Sources of uncertainty in gross primary productivity simulated by light use efficiency models: Model structure, parameters, input data, and spatial resolution

Yi Zheng, Li Zhang, Jingfeng Xiao, Wenping Yuan, Min Yan, Tong Li, Zhiqiang Zhang

A Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China
b Hainan Key Laboratory of Earth Observation, Hainan 572029, China
c Earth Systems Research Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, NH 03824, USA
d School of Atmospheric Sciences, Sun Yat-Sen University, Guangzhou 519082, China
f Zhuhai Joint Innovative Center for Climate-Environment-Ecosystem, Zhuhai Key Laboratory of Dynamics Urban Climate and Ecology, Beijing Normal University, Beijing 100875, China
g College of Soil and Water Conservation, Beijing Forestry University, Beijing 100083, China

ABSTRACT

Accurate estimation of gross primary productivity (GPP) is essential for understanding ecosystem function and global carbon cycling. However, there is still substantial uncertainty in the magnitude, spatial distribution, and temporal dynamics of GPP. Using light use efficiency (LUE) models, we conducted a comprehensive analysis of the uncertainty in GPP estimation resulting from various sources: model structure, model parameters, input data, and spatial resolution. We first evaluated the influences of model structures, namely the fraction of absorbed photosynthetically active radiation (FPAR), water scalar (WS), and temperature scalar (TS), on site-level GPP estimates. We then used the Sobol’ sensitivity analysis to quantify the relative contributions of model input variables to the uncertainty in simulated GPP. In addition, we used different land cover and meteorological datasets to examine the effects of input data and spatial resolution on the magnitude and spatiotemporal patterns of GPP.

We found that the model structures affected not only model performance but also model parameters in a manner that differed with vegetation type and region. Thus, proper model structures and rigorous model parameterization and calibration should be adopted in GPP modeling. The Sobol’ sensitivity analysis showed that the meteorological drivers including photosynthetically active radiation (PAR) and daily minimum temperature (TMIN) had larger contribution to the uncertainty in simulated GPP than did the surface reflectance-based indices including enhanced vegetation index (EVI) and normalized difference water index (NDWI). At the regional scale, different land cover datasets had the largest impacts on GPP simulations, followed by the scale effects from different spatial resolutions; changing meteorological datasets had the smallest effects. Therefore, more accurate and finer-resolution land cover maps and meteorological datasets are essential for more accurate GPP estimates. Our findings have implications for improving our understanding of the full uncertainty in carbon flux estimates and reducing the uncertainty in carbon cycle simulations.

1. Introduction

Gross primary productivity (GPP) is the amount of carbon absorbed by plants through photosynthesis. As an important component of the carbon cycle, GPP is the largest carbon flux between the terrestrial biosphere and the atmosphere. GPP also drives ecosystem services such as food, fiber, and wood production (Beer et al., 2010). Therefore, accurately quantifying GPP at various spatial and temporal scales is essential for better understanding ecosystem function and global carbon cycling. However, there is substantial uncertainty associated with the estimation of GPP, particularly at regional to global scales.

Light use efficiency (LUE) models have been widely used to estimate...
GPP (Potter et al., 1993; Running et al., 2004; Xiao et al., 2004; Yuan et al., 2007). These models are based on the classical LUE logic (Monteith, 1972; Monteith and Moss, 1977):

\[
GPP = \text{PAR} \times \text{FPAR} \times \epsilon_{\text{max}} \times f_{\epsilon}
\]  

where PAR is the incident photosynthetically active radiation (MJ m\(^{-2}\)) per time period, FPAR is the fraction of PAR absorbed by vegetation canopies, \(\epsilon_{\text{max}}\) is the maximum LUE (g C MJ\(^{-1}\)) under the condition without environmental stresses, and \(f_{\epsilon}\) represents the environmental stresses (e.g., water scalar \((W_p)\) and temperature scalar \((T_p)\)) ranging from 0 to 1. LUE models have simple model structures and require only a small number of input variables. Despite their simplicity, the LUE models can generally capture the spatial and temporal dynamics of GPP fairly well (Yuan et al., 2007; Zhao and Running, 2010). However, similar to process-based ecosystem models (Cramer et al., 1999; Xiao et al., 2009; Thorn et al., 2015; Ma et al., 2017), LUE models can also lead to significant uncertainty in regional or global carbon flux estimates (Gebremichael and Barros, 2006; Verma et al., 2014; Yuan et al., 2014).

Model simulations have several sources of uncertainty, including model structures, model parameters, and input datasets (Beck, 1987; Xiao et al., 2014). Although LUE models are all based on the LUE logic, they have been developed using different model structures. Different representations have been used for FPAR, \(W_p\), and \(T_p\). Model structure has been considered as the most important factor that affects parameter values (e.g., \(\epsilon_{\text{max}}\)) and model performance (Yuan et al., 2014). Uncertainties in model parameters, particularly \(\epsilon_{\text{max}}\), significantly influence the accuracy of simulated GPP (Wagle et al., 2016). The spatial datasets that affect the uncertainty in GPP estimation mainly include land cover maps and meteorological data. Land cover maps adopted in carbon cycle modeling are usually derived from satellite data, and substantial uncertainties exist due to their data sources, classification schemes, and classifiers (Giri et al., 2005). For a given site or grid cell, the land cover type directly determines the value of parameters, particularly \(\epsilon_{\text{max}}\) (Wang et al., 2010). Besides land cover maps, meteorological data are also critical drivers for the estimation of carbon fluxes. The uncertainty of the meteorological products may be propagated to modeling results (Gebremichael and Barros, 2006; Heinsch et al., 2006).

Many studies have examined the effects of individual source of uncertainty (e.g., model structures, parameters, or input data) on GPP modeling using LUE models. For example, Zhang et al. (2015a) evaluated the model performance of four LUE models using 51 eddy covariance flux towers and identified possible further improvements through structure optimization. Xiao et al. (2014) quantified the uncertainty of model parameters and assessed its effects on the estimation of regional carbon fluxes. A few studies also examined the influences of different meteorological data and land cover representation on GPP simulations (Zhao et al., 2006; Xiao et al., 2011; Cai et al., 2014). In addition, spatial heterogeneity and the resolution of input data (e.g., land cover map, meteorological data) may also lead to uncertainty in simulated carbon fluxes (Liu, 2014; Zhao and Liu, 2014). However, to our knowledge, no study has systematically evaluated the influences of model structure, model parameters, input data, and spatial resolution on the uncertainty in carbon fluxes.

In this paper, we assessed the uncertainty in simulated GPP resulting from the three main sources of uncertainty: model structures, model parameters, and model inputs. We first analyzed the influence of model structures on model parameters and model uncertainty at site level using different representations of FPAR, \(W_p\), and \(T_p\). We then quantified the relative contributions of model input variables to GPP uncertainty. Finally, we investigated the effects of different model input datasets (land cover maps and meteorological datasets) and spatial resolutions on the magnitude and spatiotemporal patterns of GPP at the regional scale. This study will improve our understanding of the uncertainty in GPP modeling and will potentially lead to more accurate carbon flux estimates.

2. Materials and methods

2.1. Study area

This study was carried out around the agro-pastoral ecotone of northern China (39.0° N to 46.8° N, 110.5° E to 122.8° E) (Fig. 1). From the northwest to the southeast, the annual mean temperature and precipitation range from −5°C to 10°C and 35 mm to 600 mm, respectively (calculated from the meteorological datasets in Section 2.2.3). The large temperature and precipitation gradients make the region a natural ecotone from arid and semi-arid to humid land, leading to grassland–cropland–forest mixed landscapes. The two dominant vegetation types of the ecotone are steppes and croplands, while forests (mainly deciduous broadleaf forest (DBF)) only account for a relatively small fraction of the region. The highly heterogeneous landscape accompanied with densely instrumented eddy covariance flux towers (Fig. 1) make the region a unique test bed for uncertainty analysis of ecosystem models (e.g., LUE models).

