foundation of China (41475099 and 41630530). We thank Sam Levis for revising language errors in the manuscript, and Joe Melton, Ruixu Lian, and Zhongda Lin for helpful discussions. We are also grateful to two anonymous reviewers for their valuable comments and suggestions and ERL editors for handing this paper. SPH acknowledges the support from the ERC-funded project GC2.0 (694481) and from the High-end Foreign Expert Programme of China (GDW2019100161). And acknowledge support from the Helmholtz Foundation and its ATMO programme and Impulse and Networking fund. GL is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) (338130981). Jian Ni is funded by National Natural Science Foundation of China (31870462).

ORCID iDs

Fang Li https://orcid.org/0000-0002-3686-2257
Matthew Forrest https://orcid.org/0000-0003-1858-3489

References

Ahlström A, Shurgers G, Arneth A and Smith B 2012 Robustness and uncertainty in terrestrial ecosystem carbon response to CMIP5 climate change projections Environ. Res. Lett. 7 044008–10
Andela N, Morton D C, Giglio L, Chen Y and Randerson J T 2017 A human-driven decline in global burned area Science 356 1356
Best M J et al 2011 The Joint UK Land Environment Simulator (JULES), model description-Part 1: energy and water fluxes Geosci. Model Dev. 4 677–99
Brovkin V et al 2013 Effect of anthropogenic land-use and land-cover changes on climate and land carbon storage in CMIP5 projections for the twenty-first century J. Climate 26 6859–81
Carvalhais N et al 2014 Global covariation of carbon turnover times with climate in terrestrial ecosystems Nature 514 213–7
Chen Y Y et al 2016 Evaluating the performance of land surface model ORCHIDEE-CAN v1.0 on water and energy flux estimation with a single-and multi-layer energy budget scheme Geosci. Model. Dev. 9 2951–72
Clark D B et al 2011 The Joint UK Land Environment Simulator (JULES), model description—part 2: carbon fluxes and vegetation dynamics Geosci. Model. Dev. 4 701–22
Erb M P, Jackson C S, Broccoli A J, Lea D W, Valdes P J, Crucifix M and Dineizio P N 2018 Model evidence for a seasonal bias in Antarctic ice cores Nat. Commun. 9 1361
Eyring V, Bony S, Meehl G A, Senior C A, Stevens B, Stouffer R J and Taylor K E 2016 Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization Geosci. Model. Dev. 9 1937–58
Fang J Y, Chen A P, Peng C H, Zhao S Q and Ci L J 2001 Changes in forest biomass carbon storage in China between 1949 and 1998 Science 292 2320–2
Fang J Y, Guo Z D, Piao S L and Chen A P 2007 Terrestrial vegetation carbon sinks in China, 1981–2000 Sci. China D 50 1341–50
Fang J Y, Yu G R, Liu L L, Hu S J and Chapin III F S 2018 Climate change, human impacts, and carbon sequestration in China Proc. Natl. Acad. Sci. USA 115 4015–20
Foley J A, Prentice I C, Ramankutty N, Levis S, Pollard D, Sitch S and Haxeltine A 1996 An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics Global Biogeochem Cycles 10 603–28
Godney N, Cox P M, Betts R A, Boucher O, Huntingford C and Stott P A 2006 Detection of a direct carbon dioxide effect in continental river runoff records J. Geophys. Res. 111 2015–22
Hantson S et al 2016 The status and challenge of global fire modelling Biogeosciences 13 3359–75
Hantson S et al 2019 Quantitative assessment of fire and vegetation properties in historical simulations with fire-enabled vegetation models from the FireMIP intercomparison project Geosci. Model. Dev. Submitted
Houghton R A and Hackler J L 2003 Sources and sinks of carbon from land-use change in China Global Biogeochem. Cy. 17 1034
Houghton R A, House J I, Pongratz J, Van der Werf G R, Defries R S, Hansen M C, Le Quéré C and Ramankutty N 2012 Carbon emissions from land use and land-cover change Biogeosciences 9 5125–42
