Remote Sensing of Environment 183 (2016) 154–169

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Consistency between sun-induced chlorophyll fluorescence and gross primary production of vegetation in North America

Yao Zhang [a], Xiangming Xiao [a,b,*], Cui Jin [a], Jinwei Dong [a], Sha Zhou [c], Pradeep Wagle [a], Joanna Joiner [d], Luis Guanter [e], Yongguang Zhang [f], Geli Zhang [a], Yuanwei Qin [a], Jie Wang [a], Berrien Moore III [g]

[a] Department of Microbiology and Plant Biology, Center for Spatial Analysis, University of Oklahoma, Norman, OK 73019, USA
[b] Institute of Biodiversity Science, Fudan University, Shanghai 200433, China
[c] State Key Laboratory of Hydroscience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing, China
[d] NASA Goddard Space Flight Center, Greenbelt, MD, USA
[e] Helmholtz Centre Potsdam, German Research Center for Geosciences (GFZ), Telegrafenberg A17, 14473 Potsdam, Germany
[f] Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, International Institute for Earth System Sciences, Nanjing University, 210023 Nanjing, China
[g] College of Atmospheric and Geographic Sciences, University of Oklahoma, Norman, OK 73019, USA

A R T I C L E I N F O

Article history:
Received 16 August 2015
Received in revised form 17 May 2016
Accepted 23 May 2016
Available online xxxx

Keywords:
Vegetation Photosynthesis Model (VPM)
Light use efficiency
Remote sensing
SIF
MODIS
Carbon cycle
GPP product

A B S T R A C T

Accurate estimation of the gross primary production (GPP) of terrestrial ecosystems is vital for a better understanding of the spatial-temporal patterns of the global carbon cycle. In this study, we estimate GPP in North America (NA) using the satellite-based Vegetation Photosynthesis Model (VPM), MODIS images at 8-day temporal and 500 m spatial resolutions, and NCEP-NARR (National Center for Environmental Prediction-North America Regional Reanalysis) climate data. The simulated GPP (GPP_{VPM}) agrees well with the flux tower derived GPP (GPP_{EC}) at 39 AmeriFlux sites (155 site-years). The GPP_{VPM} in 2010 is spatially aggregated to 0.5 by 0.5° grid cells and then compared with sun-induced chlorophyll fluorescence (SIF) data from Global Ozone Monitoring Instrument 2 (GOME-2), which is directly related to vegetation photosynthesis. Spatial distribution and seasonal dynamics of GPP_{VPM} and GOME-2 SIF show good consistency. At the biome scale, GPP_{VPM} and SIF shows strong linear relationships ($R^2 > 0.95$) and small variations in regression slopes (4.60–5.55 g C m^{-2} day^{-1}/mW m^{-2} nm^{-1} sr^{-1}). The total annual GPP_{VPM} in NA in 2010 is approximately 13.53 Pg C $year^{-1}$, which accounts for ~11.0% of the global terrestrial GPP and is within the range of annual GPP estimates from six other process-based and data-driven models (11.35–22.23 Pg C $year^{-1}$). Among the seven models, some models did not capture the spatial pattern of GOME-2 SIF data at annual scale, especially in Midwest cropland region. The results from this study demonstrate the reliable performance of VPM at the continental scale, and the potential of SIF data being used as a benchmark to compare with GPP models.

1. Introduction

Carbon dioxide fixed through photosynthesis by terrestrial vegetation is known as gross primary production (GPP) at the ecosystem level. Increased carbon uptake during the past decades helped offset growing CO_2 emissions from fossil fuel burning and land cover change and mitigate the increase of atmospheric CO_2 concentration and global climate warming (Ballantyne, Alden, Miller, Tans, & White, 2012). A variety of approaches have been used to estimate GPP of terrestrial ecosystems, and they can be grouped into four categories: 1) process-based GPP models; 2) satellite-based production efficiency models (PEM); 3) data-driven GPP models upscaled from eddy covariance data; and 4) models based on sun-induced chlorophyll fluorescence (SIF) (Fig. 1). However, large uncertainty still remains regarding the spatial distribution and seasonal dynamics of GPP, which limits our capability to address scientific questions related to the increasing seasonal amplitude and interannual variation of atmospheric CO_2 (Forkel et al., 2016; Graven et al., 2013; Poulter et al., 2014). An accurate estimation of GPP at regional and global scales is essential for a better understanding of the underlying mechanisms of ecosystem-climate interactions and ecosystem response to extreme climate events, such as drought, heat wave, and flood, etc. (Beer et al., 2010; Yu et al., 2013; Zhang et al., 2016).

Many process-based biogeochemical models employ the enzyme kinetics theory, most well-known as encapsulated by Farquhar, Caemmerer, and Berry (1980) and its modification for C4 plants (Collatz, Ribas-Carbo, & Berry, 1992). Some process-based models employ the light-use-efficiency (LUE) concept to estimate GPP (Zeng, Mariotti, & Wetzel, 2005). These models also take multiple

* Corresponding author at: Department of Microbiology and Plant Biology, Center for Spatial Analysis, University of Oklahoma, Norman, OK 73019, USA.
E-mail address: xiangming.xiao@ou.edu (X. Xiao).

http://dx.doi.org/10.1016/j.rse.2016.05.015
0034-4257/

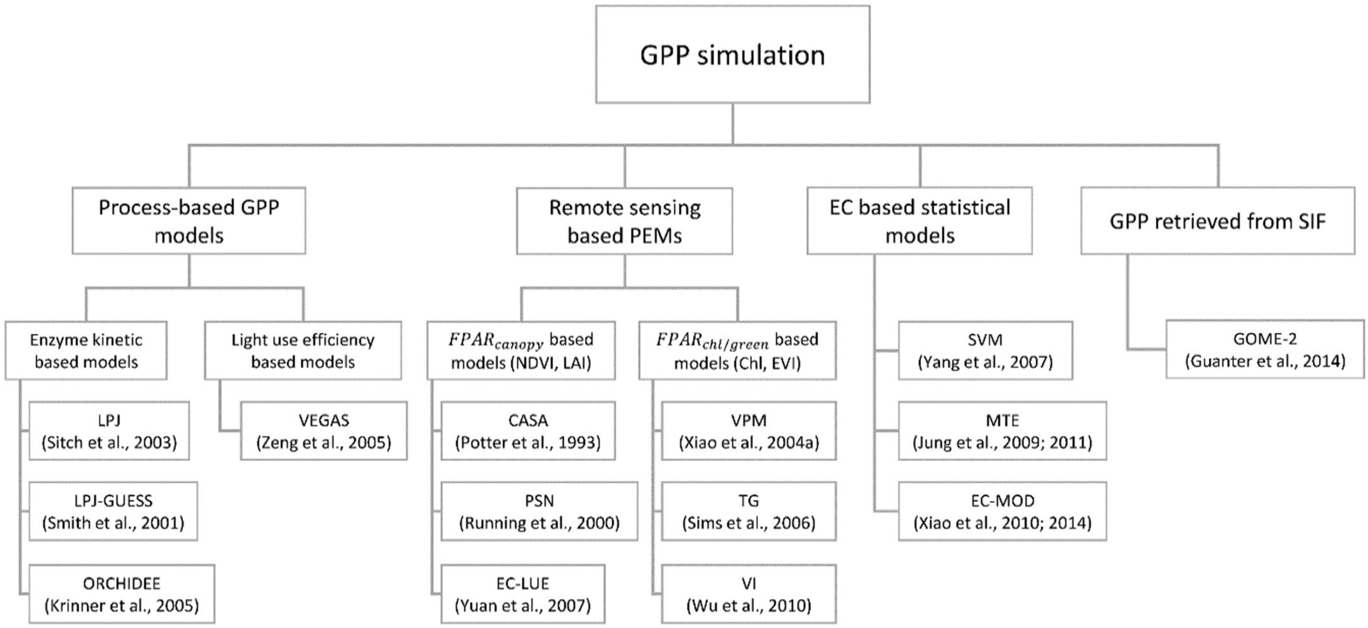

Fig. 1. A list of different approaches and models (as examples) to estimate gross primary production (GPP) of vegetation.

ecological processes into consideration so that they can be coupled with general circulation models (GCMs) to predict feedbacks related to the global warming and CO_2 fertilization (Booth et al., 2012; Keenan et al., 2012; Piao et al., 2013; Xia et al., 2014). However, these models are often run at coarse spatial resolution and the simulation results vary enormously even with the same set of meteorological input datasets (Coops, Ferster, Waring, & Nightingale, 2009).