2.2. Data

2.2.1. Eddy covariance and meteorological data from flux towers

We used carbon flux and meteorological data from six eddy covariance flux sites across the study area (Fig. 1). Our study sites consist of four grassland sites – CN-Du2 (typical steppe), CN-Xi1 (typical fenced steppe), CN-X12 (degraded steppe), and Xi5 (short grass steppe); one cropland site – CN-Du1; and one forest site – CN-Bed. The data for the forest site (CN-Bed) was obtained from the US-China Carbon Consortium (USCCC), and the data for all other five sites were obtained from the LaThuile Synthesis Dataset (http://fluxnet.fluxdata.org). More detailed information on these sites can be found in the references in Table 1.

For each site, we used daily or hourly GPP, photosynthetically active radiation (PAR), air temperature (T), and vapor pressure deficit (VPD). To match the intervals (8-day) of the data products derived from the moderate resolution imaging spectroradiometer (MODIS), we aggregated the daily or hourly values of each variable to 8-day time step. We ignored any 8-day interval with more than 5 days of missing daily values to minimize the errors and uncertainties of the flux and meteorological data.

The eddy covariance flux towers directly measure the net ecosystem exchange (NEE) of carbon dioxide between ecosystems and the atmosphere. GPP was calculated as the difference between daytime ecosystem respiration (RE\(_d\)) and daytime NEE (NEE\(_d\)). RE\(_d\) was estimated using daytime temperature and the equation between nighttime temperature and nighttime NEE. The partitioning of NEE and the gap-filling of missing or bad data were based on the methods described in Reichstein et al. (2005).

2.2.2. Remote sensing data

In this study, we used surface reflectance, FPAR, GPP, and leaf area index (LAI) products over the period 2001–2012. The surface reflectance (MOD09A1, collection 006), FPAR (MOD15A2, collection 005), and GPP (MOD17A2, collection 005) products were derived from MODIS and were obtained from NASA’s Distribute Active Archive Center (DAAC) (https://ladsweb.nascom.nasa.gov/). The LAI product was the Global Land Surface Satellite Leaf Area Index (GLASS LAI) product provided by the Center for Global Change Data Processing and Analysis of the Beijing Normal University (http://glass-product.bnu.edu.cn/). Each product is available at 8-day interval. Because the footprint of eddy covariance flux towers is generally less than 1 km\(^2\) (Schmid, 2002), we extracted the average values of all grid cells within the 1 km × 1 km area surrounding each tower for each variable and for each time step. The linear interpolation technique was used to fill the missing values or to replace the unreliable values determined by the quality assurance flags.
The MODIS surface reflectance product (MOD09A1) provides surface spectral reflectance in seven bands from blue to shortwave infrared at 500 m spatial resolution. This product was used to calculate the normalized difference water index (NDWI) (Gao, 1996) and enhanced vegetation index (EVI) (Huete et al., 2002).

\[
\text{NDWI} = \frac{\rho_{\text{nir}} - \rho_{\text{swir}}}{\rho_{\text{nir}} + \rho_{\text{swir}}} 
\]

\[
\text{EVI} = G \times \frac{\rho_{\text{nir}} - \rho_{\text{red}}}{\rho_{\text{nir}} + (C_1 \times \rho_{\text{red}} - C_2 \times \rho_{\text{blue}}) + L}
\]

where \(\rho_{\text{nir}}\), \(\rho_{\text{swir}}\), \(\rho_{\text{red}}\), and \(\rho_{\text{blue}}\) are the reflectance of near infrared (841–875 nm), shortwave infrared (1628–1652 nm), red (620–670 nm), and blue (459–479 nm) bands, respectively; \(G = 2.5\), \(C_1 = 6\), \(C_2 = 7.5\), and \(L = 1\).

The MODIS FPAR product (MOD15A2) is at 1 km spatial resolution and was derived from surface reflectance data from MODIS on board Terra using a three-dimensional radiative transfer model. When the radiative transfer model failed due to bad geometry, cloud contamination, or snow and ice, a back-up method based on the empirical relationships between normalized difference vegetation index (NDVI) and FPAR was utilized (Myneni et al., 2002). The MODIS FPAR data is

---

Table 1

<table>
<thead>
<tr>
<th>Site name</th>
<th>Vegetation type</th>
<th>Location (°N, °E)</th>
<th>Site-year</th>
<th>MAT (°C)</th>
<th>MAP (mm)</th>
<th>Dominant species</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CN-Du2</td>
<td>Grasslands</td>
<td>42.05, 116.28</td>
<td>2006</td>
<td>3.3</td>
<td>399</td>
<td>Stipa krylovi, Artemisia frigida</td>
<td>Chen et al. (2009)</td>
</tr>
<tr>
<td>CN-Xi1</td>
<td>Grasslands</td>
<td>43.55, 116.68</td>
<td>2006</td>
<td>2.0</td>
<td>360</td>
<td>Stipa grandis, Leymus chinensis</td>
<td>Chen et al. (2009)</td>
</tr>
<tr>
<td>CN-Xi2</td>
<td>Grasslands</td>
<td>43.55, 116.67</td>
<td>2006</td>
<td>2.0</td>
<td>360</td>
<td>Stipa grandis, Artemisia frigida</td>
<td>Chen et al. (2009)</td>
</tr>
<tr>
<td>CN-Du1</td>
<td>Croplands</td>
<td>42.05, 116.67</td>
<td>2005-2006</td>
<td>3.3</td>
<td>399</td>
<td>Triticum aestivum</td>
<td>Chen et al. (2009)</td>
</tr>
</tbody>
</table>

---

Fig. 1. Study area and the distribution of eddy covariance flux sites. The land cover map is simplified from the GlobeLand30 (2010) land cover map (Chen et al., 2015). In the surrounding panels, the red symbols stand for flux sites overlaid on the Landsat imagery (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).
sometimes spatially and temporally incomplete, and the quality and accuracy of this product still require improvements (Liang et al., 2013). To further investigate the potential errors of simulated GPP introduced by FPAR, we also used the GLASS LAI product to calculate FPAR. The long-term GLASS LAI product was generated by using the General Regression Neural Networks (GRNNs) trained by the fusion of the MODIS LAI and CYCLOPES LAI products (Liang et al., 2013; Zhao et al., 2013). Compared with other LAI products, the GLASS LAI product reduced uncertainty and was also spatially complete and temporally continuous (Fang et al., 2013; Xiao et al., 2016a).

The MODIS GPP product (MOD17A2) is at 1 km spatial resolution and was produced based on a LUE model, the MODIS GPP algorithm (Running et al., 2004; Zhao et al., 2005). It was used to compare against our modeled GPP.

2.2.3. Meteorological reanalysis data

The China Meteorological Forcing Dataset (CMFD) and the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) were used to analyze the effects of different meteorological datasets on GPP simulation. The CMFD is a reanalysis dataset developed specifically for China with high accuracy (http://westdc.westgis.ac.cn/data/) (He and Yang, 2011). It was produced by merging observations from China Meteorological Administration’s weather stations, Princeton forcing data, GLDAS data, GEWEX-SRB downward shortwave radiation, and TRMM satellite precipitation analysis data (Yang et al., 2010; Chen et al., 2011). The meteorological data from CMFD have a spatial resolution of 0.1° and a temporal resolution of 3 h (Chen et al., 2011). In this study, we obtained temperature and downward shortwave radiation (SWRad) data from 2001 to 2012. Photosynthetically active radiation (PAR) was determined as a fraction of the SWRad:

\[
PAR = 0.45 \times SWRad
\]

The MERRA-2 is a reanalysis dataset developed by NASA’s Global Modeling and Assimilation Office (GMAO) using an upgraded version of the Goddard Earth Observing System Version 5 (GEOS-5) (https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/) (Rienecker et al., 2011). Compared with the first generation (MERRA), MERRA-2 has an enhanced assimilation system to reduce uncertainties (Molod et al., 2015). It enables the assimilation of modern hyperspectral radiance and microwave observations along with GPS-Radio Occultation datasets. The MERRA-2 product provides data from 1980 to present and has a spatial resolution of 0.625° in longitude by 0.5° in latitude. In this study, we obtained hourly air temperature and hourly PAR from 2001 to 2012. For the two meteorological datasets, we aggregated daily minimum temperature (TMIN) and PAR to 8-day interval to match the temporal resolution of the MODIS products.