Hurtt G C et al 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 Clim. Change 109 117–61
Ichii K, Kondo M, Okabe Y, Ueyama M, Kobayashi H, Lee S-J, Saigusa N, Zhu Z and Myneni B R 2013 Recent changes in terrestrial gross primary productivity in Asia from 1982 to 2011 Remote Sens. 5 6043–62
Ito A et al 2016 Decadal trends in the seasonal-cycle amplitude of terrestrial CO2 exchange resulting from the ensemble of terrestrial biosphere models TELUS STR (Beijing: China Forestry Publishing House) pp 13–14 (in Chinese)
Liu M L and Tian H Q 2010 China’s land cover and land use change from 1700 to 2005: estimations from high-resolution satellite data and historical archives Global Biogeochem. Cy. 24 GB3003
Luo T X 1996 Patterns of net primary productivity for Chinese major forest types and their mathematical models PhD Dissertation Chinese Academy of Sciences p 230 (in Chinese)
Ma Z L, Song C S and Zhang Q H 1997 The Changes of Forest Distribution in China (Beijing: China Forestry Publishing House) (in Chinese)
Mangeon S, Voulgarakis A, Gillham R, Harper A, Sitch S and Folberth G 2016 INFERNO: a fire and emissions scheme for the UKMet Office’s Unified Model Geosci. Model. Dev. 9 2685–700
Mao J F, Wang B and Dai Y J 2009 Sensitivity of the carbon storage of potential vegetation to historical climate variability and CO2 in continental China Adv. Atmos. Sci. 26 87–100
Nj J 2001 Carbon storage in terrestrial ecosystems of China: estimates at different spatial resolutions and their responses to climate change Clim. Change 49 339–58
Norby R J et al 2005 Forest response to elevated CO2 is conserved across a broad range of productivity PNAS 102 18052–6
Oleson K W et al 2013 Technical Description of Version 4.5 Of the Community Land Model (CLM) (NCAR/TN-503+STR) (USA: NCAR Boulder)
Knorr W, Jiang L and Arneth A 2016 Climate, CO2, and human population impacts on global wildfire emissions Biogeosciences 13 267–82
Kondo M, Ichii K, Ueyama M, Mizoguchi Y, Hirata R and Saigusa N 2013 The role of carbon flux and biomeetric observations in constraining a terrestrial ecosystem model: a case study in disturbed forests in East Asia Ecol. Res. 28 893–905
Krinner G, Viovy N, de Noblet-ducoudr´e N, Og´ee J, Polcher J, Friedlingstein P, Gash J R, Sitch S and Prentice I C 2005 A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system Global Biogeochem. Cy. 19 GB1015
Lasslop G, Thonicke K and Kloster S 2014 SPITFIRE within the MPI Earth system model: model development and evaluation J. Adv. Model. Earth Syst. 6 745–55
Lawrence D W et al 2016 The Land Use Model Intercomparison Project (LUMIP) contribution to CMIP6: rationale and experimental design Geosci. Model. Dev. 9 2973–98
Le Quéré C et al 2014 Global carbon budget 2013 Earth Syst. Sci. Data 6 235–63
Lehsten V, Tansey K, Balszter H, Thonicke K, Spessa A, Weber U, Smith B and Arneth A 2009 Estimating carbon emissions from African wildfires Biogeosciences 6 349–60
Li F, Levis S and Ward D S 2013 Quantifying the role of fire in the Earth System – Part 1: Improved global fire modeling in the Community Earth System Model (CESM1) Biogeosciences 10 2293–314
Le Quéré C et al 2014 Global carbon budget 2013 Earth Syst. Sci. Data 6 235–63
Li F, Zeng X D and Levis S 2012 A process-based fire parameterization of intermediate complexity in a Dynamic Global Vegetation Model Biogeosciences 9 2761–80
Li F and Lawrence D M 2017 Role of fire in the global land water budget during the 20th century through changing ecosystems J. Clim. 30 1893–908
Li F et al 2019 Historical (1700–2012) Global multi-model estimates of the fire emissions from the fire Modeling Intercomparison Project (FireMIP) Atmos. Chem. Phys. 19 12345–67