The remote sensing based PEMs estimate GPP as the product of the energy absorbed by plants (absorbed photosynthetically active radiation, APAR) and LUE that converts energy to carbon fixed during the photosynthesis process (Monteith, 1972). These models can be further divided into two subcategories (Dong et al., 2015a; Xiao et al., 2004a). The $FPAR_{canopy}$ based models, including the Carnegie Ames Stanford Approach (CASA) (Potter et al., 1993), the MODIS GPP algorithm (Photosynthesis, PSN) (Running et al., 2004; Zhao, Heinsch, Nemani, & Running, 2005), and the EC-LUE model (Yuan et al., 2007), use the radiation absorbed by vegetation canopy. The $FPAR_{chl/green}$ based models use radiation absorbed by chlorophyll or green leaves and include the Vegetation Photosynthesis Model (VPM) (Xiao et al., 2004a; Xiao et al., 2004b), Greenness and Radiation (GR) model (Gitelson et al., 2006), and the Vegetation Index (VI) model (Wu, Niu, & Gao, 2010b).

The eddy covariance (EC) technique provides estimates of GPP by partitioning measured net ecosystem CO_2 exchange (NEE) between land and the atmosphere into GPP and ecosystem respiration (R_e) (Baldocchi et al., 2001). Over the past decades, the EC technique has been widely applied to measure NEE of various biome types throughout the world, and a large amount of GPP data (GPP_{EC}) has been accumulated (Baldocchi, 2014; Baldocchi et al., 2001). A number of statistical models have been developed to upscale GPP_{EC} from individual sites to the regional scales (Jung, Reichstein, & Bondeau, 2009; Jung et al., 2011; Xiao et al., 2010; Xiao et al., 2014; Yang et al., 2007). These algorithms, such as model tree ensembles (MTE) or regression tree approaches, build a series of rules through data mining that relate *in situ* flux observations to satellite-based indices and climate data.

Sun-induced chlorophyll fluorescence (SIF), a byproduct of the vegetation photosynthesis process, has been recently retrieved using multiple satellite platforms/instruments such as the Greenhouse gases Observing SATellite (GOSAT) (Frankenberg et al., 2011; Guanter et al., 2012; Joiner et al., 2011; Joiner et al., 2012), the Global Ozone Monitoring Instrument 2 (GOME-2) (Joiner et al., 2013), and the Orbiting Carbon Observatory-2 (OCO-2) (Frankenberg et al., 2014). Recent field studies and theory suggest that SIF contains information from both APAR and LUE that is complementary to vegetation indices such as the normalized difference vegetation index (NDVI) (Guanter et al., 2013; Rossini et al., 2015; Yang et al., 2015). A simple regression model based on space-borne SIF has been developed to estimate cropland GPP (Guanter et al., 2014). Zhang et al. (2014) have also shown the potential of SIF data to improve carbon cycle models and provide accurate projections of agricultural productivity (Guan et al., 2015).

Over the past several years, a number of studies have run the VPM with *in situ* climate data at various eddy flux tower sites. The resulting GPP_{VPM} was evaluated with GPP_{EC} at different ecosystem types, including forests (Xiao et al., 2004a, 2004b, 2005), croplands (Kalfas, Xiao, Vanegas, Verma, & Suyker, 2011; Wagle, Xiao, & Suyker, 2015), savannas (Jin et al., 2013), and grasslands (He et al., 2014; Wagle et al., 2014). Wu, Munger, Niu, and Kuang (2010a) compared GPP from four models driven by remotely sensed data at the Harvard forest site and found that VPM performed best in terms of capturing the seasonal dynamics of GPP. Yuan et al. (2014) compared seven LUE based models at 157 eddy flux sites and showed that VPM had a moderate rank of performance. Dong et al. (2015a) used four EVI-based models to estimate GPP of grasslands and croplands under normal and severe drought conditions, and reported that VPM performed better than other models in capturing the impacts of drought on GPP. This was mostly because VPM uses Land Surface Water Index (LSWI) that is sensitive to water stress (Wagle et al., 2014, 2015), while the other three models lack a water stress scalar. Recently, simulations of VPM on the regional scale, driven by regional climate data, have been carried out in the Tibetan Plateau (He et al., 2014) and China (Chen et al., 2014), where only limited GPP_{EC} data are available for model calibration and validation.

In this study, we aim to assess the feasibility and performance of the VPM model in estimating GPP across North America (NA) and explore the relationship between SIF and GPP_{VPM} at continental scale. The selection of the NA as study area is based on two facts: (1) large uncertainties exist in the GPP estimates from various models (ranging from 12.2 to 32.9 Pg C year^{-1}) (Huntzinger et al., 2012); and (2) a large number of eddy flux sites are available in NA, which provides an opportunity for a thorough validation. The specific objectives of this study are to: (1) implement the VPM simulation at the continental scale over NA; (2) evaluate the performance of VPM at individual sites using GPP_{EC} data from 39 flux tower sites (155 site-years); (3) compare GPP_{VPM} with GOME-2 SIF data at 0.5° (latitude/longitude) resolution across NA; and (4) use of GOME-2 SIF as a reference to compare with GPP estimates from other six models. In this paper, we report (1) multi-year GPP_{VPM} and GPP_{EC} at individual flux tower sites, dependent upon availability of GPP_{EC} data, and (2) GPP_{VPM} in 2010 across NA.

2. Materials and method

2.1. Regional datasets for VPM simulations across North America

2.1.1. Climate data

The VPM model uses photosynthetically active radiation (PAR) and temperature data as climate input data. We use the National Center for Environmental Prediction-North America Regional Reanalysis (NCEP-NARR) datasets (Mesinger et al., 2006) for 2000–2014. The original three hourly data are first aggregated into 8-day averages to match the temporal resolution of MODIS vegetation indices. The day-time mean air temperature is obtained by averaging the temperature between 6 am to 6 pm local time. Zhao, Running, and Nemani (2006) reported that the NCEP-NARR product overestimates the surface shortwave radiation when comparing with the *in situ* observation at the flux tower sites. Jin et al. (2015) also compared the NCEP-NARR radiation data with *in situ* radiation measurements at 37 AmeriFlux sites and reported a bias correction factor of 0.8. In this study, we applied this factor to adjust the radiation data.

In order to run VPM at a 500 m spatial resolution, we use a non-linear spatial interpolation method (Zhao et al., 2005) to downscale the NCEP-NARR radiation and temperature dataset from the spatial resolution of 0.25° × 0.25° to 500-m. It uses a fourth power of a cosine function and adopts the weighted distance from the nearest four grid cells to calculate a value for each output pixel at MODIS resolution. The distance factor (D_i) for the four nearby grid cells can be calculated as follows:

$$D_i = cos^4\left(\frac{\pi}{2}\times\left(\frac{d_i}{d_{max}}\right)\right)\ i = 1, 2, 3, 4 \tag{1}$$

where d_i and $d_{\max}$ indicate the distance between the center of the 500 m MODIS pixel and each of the four vertex grid cells from NCEP-NARR data, and the maximum distance between the four vertex NCEP-NARR grid cells, respectively. For each MODIS pixel, the weight from the four surrounding NCEP-NARR grid cells can be calculated as:

$$W_i = \frac{D_i}{\sum_{i=1}^{4} D_i} \tag{2}$$

The final value for each interpolated MODIS pixel (V) can be expressed as a weighted average:

$$V = \sum_{i=1}^{4}(W_i{*}V_i) \tag{3}$$

where V_i is the value for the four surrounding grid cell values of NCEP-NARR data.