We examined the accuracy of daily TMIN and PAR from the two meteorological datasets by comparing them with in-situ flux tower measurements at daily steps (Figs. 2–3). Both TMIN\text{CMFD} and TMIN\text{MERRA-2} were highly correlated with TMIN\text{NE} with correlation coefficient (r) over 0.96 (Fig. 2). The r of PAR varied with both dataset and vegetation type, and ranged from 0.75 to 0.93 (Fig. 3). At the forest site, PAR\text{CMFD} and PAR\text{MERRA-2} were slightly over-estimated, but with high correlation (r = 0.75 and r = 0.88, respectively) (Fig. 3(a3)–(b3)). Overall, the temperature and PAR of the two reanalysis datasets had relatively high accuracy and are suitable for being used as input datasets for GPP modeling.

2.2.4. Land cover data

To investigate the effects of different land cover maps on regional GPP modeling, we used three sources of land cover maps: GlobeLand30, MODIS land cover map (MCD12Q1), and Global Land Cover 2000 (GLC2000). The GlobeLand30 in 2010 at 30 m spatial resolution covers the land from 80°N to 80°S (http://glc30.tianditu.com/). It consists of ten land cover types: cultivated land, forests, grasslands, shrublands, wetlands, water bodies, tundra, artificial surfaces, bareland, and permanent snow (Chen et al., 2015). The 500 m resolution MCD12Q1 in 2006 with the International Geosphere Biosphere Programme (IGBP) classification scheme was obtained from NASA’s Earth Observing System Data and Information System (EOSDIS) (http://reverb.echo.nasa.gov). MCD12Q1 was derived from data from MODIS sensors on board the TERRA and AQUA satellites and was produced using an ensemble decision tree algorithm (Friedl et al., 2010). The 1 km GLC2000 was produced by the Joint Research Centre (JRC) of the European Commission using 1 km resolution, daily SPOT VEGETATION data (http://forobs.jrc.ec.europa.eu/products/gl2000/gl2000.php). It was based on the Food and Agriculture Organization (FAO) Land Cover Classification System (LCCS). These maps are based on different sources of satellite data and different classification methods and thereby exhibit significant discrepancies. These maps are also based on different classification schemes. For example, the MCD12Q1 includes evergreen forests, deciduous forests and mixed forests; the GLC2000 contains evergreen forests and deciduous forests; there are no forest subclasses in GlobeLand30. Given the inconsistency in the classification schemes and the land surface characteristic in the study area, we used four broad classes (grasslands, croplands, forests, and others) in our study.

We aggregated the 30 m GlobeLand30 land cover map to 500 m and 1 km resolution and the 500 m MCD12Q1 to 1 km by selecting the dominant type for each 500 m or 1 km pixel. Fig. 4 shows the spatial patterns and statistics of these land cover maps. For the same land cover product (e.g., GlobeLand30), there were no obvious differences in the percentage area of each vegetation type among the maps with different spatial resolutions (30 m, 500 m or 1 km) (Fig. 4(a)–(c), and (g)). However, there were considerable differences for each vegetation type among different land cover products at the same spatial resolution (e.g., 1 km), particularly for the central grasslands–croplands–forests mixed regions (Fig. 4(c), (e)–(g)). Specifically, grasslands was the dominant type, with the highest percentage in the 1 km MCD12Q1 (78.9%) and relatively lower percentage in the 1 km GlobeLand30 (58.4%) and 1 km GLC2000 (48.7%). The proportion of croplands in the 1 km GlobeLand30 and 1 km GLC2000 was higher than that in the 1 km MCD12Q1. The forests in the 1 km MCD12Q1 only accounted for 2.2% of the region, but accounted for 7.5% and 13.9% in the 1 km GlobeLand30 and 1 km GLC2000, respectively.

2.3. Methods

2.3.1. LUE model description and validation

LUE models are based on different representations for each of the three model structures: FPAR, Ws, and Ts. For example, FPAR is approximated by EVI in the VPM model (Xiao et al., 2004):

\[
\text{FPAR}_{\text{EV1}} = \text{EVI}
\]

In the EC-LUE model (Yuan et al., 2007), a linear function of NDVI is adopted to represent FPAR:

\[
\text{FPAR}_{\text{NDVI}} = 1.24 \times \text{NDVI} - 0.168
\]

FPAR in the MODIS GPP algorithm is based on the MODIS FPAR product (MOD15A2), referred to as FPAR\text{MODIS} hereafter. We also used GLASS LAI (LAI\text{GLASS}) to approximate FPAR for comparison purpose. Based on the simple Beer’s Law approach (Turner et al., 2006), the canopy FPAR can be calculated as function of LAI:

\[
\text{FPAR}_{\text{GLASS}} = 1 - e^{-K \times \text{LAI}_{\text{GLASS}}}
\]

where K is the light extinction coefficient, which was set to 0.5. Water availability is an important controlling factor of vegetation photosynthesis. Similar to FPAR, the Ws in the LUE models is based on different representations. For example, the VPM model uses a function of NDWI and NDWImax to assess the effects of water stress on photosynthesis.
where NDWImax is the maximum NDWI during the growing season. For multiple years, the longtime run averaged NDWI value for each 8-day interval was calculated for each pixel. We defined the growing season as the period when the EVI value was larger than the annual mean EVI. In some studies, Ws was modified to simulate the rapid response of vegetation photosynthesis to short-term water conditions (He et al., 2014):

$$W_{\text{NDWI}} = 0.5 + \text{NDWI}$$

In the MODIS GPP model, Ws is represented using daytime VPD (VPDd):

$$W_{\text{VPD}} = \frac{\text{VPD}_{\text{max}} - \text{VPD}_d}{\text{VPD}_{\text{max}} - \text{VPD}_{\text{min}}}$$

where VPDmax is the daylight average VPD when the actual LUE equals 0, and VPDmin is the daylight average VPD when the actual LUE equals εmax (for optimal daily TMIN). Both VPDmax and VPDmin were obtained from the Biome Property Look Up Table (BPLUT) of the MOD17 algorithm (Running and Zhao, 2015).

In many LUE models (e.g., GLO-PEM, EC-LUE, and VPM), the Ts is calculated using the equation in the Terrestrial Ecosystem Model (TEM) (Raich et al., 1991):

$$T_{\text{TEM}} = \frac{(T_{\text{min}} - T_{\text{opt}})(T_{\text{max}} - T_{\text{opt}})}{(T_{\text{min}} - T_{\text{opt}})(T_{\text{max}} - T_{\text{opt}})(T_{\text{max}} - T_{\text{min}})}$$

where Tmin, Topt, and Tmax are the minimum, optimum, maximum photosynthesis temperatures. Topt is the air temperature when the...
yearly averaged EVI reaches the peak for each pixel. $T_{\min}$ was set to 0, and $T_{\min}$ was calculated as $T_{\text{opt}} + (T_{\text{opt}} - T_{\min})$. The $T$s in the Carnegie-Ames-Stanford Approach (CASA) model (Potter et al., 1993) does not use $T_{\min}$ and $T_{\text{opt}}$ as input parameters (Eq. (12)). Therefore, we introduced $T$s in the CASA model to investigate whether a simplified number of parameters can reduce model uncertainty and improve accuracy:

$$T_{\text{CASA}} = \frac{1.1814 \times (0.8 + 0.02 \times T_{\text{opt}} - 0.0005 \times T_{\text{opt}}^2) \times (1 + \exp(0.3 \times (T_{\text{opt}} - 10)))}{1 + \exp(0.2 \times (T_{\text{opt}} - 10))}$$

where $T_{\text{CASA}}$ is a parameter between 0 and 1. If $T_{\text{CASA}}$ is less than 0, we set $T_{\text{CASA}}$ to 0; if $T_{\text{CASA}}$ is greater than 1, we set $T_{\text{CASA}}$ to 1. In the MODIS GPP model, the daily minimum temperature (TMIN) was used to represent the limitation of cold temperature on plant photosynthesis:

$$T_{\text{MODIS}} = \frac{T_{\text{MIN}} - T_{\text{MIN}_{\min}}}{T_{\text{MIN}_{\max}} - T_{\text{MIN}_{\min}}}$$

where $T_{\text{MIN}_{\min}}$ is the daily TMIN when the actual LUE equals 0, and $T_{\text{MIN}_{\max}}$ is the daily TMIN when the actual LUE equals $\epsilon_{\text{max}}$ (for optimal VPD).