Li K R, Wang S Q and Cao M K 2004 Vegetation and soil carbon storage in China Sci. China Earth Sci. 47 49–57
Li N, Xie G D, Zhang C S, Xiao Y, Zhang B, Chen W H, Sun Y Z and Wang S 2015 Biomass resources distribution in the terrestrial ecosystem of China Sustainability 7 8548–64
Li W H and Li F 1996 Research on Forest Resources in China (Beijing: China Forestry Publish House) pp 13–14 (in Chinese)
Pan Y D, Luo T X, Birdsey R, Hom J and Melillo J 2004 New estimations of carbon storage and sequestration in China's forests: effects of age-class and method on inventory-based carbon estimation Clim. Change 67 211–36
Piao S L et al 2011 Contribution of climate change and rising CO2 to terrestrial carbon balance in East Asia: A multi-model analysis Global Planet. Change 75 133–42
Rabin S S et al 2017 The Fire Modeling Intercomparison Project (FireMIP), phase 1: experimental and analytical protocols with detailed model descriptions Geosci. Model Dev. 10 1175–97
Reick C H, Raddatz T, Brovkin V and Gayler V 2013 Representation of natural and anthropogenic land cover change in MPI-ESM J. Adv. Model. Earth Syst. 5 459–82
Scheiter S, Langan L and Higgins S I 2013 Next-generation dynamic global vegetation models: learning from community ecology New Phytol. 198 937–69
Sitch S et al 2003 Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model Glob. Change Biol. 9 161–85
Sitch S et al 2015 Recent trends and drivers of regional sources and sinks of carbon dioxide Biogeosciences 12 653–79
Smith B, Prentice I C and Sykes M T 2001 Representation of vegetation dynamics in the modelling of terrestrial ecosystems: comparing two contrasting approaches within European climate space Global Ecol. Biogeogr. 10 621–37
Smith B, Wårlind D, Arneth A, Hickler T, Leadley P, Siltberg J and Zachе S 2014 Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model Biogeosciences 11 2027–54
Song X P, Hansen M C, Stehman S V, Potapov P V, Tyukavina A, Vermote E F and Townshend J R 2018 Global land change from 1982 to 2016 Nature 560 639–43
State Administration of Forestry and Grassland 2017 China Forestry Statistical Yearbook 2017 (Beijing: China Forestry Press) pp 396
Tang X L et al 2018 Carbon pools in China's terrestrial ecosystems: new estimates based on an intensive field survey PNAS 115 4021–6
Taylor K E, Stouffer R J and Meehl G A 2012 An overview of CMIP5 and the experiment design Bull. Amer. Meteor. Soc. 93 485–98
Wang S, Chen J M, Ju W M, Feng X, Chen M, Chen P and Yu G 2007 Carbon sinks and sources in China's forests during 1901–2001 J. Environ. Manage. 85 524–37
Wilks D S 1995 Statistical Methods in the Atmosphere Sciences: An Introduction (New York: Academic) pp 467
Xin Q C, Dai Y J, Li X, Liu X P, Gong P and Richardson A D 2018 A steady-state approximation approach to simulate seasonal leaf dynamics of deciduous broadleaf forests via climate variables Agr. Forest Meteorol. 239 44–56
Xu L, Yu G R, He N P, Wang Q F, Gao Y, Wen D, Li S G, Niu S L and Ge J P 2018 Carbon storage in China's terrestrial ecosystems: a synthesis Sci. Rep. 8 2806
Yue C et al 2014 Modelling the role of fires in the terrestrial carbon balance by incorporating SPITFIRE into the global vegetation model ORCHIDEE—Part 1: simulating historical global burned area and fire regimes Geosci. Model Dev. 7 2747–67
Yue C, Ciais P, Cadule P, Thonicke K and van Leeuwen T T 2015 Modelling the role of fires in the terrestrial carbon balance by incorporating SPITFIRE into the global vegetation model ORCHIDEE—part 2: carbon emissions and the role of fires in the global carbon balance Geosci. Model Dev. 8 1321–38
Zhao L M and Zhang S H 2013 The accuracy evaluation of two common global historical land use/cover datasets in China J. Northwest A&F Univ. (Natural Science Edition) 41 133–40