2.1.2. MODIS data

2.1.2.1. MODIS surface reflectance and vegetation indices. The MODIS MOD09A1 surface reflectance product (500 m spatial resolution and 8-day temporal resolution) is used to calculate the enhanced vegetation index (EVI) (Huete et al., 2002) and LSWI as inputs to the VPM. LSWI is calculated as the normalized difference between NIR (0.78–0.89 μm) and SWIR (1.58–1.75 μm) and is sensitive to water content. Therefore, LSWI is a good indicator of water stress from the vegetation canopy and soil background (Xiao, Boles, Liu, Zhuang, & Liu, 2002). These two indices are calculated as follows:

$$EVI = 2.5 \times \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + (6 \times \rho_{red} - 7.5 \times \rho_{blue}) + 1} \tag{4}$$

$$LSWI = \frac{\rho_{nir} - \rho_{swir}}{\rho_{nir} + \rho_{swir}} \tag{5}$$

A temporal gap-fill algorithm is applied to the EVI time series data. The data quality is checked using the quality flag layer, and those observations not affected by cloud and climatological aerosols are considered 'GOOD' quality (MOD35 cloud = 'clear'; aerosol quantity = 'low' or 'average'). Each pixel is temporally linearly interpolated using only good-quality EVI observations within each year. A Savitzky–Golay filter is then applied to each pixel to eliminate high frequency noise (Chen et al., 2004). If a pixel has fewer than three out of 46 good observations for one year, the original data (no gap-filled) are used. Fortunately, this happens only for <0.5% of the total pixels and the majority of those are in less productive, boreal areas.

2.1.2.2. MODIS land cover data. The MODIS MCD12Q1 land cover product at 500-m spatial resolution (Friedl et al., 2010) includes annual land cover types from 2001 to 2013. We use MCD12Q1 data in 2001 to represent year 2000, and MDD12Q1 data in 2013 to represent year 2014, which allows us to have a full time series of land cover types for 2000–2014. The IGBP land cover classification scheme in the dataset is used to provide biome specific information for the VPM. A lookup-table (LUT) is used to get the essential parameters including maximum LUE as well as the maximum, minimum, and optimum temperatures for vegetation photosynthesis (see Appendix Table A1).

In order to investigate the relationship between GPP_{VPM} and SIF (0.5° latitude and longitude resolution) in different vegetation/biome types, we also aggregate the original 500 m land cover data to 0.5° grid cells using the following procedure. The original IGBP land cover data are first merged and reprojected onto the longitude-latitude projection with the original spatial resolution. We calculate the frequency (number of 500-m pixels) of individual vegetation types within a 0.5° × 0.5° grid cell. Then, for each 0.5° × 0.5° grid cell, if one vegetation type is dominant (>75% of the grid cell), this grid cell is assigned that vegetation type; if no land cover type is dominant, the grid cell is not assigned a type.

2.1.2.3. MODIS land surface temperature data. The MODIS MYD11A2 land surface temperature dataset is used to derive the thermal growing season and eliminate the snow cover period, which avoids the effect of snow cover in retrieving the yearly maximum LSWI. The MYD11A2 dataset is chosen because it provides observations at 1:30 am, which is close to the daily minimum temperature. For each pixel each year, the thermal growing season is defined using the nighttime land surface temperature (Dong et al., 2015b). Once three consecutive 8-day's in the spring have nighttime temperatures above 5°C, the thermal growing season begins; when three consecutive 8-day's in the fall have nighttime temperatures below 10°C, the thermal growing season ends. A detailed application of this temperature-based phenology was recently reported (Zhang et al., 2015).

2.2. Datasets used to evaluate and compare VPM simulations across North America

2.2.1. CO_2 eddy flux data from AmeriFlux tower sites

CO_2 flux data from 39 AmeriFlux sites are downloaded from the AmeriFlux data portal (http://ameriflux.ornl.gov/). These flux sites cover most of the major biomes in NA (DBF, ENF, MF, GRA, CRO, CSH, OSH, WET and WSA) (Table 1). The 8-day level-4 gap-filled flux data with the Marginal Distribution Sampling (MDS) method is used (Reichstein et al., 2005). GPP_{EC} estimates from individual sites are used to evaluate GPP_{VPM}.

2.2.2. Sun-induced chlorophyll fluorescence (SIF) data from GOME-2

The latest version (v26) of monthly SIF data from the GOME-2 instrument onboard Eumetsat's MetOp-A satellite is used in this study and available to the public at http://acdb-ext.gsfc.nasa.gov/People/Joiner/my_gifs/GOME_F/GOME-F.htm (Joiner et al., 2014). GOME-2 captures earth radiation in the range from ~600 to 800 nm with a spectral resolution of ~0.5 nm at a nominal nadir footprint of $40 \times 80\ km^2$ in the nominal observing configuration. Wavelengths around 740 nm at the far-red peak of the SIF emission are used for SIF retrievals with a principal component analysis approach to account for atmospheric absorption. The results are then quality-controlled (e.g., heavily cloud contaminated data removed) and aggregated to monthly means at $0.5° \times 0.5°$ spatial resolution (Joiner et al., 2013). In this study, we use GOME-2 SIF data for the period from January 2010 to February 2011.

2.2.3. GPP data from other six models

The GPP data from the four process-based models (LPJ, LPJ-GUESS, ORCHIDEE, and VEGAS) are part of the TRENDY projects (Sitch et al., 2008), which intended to compare trends in net land-atmosphere carbon exchange over the period 1980–2010 (Table 3). These four models, driven by the CRU + NCEP climate data and global annual atmospheric CO_2, are chosen because they have different algorithms to simulate GPP at $0.5° \times 0.5°$ spatial resolution.

Another two models involved in the comparison are the MPI-BGC and MODIS PSN. The MPI-BGC estimates GPP by upscaling global CO_2 flux observations using a Model Tree Ensemble approach (Jung et al., 2009). MODIS PSN employs a production-efficiency approach and uses the MODIS fraction of photosynthetically active radiation product (MOD15A2) and meteorological data (Running et al., 2004). The C55 version of MODIS PSN product (MOD17A2 C55) is used.

Table 1
Descriptions of the 39 flux tower sites used in this study. IGBP class, R^2, and RMSE are the International Geosphere-Biosphere Programme land cover classification, coefficient of determination, and root mean square error of the regression analysis between tower-based gross primary production (GPP_{EC}) and simulated GPP (GPP_{VPM}) using vegetation photosynthesis model.