As in Fig. 5, we compared the effects of the model structures described above on GPP modeling. For each structure, one representation was selected at a time, and a total of 36 combinations of model structures were compared. We used all the site-years of flux data for model calibration and then adopted the leave-one-out cross-validation method (Cawley and Talbot, 2003) to evaluate the performance of the optimized GPP model.

We optimized $\epsilon_{\text{max}}$ but not the parameters in environmental scalars for several reasons. First, $\epsilon_{\text{max}}$ determines the expected rate of photosynthesis assuming optimal conditions and is the most important parameter in light use efficiency (LUE) models. Second, $\epsilon_{\text{max}}$ is the parameter which may introduce the largest uncertainties compared with other parameters in temperature and water scalars. For example, using ten crop flux towers in northern China, Wang et al. (2013) found that the MODIS GPP product was underestimated by 70% compared to tower measured GPP mainly because of the low $\epsilon_{\text{max}}$ value in the MODIS GPP algorithm for croplands. Third, more importantly, $\epsilon_{\text{max}}$ is a common parameter for all the LUE models, while other parameters only exist in one or two LUE models. We optimized $\epsilon_{\text{max}}$ in our study using the least-squares method.

The Pearson’s correlation coefficient ($r$) and the root mean square error (RMSE) were used to evaluate the performance of the GPP models. Standard deviation (SD) and relative uncertainty (RU) were adopted to evaluate the uncertainty in the GPP estimation. RU can be expressed as follow:

$$RU = \frac{2 \times SD}{\text{mean}} \times 100\%$$

2.3.2. Sensitivity analysis

Sensitivity analysis can quantify the degree of uncertainty in model predictions determined by uncertainties in model inputs (Saltelli et al., 2004). It can be used to corroborate the model structures, identify the critical model parameters and inputs, determine minimum data standards, and simplify models (Lilburne and Tarantola, 2009). Sensitivity analysis includes local sensitivity analysis and global sensitivity analysis. The Sobol’ sensitivity analysis is a widely used global sensitivity analysis method. It can not only obtain first order effects but also identify the interaction effects of parameters. To determine the most influential input variables, the Sobol’ sensitivity analysis was used to evaluate the effects of model input uncertainties (i.e., PAR, TMIN, EVI, and NDWI) on GPP simulation.

The Sobol’ method is a variance-based technique. The total variance of the model output ($V$) can be decomposed into the effects of each individual model input variable and their interactions.

$$V = \sum_i V_i + \sum_{ij} V_{ij} + \sum_{ijk} V_{ijk} + \cdots + V_{12...n}$$

where $V_i$ is the output variance explained by the $i$th model input; $V_{ij}$ is the output variance explained by the interaction of the $i$th and the $j$th input; and $n$ is the input variable number. Accordingly, the first order index is $S_i = V_i/V$, and the higher order indices are $S_{ij} = V_{ij}/V - S_i - S_j - V_{12...n} = V_{12...n}/V$. The total order index is the sum of the first order and the higher order indices (all...
interactions between the variable and other variables).

The Sobol’ sensitivity analysis contains the following three procedures. First, it defines the ranges and probability distribution functions (PDFs) of the model input variables. According to the literature (Li et al., 2012) and statistics from tower measured data and spatial data, we assumed that the model inputs (PAR, TMIN, EVI, and NDWI) were uniformly distributed and determined the variable ranges for each land cover type, as in Table 2. Second, this approach generates a series of samples. By using the defined ranges and PDFs, 12,288 random samples were generated for each variable. Third, it estimates the simulated GPP based on the samples, and outputs the results of the sensitivity analysis.

2.3.3. Regional GPP estimation and uncertainty quantification

To identify and quantify the influences of different spatial resolutions, land cover maps, and meteorological datasets on GPP estimation, we applied the optimized GPP model to estimate GPP for each grid cell for the period from 2001 to 2012 using the different schemes in Table 3. The reference scheme (Sch0) used a relatively higher spatial resolution (500 m), a more accurate meteorological dataset (CMFD), and a land cover map with higher native spatial resolution (GlobeLand30). The altered schemes (Sch1-Sch9) were designed by modifying one, two or three input factors (e.g., spatial resolution, meteorology dataset, or land cover map).

The effects of spatial resolutions, meteorological datasets, and land cover maps on the uncertainty in regional GPP estimations were quantified by comparing the altered schemes with the reference scheme. We first calculated the GPP difference between the altered and reference schemes. Furthermore, we evaluated the effects on the magnitude, spatial pattern, and temporal pattern of GPP following the method by Jung et al. (2007). The effect on magnitude was calculated in percentage as the absolute difference between the altered and reference schemes relative to the temporal mean value of the reference scheme for all pixels (Eq. (16)). The effect on spatial pattern was quantified as the percentage of the spatial correlation coefficient (Eq. (18)), and finally the effect on temporal pattern was calculated as the mean of all pixels.

\[
\text{EFFECT}_{\text{Magnitude}} = \frac{\sum_{i=1}^{n} (\text{Sch}_i - \text{Ref}_i)}{\sum_{i=1}^{n} \text{Ref}_i} \times 100 \% \tag{16}
\]

\[
\text{EFFECT}_{\text{Spatial}} = 1 - \left( \frac{\sum_{i=1}^{n} (\text{Sch}_i - \text{Ref}_i)^2}{\sum_{i=1}^{n} (\text{Sch}_i - \text{Sch}_i)^2 \times \sum_{i=1}^{n} (\text{Ref}_i - \text{Ref}_i)^2} \right)^{1/2} \tag{17}
\]

\[
\text{EFFECT}_{\text{Temporal}(i)} = 1 - \left( \frac{\sum_{y=2001}^{2012} (\text{Sch}_y - \text{Ref}_y)^2}{\sum_{y=2001}^{2012} (\text{Sch}_y - \text{Sch}_y)^2 \times \sum_{y=2001}^{2012} (\text{Ref}_y - \text{Ref}_y)^2} \right)^{1/2} \tag{18}
\]

where Ref and Sch represent the reference and altered schemes, respectively, i is the pixel index, n is the number of valid pixels, and y is year. The one and two overbars denote the pixel-based temporal mean and the spatial pixel mean of the temporal mean, respectively.

3. Results

3.1. Parameter estimation, model structure comparison and model validation

The values of model performance indices and \( \varepsilon_{\text{max}} \) are shown in Table 4. The performance of LUE models varied with model structure (FPAR, \( W_s \), and \( T_s \)) and vegetation type. FPARNDVI was generally better than FPARMODIS, FPAREVI, and FPARGLASS for grassland GPP estimation. For forests, FPARGLASS was better than FPARMODIS, FPARMODIS, and FPARNDVI. For crop lands, the FPARGLASS was better than FPARMODIS, FPAREVI, and FPARNDVI (Table 4).

As for the water scalar, the model performance obviously improved for grasslands and croplands but decreased slightly for forests by changing \( \text{W}_s_{\text{NDWI}} \) to \( \text{W}_s_{\text{VPD}} \). The \( \text{W}_s_{\text{NDWI}} \) was inferior to \( \text{W}_s_{\text{NDWI}} \) for the three vegetation types (Table 4). For the temperature scalar, there was no evident difference in model performance between \( \text{T}_s_{\text{TEM}} \) and \( \text{T}_s_{\text{CASA}} \). The \( \text{T}_s_{\text{MODIS}} \) improved the model performance for croplands, and was comparable to \( \text{T}_s_{\text{TEM}} \) and \( \text{T}_s_{\text{CASA}} \) for grasslands and forests (Table 4).