ID	NAME	LAT	LON	IGBP class	Years used	R^2	RMSE	Reference
US-Bo1	Bondville	40.0062	−88.2904	CRO	2001–2006	0.83	2.20	Hollinger, Bernacchi, and Meyers (2005)
US-Ne1	Mead irrigated continuous	41.1651	−96.4766	CRO	2001–2005	0.91	3.06	Suyker, Verma, Burba, and Arkebauer (2005)
US-Ne2	Mead irrigated rotation	41.1649	−96.4701	CRO	2001–2005	0.91	2.71	Suyker et al. (2005)
US-Ne3	Mead rainfed rotation	41.1797	−96.4397	CRO	2001–2005	0.85	2.76	Suyker et al. (2005)
US-Ro1	Rosemount- G21	44.7143	−93.0898	CRO	2004–2006	0.80	2.45	Griffis, Baker, and Zhang (2005)
US-Ro3	Rosemount- G19	44.7217	−93.0893	CRO	2004–2006	0.81	2.22	Griffis et al. (2005)
US-KS2	Kennedy Space Center	28.6086	−80.6715	CSH	2004–2005	0.72	0.96	Dijkstra et al. (2002)
US-Los	Lost Creek	46.0827	−89.9792	CSH	2001–2002	0.90	1.59	Sulman, Desai, Cook, Saliendra, and Mackay (2009)
US-Bar	Bartlett Experimental Forest	44.0646	−71.2881	DBF	2004–2006	0.93	1.33	Jenkins et al. (2007)
US-Ha1	Harvard Forest	42.5378	−72.1715	DBF	2000–2006	0.83	2.05	Urbanski et al. (2007)
US-LPH	Little Prospect Hill	42.5419	−72.1850	DBF	2001–2005	0.91	1.30	Vanderhoof, Williams, Pasay, and Ghimire (2013)
US-MMS	Morgan Monroe State Forest	39.3232	−86.4131	DBF	2005–2007	0.91	1.59	Schmid, Grimmond, Cropley, Offerle, and Su (2000)
US-MOz	Missouri Ozark Site	38.7441	−92.2000	DBF	2000–2006	0.89	1.37	Gu et al. (2006)
US-UMB	Univ. of Mich. Biological Station	45.5598	−84.7138	DBF	2000–2006	0.97	0.78	Gough, Vogel, Schmid, Su, and Curtis (2008)
US-WCr	Willow Creek	45.8059	−90.0799	DBF	2002–2005	0.96	1.05	Cook et al. (2004)
CA-NS1	UCI-1850 burn site	55.8792	−98.4839	ENF	2003–2005	0.65	1.00	Goulden et al. (2006)
CA-NS2	UCI-1930 burn site	55.9058	−98.5247	ENF	2002–2005	0.70	0.88	Goulden et al. (2006)
CA-NS3	UCI-1964 burn site	55.9117	−98.3822	ENF	2002–2005	0.92	1.49	Goulden et al. (2006)
CA-NS4	UCI-1964 burn site wet	55.9117	−98.3822	ENF	2003–2004	0.84	1.08	Goulden et al. (2006)
CA-NS5	UCI-1981 burn site	55.8631	−98.4850	ENF	2002–2005	0.89	1.13	Goulden et al. (2006)
US-Blo	Blodgett Forest	38.8953	−120.6328	ENF	2000–2006	0.74	1.58	Goldstein et al. (2000)
US-Fmf	Flagstaff Managed Forest	35.1426	−111.7273	ENF	2007	0.63	0.95	Dore et al. (2008)
US-Ho1	Howland Forest (main tower)	45.2041	−68.7402	ENF	2000–2004	0.88	0.84	Hollinger et al. (2004)
US-Ho2	Howland Forest (west tower)	45.2091	−68.7470	ENF	2000–2004	0.69	0.98	Hollinger et al. (2004)
US-Me2	Metolius-intermediate aged pine	44.4523	−121.5574	ENF	2002, 2004–2007	0.91	1.03	Law et al. (2004)
US-Me3	Metolius-second young aged pine	44.3154	−121.6078	ENF	2004–2005	0.69	1.26	Law, Williams, Anthoni, Baldocchi, and Unsworth (2000)
US-Me5	Metolius-first young aged pine	44.4372	−121.5668	ENF	2000–2002	0.94	0.60	Law et al. (2000)
US-NC1	North Carolina Clearcut	35.8115	−76.7115	ENF	2005–2006	0.95	0.93	Noormets et al. (2010)
US-Wi0	Wisconsin young red pine	46.6188	−91.0814	ENF	2002	0.81	1.79	Sun, Noormets, Chen, and McNulty (2008)
US-Wi4	Wisconsin mature red pine	46.7393	−91.1663	ENF	2002–2005	0.92	0.81	Sun et al. (2008)
US-ARb	ARM SGP burn	35.5497	−98.0402	GRA	2005–2006	0.91	1.99	Fischer, Billesbach, Berry, Riley, and Torn (2007)
US-ARc	ARM SGP control	35.5465	−98.0400	GRA	2005–2006	0.91	2.07	Fischer et al. (2007)
US-Goo	Goodwin Creek	34.2547	−89.8735	GRA	2004–2006	0.68	1.93	Wilson and Meyers (2007)
US-Wlr	Walnut River Watershed	37.5208	−96.8550	GRA	2002–2004	0.94	0.81	Coulter et al. (2006)
US-Syv	Sylvania Wilderness Area	46.2420	−89.3477	MF	2001–2006	0.92	1.12	Desai, Bolstad, Cook, Davis, and Carey (2005)
CA-NS6	UCI-1989 burn site	55.9167	−98.9644	OSH	2002–2005	0.87	0.69	Goulden et al. (2006)
CA-NS7	UCI-1998 burn site	56.6358	−99.9483	OSH	2002–2005	0.86	0.63	Goulden et al. (2006)
US-Ivo	Ivotuk	68.4865	−155.7503	WET	2004, 2006	0.67	0.80	Epstein, Calef, Walker, Chapin, and Starfield (2004)
US-FR2	Freeman Ranch-Mesquite Juniper	29.9495	−97.9962	WSA	2004–2006	0.73	1.13	Heinsch et al. (2004)

CRO: cropland; CSH: closed shrublands; DBF: deciduous broadleaf forests; ENF: evergreen needleleaf forest; GRA: grassland; MF: mixed forest; OSH: open shrublands; WET: wetland; WSA: woody savannas.

2.3. A brief description of the Vegetation Photosynthesis Model (VPM)

The satellite-based VPM (Xiao et al., 2004a, 2004b) uses the product of light use efficiency (LUE, ε_g), and absorbed photosynthetically active radiation by chlorophyll ($APAR_{chl}$) to estimate GPP as follows (Fig. 2):

$$GPP = \varepsilon_g \times APAR_{chl} \tag{6}$$

VPM uses the fraction of absorbed photosynthetic active radiation by chlorophyll ($fAPAR_{chl}$) to estimate $APAR_{chl}$. The $fAPAR_{chl}$ is estimated from a linear function of EVI where the coefficient α is set to be 1.0 (Xiao et al., 2004a).

$$APAR_{chl} = fAPAR_{chl} \times PAR \tag{7}$$

$$fAPAR_{chl} = \alpha \times EVI \tag{8}$$

The light-use-efficiency (ε_g) in the VPM is a down-regulation of maximum LUE (ε_0) by temperature (T_{scalar}) and water stress limitation (W_{scalar}) on photosynthesis as follows:

$$\varepsilon_g = \varepsilon_0 \times T_{scalar} \times W_{scalar} \tag{9}$$

ε_0 is a biome-specific parameter and differs for C3 and C4 plants. The ε_0 values are obtained from a lookup-table (LUT) using the MODIS land cover data. T_{scalar} is estimated from the equation used in the Terrestrial Ecosystem Model (TEM) (Raich et al., 1991).

$$T_{scaler} = \frac{(T - T_{max}) \times (T - T_{min})}{(T - T_{max}) \times (T - T_{min}) - (T - T_{opt})^2} \tag{10}$$

where T_{min}, T_{max} and T_{opt} are the minimum, maximum, and optimum temperatures for vegetation photosynthesis, respectively. These parameters are biome specific and are also obtained from the LUT. The limitation from water stress is estimated from LSWI:

$$W_{scalar} = \frac{1 + LSWI}{1 + LSWI_{max}} \tag{11}$$

$LSWI_{max}$ is the maximum LSWI during the growing season over several years. We delineate the $LSWI_{max}$ for plant growing season from the following steps: (1) during the growing season period pre-defined by the LST, $LSWI_{max}$ is retrieved as the yearly maximum LSWI. If temperature-based identification of the growing season fails in the boreal region where nighttime temperature is always below 10°C, the growing season is set to be June to August. (2) LSWI will have an abnormally high value if snow exists and a lower value during drought periods. To eliminate these abnormal values and take the land cover change into consideration, we further calculate the $LSWI_{max}$ using a moving-window statistical algorithm: we select a window of five years and pick the second largest maximum LSWI in this period.

3. Results

3.1. Seasonal dynamics of GPP at individual flux tower sites

Fig. 3 shows the seasonal dynamics and interannual variations of GPP_{EC} and GPP_{VPM} across the 39 flux tower sites. The VPM accurately predicts the seasonality and magnitude of GPP for most natural vegetation (vegetation types other than cropland and cropland/natural vegetation mosaic in IGBP classification) (Fig. 3). Table 1 summarizes the correlation between GPP_{EC} and GPP_{VPM} at individual sites over years. Nearly two thirds of the natural biomes sites have a RMSE <1.5 g C m^{-2} day^{-1}. Cropland sites have slightly larger RMSE values of 2.20–3.06 g C m^{-2} day^{-1}.

Fig. 4 shows the comparison between GPP_{EC} and GPP_{VPM} at biome levels. When compared to GPP_{EC}, GPP_{VPM} underestimate by 4% (according to regression slope and hereafter) for deciduous broadleaf forests (DBF), 8% for mixed forests (MF), and 16% for evergreen needleleaf forests (ENF). GPP_{VPM} and GPP_{EC} agree well for closed shrubland (2%) and open shrubland (4%). For grassland and woody savannas (WSA), the biases are <8%. When all natural biome sites are combined, GPP_{VPM} is

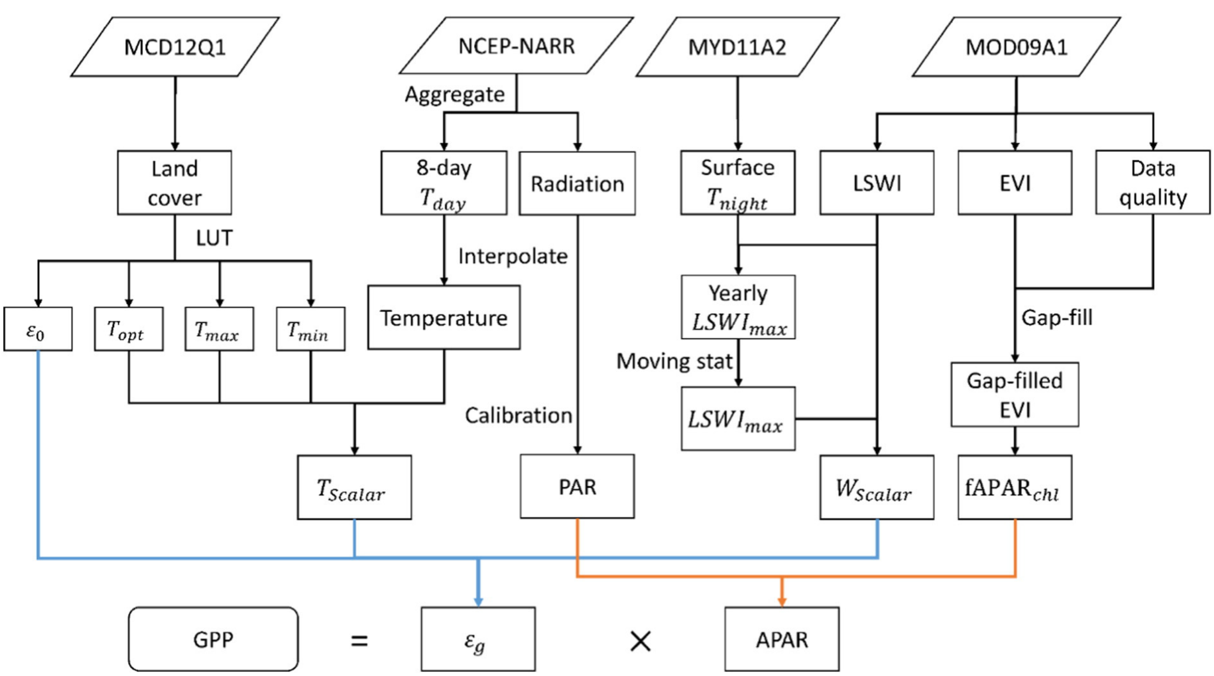

Fig. 2. Flowchart of the data processing procedures for vegetation photosynthesis model (VPM).

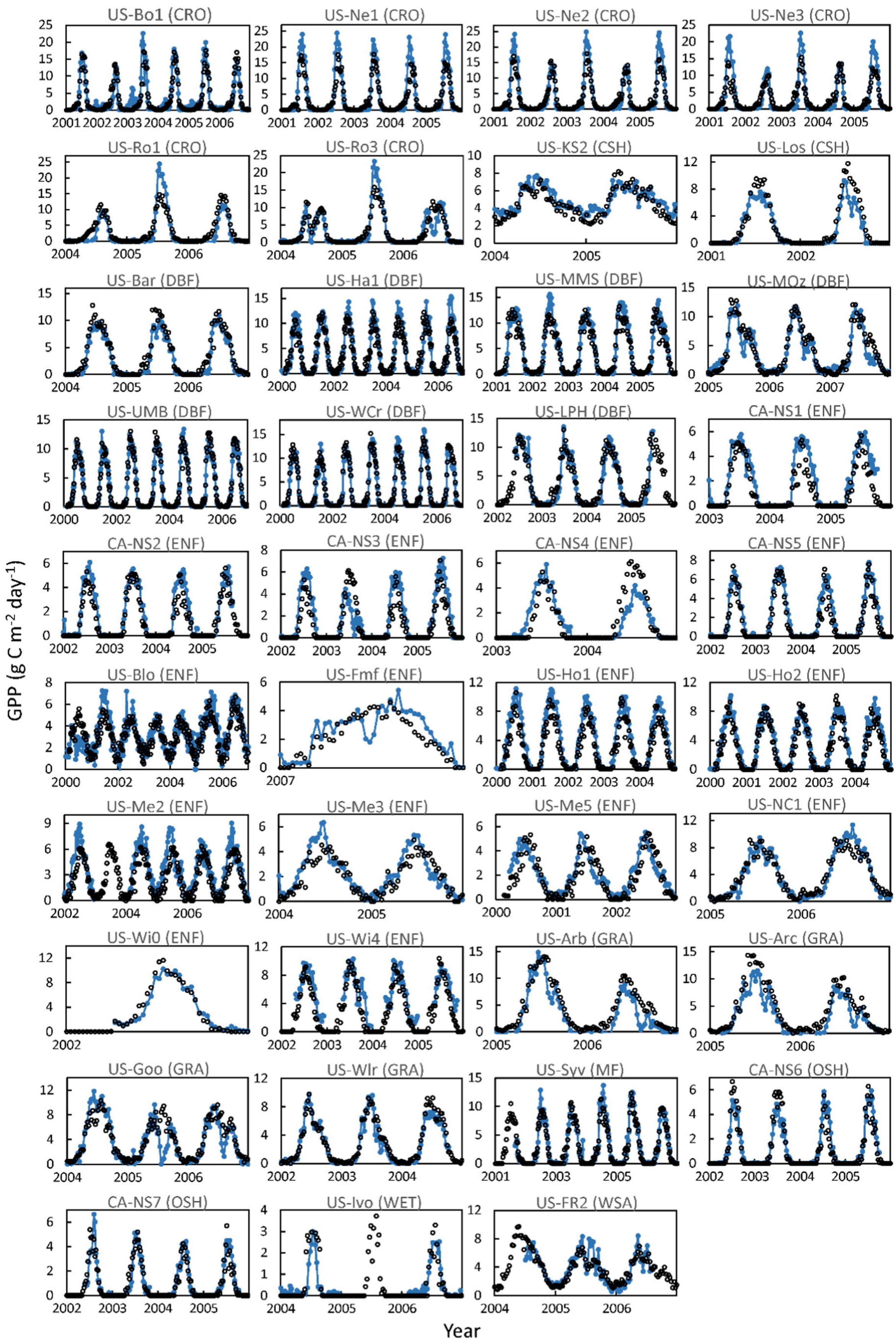

Fig. 3. Seasonal dynamics and interannual variations of the tower-based (GPP_{EC}) and the modeled (GPP_{VPM}) gross primary production at 39 flux sites at 8-day intervals. The blue lines represent the GPP_{EC} and the black circles represent the GPP_{VPM}. The ticks on the x-axis represent the first date of the corresponding year. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