In addition, the values of \( \varepsilon_{\text{max}} \) changed substantially with model structure, particularly for \( \text{W}_s \) and FPAR. The \( \varepsilon_{\text{max}} \) value ranged from 0.70 to 1.98 g C MJ\(^{-1}\) for grasslands, 0.86 to 2.44 g C MJ\(^{-1}\) for croplands, and 2.59 to 5.39 g C MJ\(^{-1}\) for forests. It implies that the model structures largely affected the model parameters (e.g., \( \varepsilon_{\text{max}} \) (Table 4).

Based on both r and RMSE, the optimized GPP model using FPAREVI,
croplands (was fairly good for grasslands (grasslands) or each year (for croplands and forests) (Table 5). The predicted GPP matched the tower GPP well and generally captured validation and uncertainty analysis for the three vegetation types. Overall, validation of LUE models based on different data sources and spatial resolutions. The altered schemes (Sch1-Sch9) are modified based on the reference scheme (Sch0). The short lines indicate no modifications.

Table 3
Schemes designed for regional GPP estimation according to different data sources and spatial resolutions. The altered schemes (Sch1-Sch9) are modified based on the reference scheme (Sch0). The short lines indicate no modifications.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Modification</th>
<th>NO.</th>
<th>Spatial Resolution</th>
<th>Meteorological Data</th>
<th>Land Cover Map</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference Scheme</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Altered Schemes</td>
<td>Modified one factor</td>
<td>Sch0</td>
<td>500 m</td>
<td>CMFD</td>
<td>GlobeLand30</td>
</tr>
<tr>
<td></td>
<td>Sch1</td>
<td>1000 m</td>
<td>—</td>
<td>MERRA-2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sch2</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sch3</td>
<td>—</td>
<td>—</td>
<td>MCD12Q1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Modified two factors</td>
<td>Sch4</td>
<td>1000 m</td>
<td>MERRA-2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sch5</td>
<td>1000 m</td>
<td>—</td>
<td>MCD12Q1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sch6</td>
<td>1000 m</td>
<td>—</td>
<td>GLC2000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sch7</td>
<td>—</td>
<td>—</td>
<td>MCD12Q1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Modified three factors</td>
<td>Sch8</td>
<td>1000 m</td>
<td>MERRA-2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sch9</td>
<td>1000 m</td>
<td>—</td>
<td>MCD12Q1</td>
<td></td>
</tr>
</tbody>
</table>

Table 4
Validation of LUE models based on different structures for the three vegetation types.

<table>
<thead>
<tr>
<th>No.</th>
<th>Model Structure</th>
<th>Grasslands</th>
<th>Croplands</th>
<th>Forests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Indicators</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
|      | W5,NDVI2, and T5,MODIS (NO. 6 in Table 4) was selected for further validation and uncertainty analysis for the three vegetation types. Overall, the predicted GPP matched the tower GPP well and generally captured the seasonal variations, with r ranging from 0.89 to 0.92 and RMSE varying from 0.50 to 1.47 g C m⁻² d⁻¹ (Fig. 6). The model performance was fairly good for grasslands (r = 0.89; RMSE = 0.50 g C m⁻² d⁻¹), croplands (r = 0.91; RMSE = 0.89 g C m⁻² d⁻¹) and forests (r = 0.92; RMSE = 1.47 g C m⁻² d⁻¹), and the data points were all close to the 1:1 lines for the three vegetation types (Fig. 6).

We also used leave-one-out cross-validation to evaluate the performance of the optimized GPP model by withholding each site (for grasslands) or each year (for croplands and forests) (Table 5). The r varied between 0.80 and 0.96, and the RMSE ranged from 0.44 to 0.64 g C m⁻² d⁻¹ for grasslands. Croplands also exhibited high correlation between tower and estimated GPP, with r = 0.91 and RMSE between 0.89 and 0.90 g C m⁻² d⁻¹. Forests had the highest r (0.88–0.97) and RMSE (0.95–1.98 g C m⁻² d⁻¹). Overall, the cross-validation demonstrated that the performance of the optimized GPP models was robust and stable.

3.2. Sobol’ sensitivity analysis of model input variables

The Sobol’ sensitivity analysis was used to quantify the contributions of different model inputs to model uncertainty. Table 6 shows the

<table>
<thead>
<tr>
<th>No.</th>
<th>Model Structure</th>
<th>Grasslands</th>
<th>Croplands</th>
<th>Forests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Indicators</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
- For different combinations of model structures, the valid number of testing points were 136–156 for grasslands, 57–63 for croplands, and 68–76 for forests.
- Units: $E_{\text{max}}$ (g C MJ⁻¹), RMSE (g C m⁻² d⁻¹).

|      | Indicators       |            |           |         |
first order indices, second order indices, and the total order indices simulated by input variables using the Sobol’ sensitivity analysis. Generally, the GPP models were most sensitive to PAR with the highest first order index and total order index, followed by EVI and TMIN. The sensitivities of PAR, EVI and TMIN were higher with first order indices over 0.20, but the sensitivities of NDWI were relatively low with first order indices under 0.10 for the three vegetation types. In addition, all the total order indices were larger than the corresponding first order indices, indicating that the interaction effects between different model inputs on GPP simulation could not be neglected.

For the interactions, the second order indices of PAR, EVI and TMIN were similar among the three vegetation types with the values ranging from 0.10–0.14 (Table 6), and all the third order indices were less than 0.01. Furthermore, a detailed pairwise second order index between each two variables can better exhibit their interactions. Similar to the second order indices in Table 6, the pairwise second order indices were higher when the variables were more sensitive with higher first and total order indices. For example, the pairwise second order indices were higher for PAR&EVI, PAR&TMIN and EVI&TMIN with the values ranging from 0.04 to 0.07, but lower for NDWI&PAR, NDWI&EVI and NDWI&TMIN with the values ranging from 0 to 0.02 (data are not shown).

Table 5
Leave-one-out validation of the optimized GPP model (No. 6 in Table 4 using FPAREVI,W S_NDWI2, and TS_MODIS).

<table>
<thead>
<tr>
<th>Vegetation Type</th>
<th>Withheld Site/Year</th>
<th>εmax (g C MJ⁻¹)</th>
<th>r</th>
<th>RMSE (g C m⁻² d⁻¹)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grasslands</td>
<td>Dz2</td>
<td>1.84</td>
<td>0.96</td>
<td>0.60</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Xi1</td>
<td>2.05</td>
<td>0.80</td>
<td>0.57</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Xi2</td>
<td>1.87</td>
<td>0.89</td>
<td>0.64</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Xi5</td>
<td>2.11</td>
<td>0.83</td>
<td>0.44</td>
<td>73</td>
</tr>
<tr>
<td>Croplands</td>
<td>Dz1-2005</td>
<td>2.38</td>
<td>0.91</td>
<td>0.89</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Dz1-2006</td>
<td>2.51</td>
<td>0.91</td>
<td>0.90</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Bed-2006</td>
<td>5.18</td>
<td>0.88</td>
<td>1.98</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Bed-2007</td>
<td>5.38</td>
<td>0.97</td>
<td>0.95</td>
<td>43</td>
</tr>
<tr>
<td>Forests</td>
<td>Bed-2006</td>
<td>2.38</td>
<td>0.91</td>
<td>0.89</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Bed-2007</td>
<td>2.51</td>
<td>0.91</td>
<td>0.90</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Bed-2007</td>
<td>5.18</td>
<td>0.88</td>
<td>1.98</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Bed-2007</td>
<td>5.38</td>
<td>0.97</td>
<td>0.95</td>
<td>43</td>
</tr>
</tbody>
</table>

* N denotes the valid number of testing points used in the validation for each site or year.

Table 6
The Sobol’ sensitivity indices of model input variables of the optimized model (No. 6 in Table 4 using FPAREVI,W S_NDWI2, and TS_MODIS) for the three vegetation types. Higher values indicate higher uncertainty resulting from the variables.