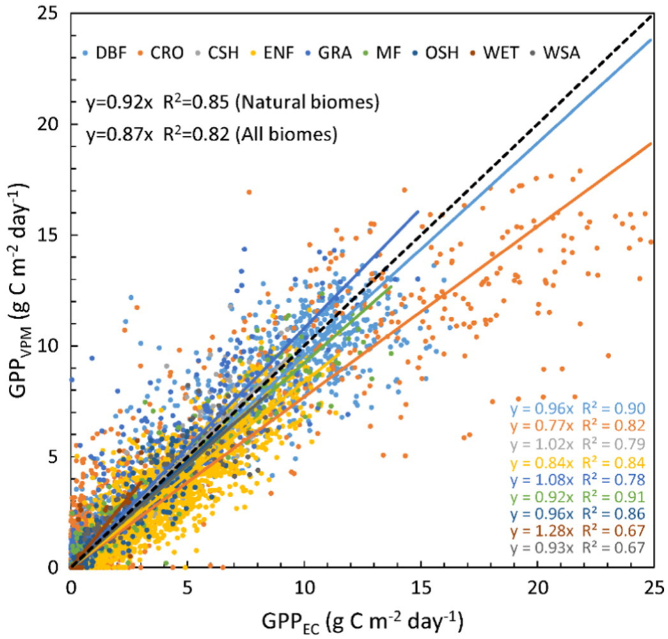

Fig. 4. A comparison of the tower-based (GPP_{EC}) and the modeled (GPP_{VPM}) gross primary production by biome types. Data are pooled across the study period for each biome. The dash line is 1:1 line and solid lines are linear regression lines forced to pass the origin.

slightly lower than GPP_{EC}, approximately 8% ($y = 0.92x$, $R^2 = 0.85$) (Fig. 4). For cropland sites (cropland and cropland/natural vegetation mosaic in IGBP classification), GPP_{VPM} is lower than GPP_{EC} by 23% ($y = 0.77x$, $R^2 = 0.82$). When all 39 sites are lumped together, the difference between GPP_{VPM} and GPP_{EC} is approximately 13% ($y = 0.87x$, $R^2 = 0.82$). The LUE parameter in VPM improves the predictability of GPP, as represented by the decreased coefficient of determination (R^2) in the VPM model sensitivity analysis for both natural biomes and all biomes sites when LUE parameter is removed (Fig. A1).

3.2. Spatial patterns of GPP_{VPM} across North America in 2010 at 500-m spatial resolution

Fig. 5A shows the spatial distribution of annual GPP_{VPM} for 2010 across NA. The highest GPP_{VPM} (>2000 g C m^{-2} $year^{-1}$) occurs in the southernmost tropical regions. GPP_{VPM} decreases along a latitudinal gradient in the eastern region, owing to the decreasing temperature and growing season length. GPP_{VPM} also decreases along a longitudinal gradient from east (dominated by forest) to west (dominated by grasslands and desert). Fig. 5B shows the spatial distribution of the maximum daily GPP_{VPM} in 2010. The highest value is ~20 g C m^{-2} day^{-1} for the Midwest Corn Belt. The southeastern U.S. has a relatively low value as compared with the mid-latitude region (35°N–45°N). The biggest contrast between annual GPP_{VPM} and maximum daily GPP_{VPM} is found in the tropical and western coastal regions, where annual GPP_{VPM} is highest while the maximum daily GPP_{VPM} is moderate.

GPP_{VPM} varies significantly across biomes (Table 2). The most productive ecosystem is the evergreen broadleaf forest with an annual GPP_{VPM} of >2000 g C m^{-2} $year^{-1}$. Open shrubland and savannas are the least productive with an annual GPP_{VPM} < 375 g C m^{-2} $year^{-1}$. Grassland, savannas, and shrublands have relatively high spatial variance because of the extensive distribution and high sensitivity to soil water. All natural vegetation contribute about 70% of the total GPP_{VPM}, with an average of 600.88 g C m^{-2} $year^{-1}$. Croplands accounts for about 27% of the total GPP but with a nearly doubled photosynthetic capacity (1194.27 g C m^{-2} $year^{-1}$) compared with the mean of natural vegetation. The maximum daily GPP_{VPM} for different biomes varies from 3.59 to 12.00 g C m^{-2} day^{-1}. Croplands have the largest GPP_{VPM} magnitudes (9.94 to 12.00 g C m^{-2} day^{-1}). Forest ecosystems have a relatively higher maximum photosynthetic rate (8.79 g C m^{-2} day^{-1}) compared with other natural vegetation types (4.65 g C m^{-2} day^{-1}).

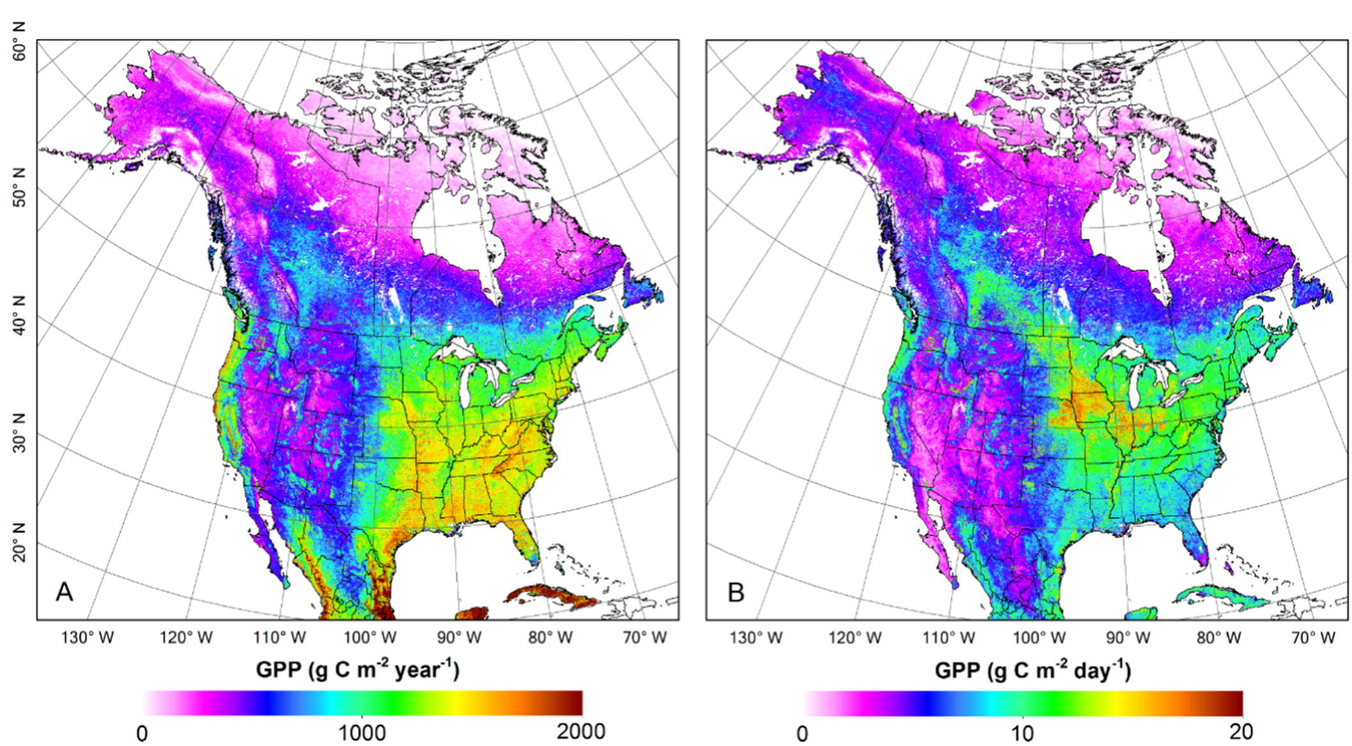

Fig. 5. Spatial distribution of modeled (A) annual GPP_{VPM} and (B) maximum daily GPP_{VPM} for year 2010.

Table 2
The magnitudes and annual sums of simulated gross primary production (GPP_{VPM}) of different biomes in North America (170°–50°W, 20°–80°N) for year 2010.

IGBP class	Average annual GPP ($g\ C\ m^{-2}\ year^{-1}$)	Standard deviation of annual GPP ($g\ C\ m^{-2}\ year^{-1}$)	Average maximum daily GPP ($g\ C\ m^{-2}\ day^{-1}$)	Standard deviation of maximum daily GPP ($g\ C\ m^{-2}\ day^{-1}$)	Total ($Pg\ C\ year^{-1}$)
ENF	638.45	255.53	5.90	1.55	1.32
EBF	2038.76	448.32	9.63	1.71	0.16
DBF	1443.95	188.49	11.09	1.47	0.75
MF	1030.24	330.46	8.53	1.78	1.94
OSH	349.30	224.44	3.59	1.31	1.48
WSA	815.81	543.79	6.27	2.29	1.50
SAV	377.65	267.02	4.17	1.27	0.20
GRA	457.50	380.74	4.24	2.59	2.00
WET	539.26	253.98	5.00	1.41	0.21
CRO	1157.99	390.54	12.00	3.09	2.15
CNV	1248.95	317.55	9.94	1.67	1.54

ENF: evergreen needleleaf forest; EBF: evergreen broadleaf forest; DBF: deciduous broadleaf forest; MF: mixed forest; OSH: open shrubland; WSA: woody savannas; SAV: savannas; GRA: grassland; WET: wetland; CRO: cropland; CNV: cropland/natural vegetation mosaic.

The inconsistency between annual GPP_{VPM} sums and maximum daily GPP_{VPM} may be mainly attributed to different growing season lengths that are affected by temperatures and rainfall.

Fig. 6 shows the frequency distribution of annual GPP_{VPM} and maximum daily GPP_{VPM} for all pixels in NA and their distribution in the climate space. >70% of pixels have relatively low productivity, i.e., annual GPP_{VPM} < 1000 g C m^{-2} $year^{-1}$ or maximum daily GPP_{VPM} < 10 g C m^{-2} day^{-1}. We also plot the distribution of the 39 flux tower sites in NA based on the annual and maximum daily GPP_{EC} (Fig. 6). The distribution of the flux tower sites cover the broad range of maximum daily GPP_{VPM}, and most of them are located in regions with moderate annual GPP (1000–1800 g C m^{-2} $year^{-1}$). In the two-dimensional climate space described by mean annual temperature (MAT) and mean annual precipitation (MAP) (Fig. 6C, D), the flux tower sites distribution covers most of the climate space. The annual GPP_{VPM} generally increases with MAT mad MAP, while the daily maximum GPP_{VPM} is highest in moderate MAT and MAP regions.

3.3. Spatial-temporal comparison between GPP_{VPM} and SIF across NA in 2010 at 0.5 degree spatial resolution

We aggregate the 8-day 500-m GPP_{VPM} estimates to the seasonal (3-month interval) and 0.5° latitude/longitude grid to compare with

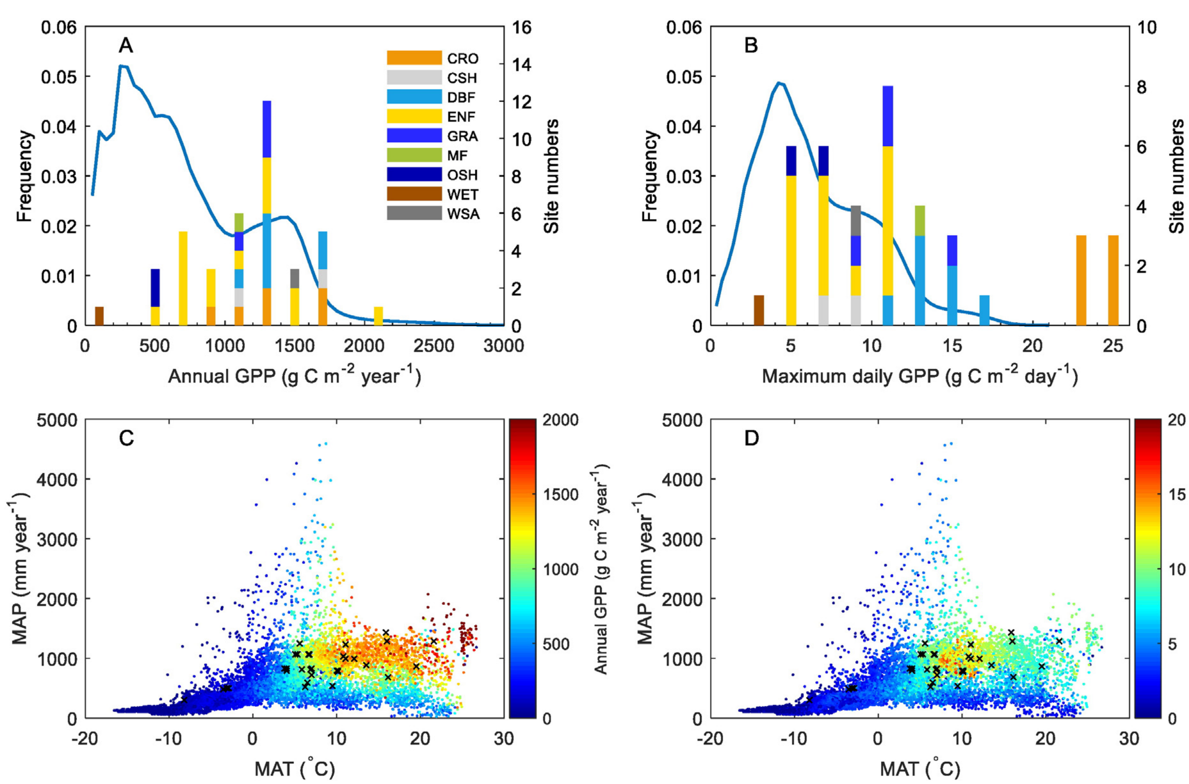

Fig. 6. The frequency distribution of GPP_{VPM} of the (A) annual GPP and (B) maximum daily GPP compared to the flux site distribution and their distribution in the climate space defined by mean annual temperature (MAT) mean annual precipitation (MAP) (C, D). The blue curves in (A and B) indicate the frequency distribution calculated from Fig. 5. The annual and maximum daily GPP for the flux tower sites are from the 39 sites used in our study. Black crosses in (C and D) represent the location of 39 flux tower sites in the climate space. Precipitation data from GPCC (Global Precipitation Climatology Centre) and temperature data from NCEP-NARR are used to generate the climate space. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