<table>
<thead>
<tr>
<th>Vegetation Type</th>
<th>Sensitivity Index</th>
<th>PAR</th>
<th>EVI</th>
<th>TMIN</th>
<th>NDWI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grasslands</td>
<td>First order index</td>
<td>0.26</td>
<td>0.21</td>
<td>0.20</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>Second order index</td>
<td>0.14</td>
<td>0.14</td>
<td>0.13</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Total order index</td>
<td>0.43</td>
<td>0.37</td>
<td>0.33</td>
<td>0.16</td>
</tr>
<tr>
<td>Croplands</td>
<td>First order index</td>
<td>0.29</td>
<td>0.20</td>
<td>0.22</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Second order index</td>
<td>0.14</td>
<td>0.13</td>
<td>0.12</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>Total order index</td>
<td>0.46</td>
<td>0.35</td>
<td>0.33</td>
<td>0.12</td>
</tr>
<tr>
<td>Forests</td>
<td>First order index</td>
<td>0.29</td>
<td>0.23</td>
<td>0.17</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Second order index</td>
<td>0.14</td>
<td>0.13</td>
<td>0.10</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Total order index</td>
<td>0.46</td>
<td>0.37</td>
<td>0.26</td>
<td>0.16</td>
</tr>
</tbody>
</table>
3.3. Regional GPP estimation and uncertainty analysis

We compared the spatial distribution of the averaged mean annual GPP of the ten schemes with mean annual GPP of the MOD17A2 GPP product during the period 2001 to 2012 (Fig. 7(a)–(b)). Generally, the spatial patterns of the two GPP products were similar, with GPP decreasing from southeast to northwest. However, MODIS GPP showed less spatial variations and lower values than our GPP map, particularly in the central areas with mixed vegetation types.

The central mixed areas, mainly covered by forests, had the highest GPP for our GPP model (800–1400 g C m\(^{-2}\) yr\(^{-1}\)) and MODIS GPP (400–1000 g C m\(^{-2}\) yr\(^{-1}\)). For the croplands in the southeast, our estimated GPP (400–800 g C m\(^{-2}\) yr\(^{-1}\)) was higher than MODIS GPP (less than 600 g C m\(^{-2}\) yr\(^{-1}\)). For the Xilingol grasslands, our estimated GPP showed higher value (200–400 g C m\(^{-2}\) yr\(^{-1}\)) in the central and eastern areas and lower values (less than 200 g C m\(^{-2}\) yr\(^{-1}\)) in the west; and the MODIS GPP exhibited similar patterns with our GPP model in the east and west.

The SD, indicating the GPP uncertainties of the ten schemes (Fig. 7(c)), had higher values (180–240 g C m\(^{-2}\) yr\(^{-1}\)) in the central mixed areas. The lower SD values (0–60 g C m\(^{-2}\) yr\(^{-1}\)) were found in the homogeneous grassland areas in Xilingol. The croplands in the southeast showed medium GPP uncertainty of 90–180 g C m\(^{-2}\) yr\(^{-1}\).

The regional summed GPP and its uncertainty measured by SD are shown in Fig. 8. Grasslands are the dominant type that exceeds approximately 50% of the study area (Fig. 4) and had the highest regional GPP compared with croplands and forests, which varied between 101.7 Tg C yr\(^{-1}\) (year 2007) and 147.8 Tg C yr\(^{-1}\) (year 2012). The uncertainty of the grasslands measured by SD ranged from 23.6 Tg C yr\(^{-1}\) (year 2007) to 31.4 Tg C yr\(^{-1}\) (year 2012). The GPP for croplands ranged from 78.7 Tg C yr\(^{-1}\) (year 2001) to 96.2 Tg C yr\(^{-1}\) (year 2012), and had the smallest uncertainty between 8.8 Tg C yr\(^{-1}\) (2001) and 13.7 Tg C yr\(^{-1}\) (2012). Forests only account for approximately 10% of the study area. Although the annual variations in forest GPP were the smallest, with 47.9 Tg C yr\(^{-1}\) in 2006 and 56.5 Tg C yr\(^{-1}\) in 2012, the uncertainty measured by the SD was the largest, with 27.4 Tg C yr\(^{-1}\) (year 2007) to 33.5 Tg C yr\(^{-1}\) (year 2012). The high GPP uncertainty of forests may be attributed to the large inconsistency in land cover classification among different land cover maps for forests (Fig. 4).

In addition, we averaged the 8-day interval GPP of the ten schemes during the period from 2001 to 2012 (Fig. 9). The growth curves (or seasonal cycles) varied with vegetation type and year. The red line represents the relative uncertainty (RU) of the ten schemes with larger values indicating larger discrepancies among the ten schemes. Overall, croplands had the highest RU (20–60%), followed by forests and grasslands (Fig. 9). At the beginning and end of the growing season (the steepest segments of the blue curves), the grassland GPP showed minor fluctuations with lower RU. Along with increasing GPP, the RU of grasslands increased and reached their highest values at the peak of the growing season. However, croplands and forests showed relatively high RU at the beginning and end of the growing season but relatively low RU at the peak of the growing season. This is primarily because croplands and forests are vulnerable to disturbances (such as drought, heat waves, or insect outbreaks) during the green up stage in early spring and tend to be stable while in the growing season. However, grasslands are ecologically fragile and susceptible to water and temperature changes in the entire growing season.
In addition, as shown in Figs. 10–11, modifying one factor (Sch1–3) had less influence than modifying two (Sch4–7) or three factors (Sch8–9). However, there were no obvious differences between the effects of modifying two and three factors.

4. Discussion

4.1. Influence of model structures and sensitivity of model input variables

Model structure is a large source of uncertainty in GPP estimates. LUE models with different model structures (i.e., different representations of FPARs and environmental stresses) have been used to estimate GPP at the site or ecosystem levels, but there are controversies on the relative performances of different representations for each model structure (Rossini et al., 2012; Yuan et al., 2014; Zhou et al., 2014; Liu et al., 2017). Our analysis elucidated the effects of different model re-
structures and representations on the uncertainty in GPP estimates.

FPAR is usually represented by vegetation indices (Vis), such as NDVI, EVI, LAI, and FPAR products (e.g., MODIS) in LUE models. Generally, there may be no significant differences in the correlation between different FPARs and GPP (Yuan et al., 2014). In our study, the four FPAR structures acted differently across vegetation types in GPP simulation. Generally, NDVI is sensitive to grassland greenness (Rossini et al., 2012), but it easily saturates in multi-layer closed canopies compared with EVI (Xiao et al., 2004). Some studies found that the FPAR reconstructed from the GLASS LAI using Beer’s Law improved the performance of LUE models (Yan et al., 2016; Zhu et al., 2016). In our study, however, the performance of the GPP model based on FPARGLASS slightly improved for croplands and decreased for grasslands and forests. The GLASS LAI product has higher quality and accuracy but still lacks rigorous validation because of the scarcity of ground measurements (Xiao et al., 2016b). The discrepancies of optimal FPAR among different studies suggested that vegetation properties and regional characteristics should be fully considered in LUE models.

The environmental stresses in LUE models mainly refer to water and temperature stresses that regulate GPP. The water scalar can be quantified by different methods, such as plant moisture indicators (e.g., NDWI and evaporative fraction (EF)) (Xiao et al., 2004; Yuan et al., 2007), atmospheric indicator (VDP) (Zhao et al., 2005), or soil moisture. In our study, the actual LUE was more responsive to NDWI than to VDP with higher GPP model performance by using WSNDWI2 (Eq. (9)) for the three vegetation types. Previous studies also indicated that plant moisture indicator - NDWI was more sensitive to changes in leaf water content and could better represent changes in available water in vegetation than atmospheric indicator VPD and soil moisture, especially in spatial heterogeneous areas (Yuan et al., 2007; Zhang et al., 2015b). Additionally, the performance of the GPP model by changing WS from WSNDWI (Eq. (8)) to WSNDWI2 (Eq. (9)) improved for grasslands and croplands but decreased for forests, primarily because canopy greenness and light use efficiency are more sensitive to short-term water changes for grasslands and croplands than for forested ecosystems (Zhang et al., 2011; He et al., 2014; Sims et al., 2014; Zhang et al., 2015b).

The daily average temperature and daily minimum temperature are usually used to assess the thermal stress on GPP. The TS MODIS using the daily minimum temperature exhibited better model performance than TS TEM and TS CASA, particularly for croplands. TS TEM uses three parameters including minimum, optimum, and maximum photosynthesis temperature (Raich et al., 1991), whereas TS CASA only utilizes optimum temperature (Potter et al., 1993). Our results, however, suggested that there was no notable difference in model performance between the two temperature scalars. It implies that fewer parameters for TS may not always enhance model performance.

Overall, the combination of FPAR EVI, WSNDWI2, and TS MODIS exhibited better performance in northern China for the three vegetation types and was selected as the optimized model in our study. Using the
optimized model, the sensitivity analysis showed that the total effects of meteorological datasets (PAR and TMIN) had larger contribution to the uncertainty of simulated GPP than the effects of surface reflectance-based indices (EVI and NDWI). Compared with other studies, our optimized model using the TS_MODIS with daily minimum temperature obviously improved the sensitivity of temperature (Li et al., 2012; He et al., 2014), indicating that the modeled GPP based on TS_MODIS can better capture the effects of small variations in temperature stress. For water scalars, the Ws_NDWI improved the model performance for grasslands and croplands but still had low sensitivity, indicating that the effects of slight fluctuation in water stress on GPP cannot be directly reflected by Ws_NDWI. The future incorporation of Ws structure with

---

**Fig. 10.** Spatial differences in the estimated annual GPP between the altered schemes (Sch1 to Sch9) and the reference scheme over the period 2001–2012. The modifications of the altered schemes (Sch1 to Sch9) are shown in the corresponding brackets.

**Fig. 11.** Effects of the altered schemes (modifying the spatial resolutions, land cover maps, and meteorological datasets) and the reference scheme on the magnitude, spatial pattern, and temporal pattern of simulated GPP over the period 2001–2012. The modifications in the altered schemes (Sch1 to Sch9) are shown in the corresponding brackets.
higher sensitivity in LUE models is expected to reduce the uncertainty in GPP estimation.

The sensitivity analysis is also useful for identifying interactions between variables in LUE models. The higher order indices (second and third order) showed that the interactions among model input variables (PAR, EVI, TMIN, and NDWI) could not be neglected. Thus, the contribution of one input variable to GPP was related to the values of other inputs. For example, the contribution of EVI to simulated GPP in tropical regions (with abundant solar radiation and higher temperature) may differ with that in boreal regions (with less abundant solar radiation and lower temperature). Our analysis indicated that the effects of an input variable on GPP may vary with different vegetation (e.g., vegetation type, EVI), meteorological conditions (e.g., PAR, TMIN), or hydrological conditions (e.g., NDWI, VPD).

4.2. Uncertainty of parameters and land cover maps on GPP estimation

The variation of parameters can result in considerable uncertainty in GPP estimates (Xiao et al., 2014; Wagle et al., 2016). In LUE models, $E_{\text{max}}$ is the parameter that determines the expected rate of photosynthesis assuming optimal conditions and introduces the largest uncertainty to GPP (Wang et al., 2013; Yuan et al., 2015). However, large discrepancies in $E_{\text{max}}$ exist among different models and studies due to different model structures, parameters and regions. In the MODIS GPP/NPP algorithm, the $E_{\text{max}}$ is a biome-specific parameter and varies from 0.604 to 1.259 g C MJ$^{-1}$ (Running and Zhao, 2015). EC-LUE (Yuan et al., 2010) and CASA (Potter et al., 1993) assumed biome-independent $E_{\text{max}}$ of 2.25 g C MJ$^{-1}$ and 0.389 g C MJ$^{-1}$, respectively, for large scale carbon flux simulations. Our study implied that the optimized $E_{\text{max}}$ varied with model structure and vegetation type. Therefore, for a certain model or a specific region, fixed or empirical values for $E_{\text{max}}$ will lead to systematic uncertainties or errors in GPP (Wagale et al., 2016), and rigorous model parameterization and calibration are critical for accurate carbon flux simulation. Wang et al. (2010) proposed that spatially explicit $E_{\text{max}}$ values should be used for GPP modeling in highly heterogeneous regions characterized by a variety of plant functional types and mixed landscapes.

Land cover maps were more influential on GPP modeling than spatial resolutions and meteorological datasets, especially in heterogeneous areas. The land cover maps were based on different sources of data, resolutions, classification schemes, and classifiers, and therefore exhibited obvious discrepancies, especially for forest (Giri et al., 2005). It was reported that the overall accuracy was 83.15% globally for Globeland30-2010 (Chen et al., 2015) and was approximately 75% for the MODIS land cover map (Friedl et al., 2010). The GLC2000 in China was verified using national statistics data with high accuracy for grasslands (97.68%), croplands (70.29%), and forests (88.39%) (Xu et al., 2005). The biases or misclassifications of land cover maps could introduce uncertainty to regional GPP estimation through parameters, especially $E_{\text{max}}$. High-accuracy land cover maps are needed for regional-scale GPP estimation. In addition, land cover types may change over time due to land cover/land use change (e.g., deforestation, afforestation, urbanization, stand-replacing disturbances). Therefore, the availability and adoption of dynamic land cover maps are expected to reduce the uncertainty in GPP estimation at regional to global scales.

Besides the uncertainty in land cover maps, the spatial resolutions of input data particularly land cover maps can also introduce uncertainty in GPP estimates at the regional scale. With the development of remote sensing technology, multi-source and multi-scale datasets are becoming available and necessary in carbon flux models. Given spatial heterogeneity and model nonlinearity, spatial scale effects exist for variables derived from remote sensing data with different resolutions (Liu, 2014). Studies have indicated that the use of land cover maps with coarse spatial resolutions can lead to large uncertainty in regional carbon flux estimates in heterogeneous regions (Quaife et al., 2008; Xiao et al., 2011). Few previous studies examined the uncertainty in GPP introduced by both land cover and meteorological data with different spatial resolutions. Our results showed that the scale effects resulting from different spatial resolutions on the uncertainty in GPP could not be neglected although different spatial resolutions had smaller effects than land cover datasets. Meanwhile, there is a tradeoff between spatial resolution and computational efficiency, and finer spatial resolution has a higher demand for computation. Therefore, the optimum spatial resolutions for various cases is increasingly important.

4.3. Uncertainty of meteorological datasets on GPP estimation

Meteorological reanalysis datasets have been major climate input and critical drivers for regional GPP estimation. However, reanalysis datasets have uncertainties, especially for heterogeneous areas with sparse climate stations and small-scale convection processes (Running and Zhao, 2015). Previous studies found that the bias of meteorological reanalysis datasets introduced considerable uncertainties to GPP simulation (Heinsch et al., 2006; Cai et al., 2014; Jin et al., 2015). In our study, input meteorological variables including PAR and TMIN had large contribution to GPP uncertainty with relatively high first and total order indices over 0.20. PAR is the energy source of photosynthesis and has been considered to be the most influential driver in carbon flux modeling (Ichii et al., 2005; Mercado et al., 2009). Meteorological reanalysis datasets tend to overestimate radiation (Zhao et al., 2006; Cai et al., 2014) because of the underestimate of atmospheric solar absorption, particularly water vapor absorption (Wild, 2005). Clouds and aerosols can also impact the uncertainty in GPP estimation by altering the proportions of direct and diffuse PARs because diffuse PAR is more easily absorbed by the bottom canopy in shadow (Gu et al., 2002; Li and Fang, 2015).

Although PAR and TMIN were highly sensitive variables in our GPP model, the effects of the two meteorological datasets (CMFD and MERRA-2) on the uncertainty in regional GPP estimates were smaller than those of land cover maps and spatial resolutions. The smaller effects of meteorological datasets were perhaps because of the relatively high accuracy and consistency of PAR$_{CMFD}$ and PAR$_{MERRA_2}$ in our study area. Overall, meteorological reanalysis datasets should be carefully evaluated and/or calibrated, and accurate datasets should be used to reduce the biases of GPP (Jin et al., 2015), particularly at regional or global scales.

4.4. Other sources of uncertainty and further improvements

Additional sources of uncertainty in simulated GPP exist, and further improvements should be considered in future studies. First, the eddy covariance flux data still have significant uncertainties, although we considered them as ground truth. The flux tower GPP data have random and system errors due to the limitations of measurement technique and instrument, the stochastic nature of turbulence, partitioning methods and gap-filling techniques (Richardson et al., 2006; Moffat et al., 2007). These uncertainties could propagate to the parameter estimation in modeling. Second, uncertainties due to scale mismatches between image pixels and tower footprints should be considered, especially in heterogeneous areas (Tan et al., 2006; Sjostrom et al., 2013). Although these uncertainties are inevitable in GPP modeling, their influence on our uncertainty quantification is relatively small. This is because we focused on how different model structures, different input datasets and model parameterizations influence the uncertainty in GPP based on the same sites. Third, we made simplifications on vegetation types (or plant functional types). We simplified the land cover maps by combining the classes into four broader vegetation types, which may introduce uncertainty to GPP simulation while make it possible for us to compare the influence of the three land cover maps. Fourth, we used one forest site (DBF site) to represent the forests across the study area in model parameterization. On one hand, our study area is mainly covered by DBF; on the other hand, the differences...
of parameters in LUE models among different forest types in adjacent regions are relatively small. Nevertheless, the availability of more forested flux sites will benefit our analysis by reducing the uncertainty in parameter estimation. Finally, our study did not distinguish C3/C4 grasslands and croplands in GPP simulation due to the lack of spatially explicit C3/C4 maps. The northern China is characterized by mixed C3/C4 grasslands (Guan et al., 2012) and rotated C3/C4 croplands. C4 plants have stronger photosynthetic capacities than C3 plants under similar climate conditions, therefore the variations in the C3 and C4 components may introduce uncertainty to GPP estimates (Suyker et al., 2005; Zhang et al., 2014; Yuan et al., 2015).

5. Conclusions

In this study, we examined various sources of uncertainty in GPP simulated by LUE models: model structure, parameters, input data, and spatial resolution. We found that model structures (i.e., FPAR, $W_b$, and $T_s$) impacted both model parameters (e.g., $\epsilon_{\text{max}}$) and model performance in a manner that varied with vegetation type. Proper model structure and rigorous model parameterization and calibration are critical for accurate GPP estimation. At the regional scale, land cover maps had the largest effects on the uncertainty in GPP estimation, followed by spatial resolutions and meteorological datasets. Our study highlights the importance of high-accuracy land cover maps for GPP estimation at regional scales. In addition, the scale effects of different spatial resolutions should be paid more attention in GPP estimation, especially in heterogeneous areas. Furthermore, meteorological drivers including PAR and TMIN were sensitive input variables in GPP simulation. Meteorological data should be carefully evaluated and/or calibrated, and accurate meteorological data can reduce the uncertainty in modeled GPP, particularly at regional or global scales.

Our findings can help us better understand the potential sources of uncertainty and their relative effects on GPP estimates, and have implications for reducing uncertainty in carbon flux simulation and global carbon budget. The use of relatively simple LUE models allows us to conduct various experiments to quantify the uncertainties in GPP from different sources. Process-based ecosystem (or terrestrial biosphere) models (Cramer et al., 1999; Xiao et al., 2009; Thorn et al., 2015; Ma et al., 2017), however, are more complex with a large number of parameters, more input datasets, and higher computational demand. Uncertainty analysis should be extended to these more complex ecosystem models to diagnose and quantify the various sources of uncertainty and thereby to better quantify the full uncertainty in regional or global carbon fluxes.

Acknowledgments

This study is supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA19030302), the Major Projects of the International Partnership Program of Chinese Academy of Sciences (Grant No. 131011KYSB20160061), the National Natural Science Foundation of China (Grant No. 41590853), the One-Three-Five Planning of RADI (Grant No. Y75G1200CX). J. Xiao was supported by National Aeronautics and Space Administration (NASA) (Carbon Cycle Science Program: Grant No. NNX14AJ18G; Climate Indicators and Data Products for Future National Climate Assessments: Grant No. NNX16AG61G). This work used eddy covariance data acquired by the FLUXNET community, particularly by ChinaFlux and USCCC networks. The China Meteorological Forcing Dataset is developed by Data Assimilation and Modeling Center for Tibetan Multi-spheres, Institute of Tibetan Plateau Research, Chinese Academy of Sciences. The GlobeLand30 dataset is provided by National Geomatics Center of China. We thank the anonymous reviewers and the Editor for their constructive comments on the manuscript.

Appendix A. List of abbreviations or variables in this paper

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\epsilon_{\text{max}}$</td>
<td>maximum light use efficiency (g C MJ$^{-1}$)</td>
</tr>
<tr>
<td>CMFD</td>
<td>China Meteorological Forcing Dataset</td>
</tr>
<tr>
<td>EVI</td>
<td>enhanced vegetation index</td>
</tr>
<tr>
<td>FPAR</td>
<td>fraction of absorbed photosynthetically active radiation</td>
</tr>
<tr>
<td>GLASS LAI</td>
<td>Global Land Surface Satellite Leaf Area Index</td>
</tr>
<tr>
<td>GPP</td>
<td>gross primary productivity (g C m$^{-2}$ d$^{-1}$; g C m$^{-2}$ yr$^{-1}$; Tg C yr$^{-1}$)</td>
</tr>
<tr>
<td>LAI</td>
<td>leaf area index</td>
</tr>
<tr>
<td>NDVI</td>
<td>normalized difference water index</td>
</tr>
<tr>
<td>LUE</td>
<td>light use efficiency (g C MJ$^{-1}$)</td>
</tr>
<tr>
<td>MAT</td>
<td>mean annual temperature (°C)</td>
</tr>
<tr>
<td>MAP</td>
<td>mean annual precipitation (mm)</td>
</tr>
<tr>
<td>MERRA-2</td>
<td>Modern-Era Retrospective analysis for Research and Applications, Version 2</td>
</tr>
<tr>
<td>MODIS</td>
<td>moderate resolution imaging spectroradiometer</td>
</tr>
<tr>
<td>NEE</td>
<td>net ecosystem exchange (g C m$^{-2}$ d$^{-1}$)</td>
</tr>
<tr>
<td>NEE$_d$</td>
<td>daytime NEE (g C m$^{-2}$ d$^{-1}$)</td>
</tr>
<tr>
<td>PAR</td>
<td>photosynthetically active radiation (MJm$^{-2}$ d$^{-1}$)</td>
</tr>
<tr>
<td>RE</td>
<td>ecosystem respiration (g C m$^{-2}$ d$^{-1}$)</td>
</tr>
<tr>
<td>RE$_d$</td>
<td>daytime ecosystem respiration (g C m$^{-2}$ d$^{-1}$)</td>
</tr>
<tr>
<td>RMSE</td>
<td>root mean square error</td>
</tr>
<tr>
<td>RU</td>
<td>relative uncertainty</td>
</tr>
<tr>
<td>SWRad</td>
<td>downward shortwave radiation (MJ m$^{-2}$ d$^{-1}$)</td>
</tr>
<tr>
<td>$T$</td>
<td>air temperature (°C)</td>
</tr>
<tr>
<td>$T_s$</td>
<td>temperature scalar</td>
</tr>
<tr>
<td>$T_{\text{min}}/T_{\text{opt}}/T_{\text{max}}$</td>
<td>minimum/ optimum/ maximum photosynthesis temperature (°C)</td>
</tr>
<tr>
<td>TMIN</td>
<td>daily minimum temperature (°C)</td>
</tr>
<tr>
<td>TMIN$_{\text{min}}$</td>
<td>daily TMIN when the actual LUE equals 0 (°C)</td>
</tr>
<tr>
<td>TMIN$_{\text{max}}$</td>
<td>daily TMIN when the actual LUE equals $\epsilon_{\text{max}}$ (°C)</td>
</tr>
</tbody>
</table>
References