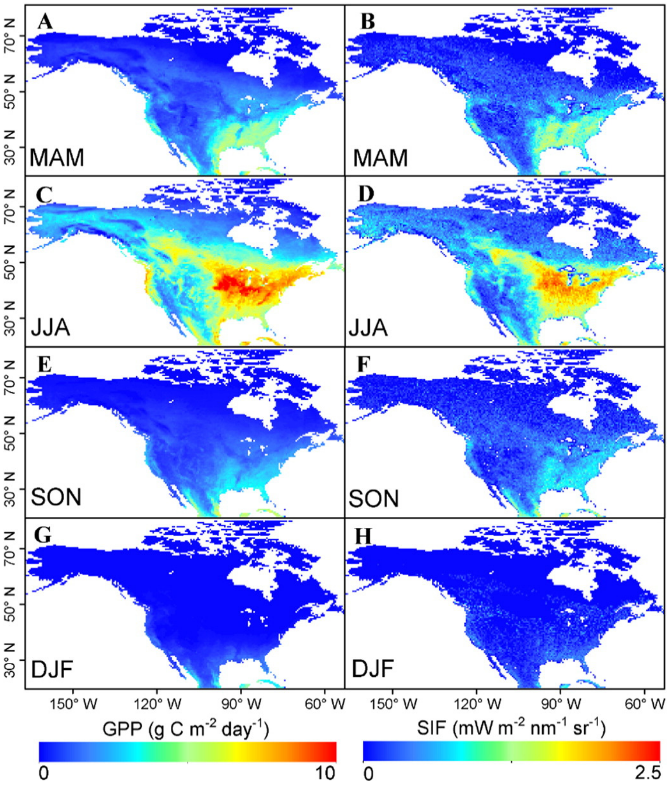

Fig. 7. A comparison of seasonal average sun-induced fluorescence (SIF) from the GOME-2 satellite instrument and simulated gross primary production (GPP_{VPM}) during the period of March 2010 through February 2011. MAM, JJA, SON, and DJF correspond to spring, summer, fall, and winter, respectively.

the seasonal SIF data. Both GPP_{VPM} and GOME-2 SIF data have strong seasonal dynamics and spatial variation across NA (Figs. 7, 8).

During spring (March to May), both GPP_{VPM} and GOME-2 SIF are relatively high in the southeastern part of the United States (Fig. 7), where forests dominate and plants grow through the spring. Both GPP_{VPM} and GOME-2 SIF are also high in California, where the Mediterranean climate (warm and wet spring and dry summer) is located (Ma, Baldocchi, Xu, & Hehn, 2007; Xu & Baldocchi, 2004). In comparison, the rest of lands with low temperature and/or rainfall in NA have low GPP_{VPM} and GOME-2 SIF values.

In summer months (June to August), the Corn Belt in midwest U.S. and southwestern Canada has the highest GPP_{VPM} and SIF. This is supported by the eddy flux data: GPP_{EC} for maize is >25 g C m^{-2} day^{-1} during summer, much higher than that of the forest ecosystems. Overall, summer months contribute >62% of the annual GPP in NA, 42% of which come from Canada and 45% from the conterminous U.S. SIF data also show the highest values in the Corn Belt and lowest in the western and northern regions, consistent with the GPP_{VPM}.

In the fall (September to November), both GPP_{VPM} and SIF drop substantially in the mid-west region due to crop harvesting. Similar to spring, the high photosynthesis rate also corresponds to a long growing season in the southeastern U.S., but the value is smaller than spring. The eastern and western coasts of Mexico as well as Cuba still fix carbon at a rate of >5 g C m^{-2} day^{-1}. In Alaska and northern Canada, all vegetation goes to dormancy, and both GPP_{VPM} and SIF values are close to 0. These spatial patterns are also evident in the SIF data.

During the winter (December through February), only the very southern part of the U.S., California, and coastal regions of Mexico and Cuba have moderate GPP_{VPM} and SIF values. All the other regions do not show any sign of photosynthesis activities, and both GPP_{VPM} and SIF values are close to zero.

4. Discussion

4.1. The relationship between SIF and GPP

SIF is emitted during the vegetation photosynthetic process. Absorbed energy by chlorophyll is partitioned into SIF, photochemical quenching (PQ, energy used for photosynthesis), non-photochemical quenching (NPQ, energy partitioned to heating), and efficiency loss (Baker, 2008). Previous studies have shown that SIF is positively correlated with PQ when light is moderate or high or environmental stress exists (Flexas, Briantais, Cerovic, & Medrano, 2000; Lee et al., 2015; Porcar-Castell, Bäck, Juurola, & Hari, 2006; Soukupová, Cséfalvay, Urban, & Košvancová, 2008). However, the relationship between GPP and SIF emission at far-red peak (SIF_{740} used in our study) is also affected by the SIF contribution from photosystem II and photosystem I, alternative sinks of energy, photorespiration, internal CO_2 concentration of leaves and enzyme activities, etc. (Porcar-Castell et al., 2014). Although SIF measurements from satellite provide a direct and independent estimations of photosynthetic activity which is different from reflectance based vegetation indices, the GPP-SIF relationship still needs intensive investigation.

Several studies (Joiner et al., 2014; Wagle, Zhang, Jin, & Xiao, 2016; Zhang et al., 2014) have reported on the direct comparison between satellite-derived SIF data (0.5° grid cell) and *in situ* GPP_{EC} from flux sites that often have footprint sizes of a few hundreds of meters, but such comparisons is problematic owing to spatial mismatches and heterogeneity due to mixed land cover types within a given 0.5° grid cell (Zhang et al., 2014). In this study, the VPM simulations are aggregated to the same spatial resolution as the GOME-2 SIF data. Fig. 8 shows the correlation between GPP_{VPM} and the SIF data for the four seasons. In spring, summer, and fall, GPP_{VPM} shows a very high correlation with SIF. The coefficient of determination ranges from 0.74 to 0.86, and the GPP_{VPM}-SIF correlation increases with the increase in daily GPP or SIF value (from early to peak growing season). This high spatial correlation confirms our comparison in Section 3.3 and can be further explained by the $APAR_{chl}$ used in the VPM. Both $APAR_{NDVI}$ (NDVI × PAR) and $APAR_{fPAR}$ (fPAR × PAR) have lower correlation with SIF compared with $APAR_{chl}$; an obvious saturation can be found in summer where SIF continues to increase while $APAR_{NDVI}$ and $APAR_{fPAR}$ tend to saturate. The regression slope between $APAR_{chl}$ and SIF are also more stable during the growing season (2.82 ± 0.13). As SIF is reemitted from the photosystem II, the higher correlation between SIF and $APAR_{chl}$ also suggests that EVI can be a good proxy of light absorbed by chlorophyll. In the winter, however, the correlations between SIF and GPP_{VPM} and APAR are much weaker mostly due to the very low SIF signal and relatively lower signal-to-noise ratio. We also calculate the regression between GPP_{VPM} and SIF for points with $GPP_{VPM} > 1$ g C m^{-2} day^{-1} (to eliminate some low values with relatively higher bias during the non-growing season). The range of the regression slopes are narrower when only data for the period of $GPP_{VPM} > 1$ g C m^{-2} day^{-1} are used as compared to all data points ($SD_{slope} = 0.42$ vs. 0.74).

4.2. Comparison of SIF and GPP estimates in North America from several models

A number of models have reported annual total GPP in NA (Huntzinger et al., 2012; Xiao et al., 2014). The annual GPP_{VPM} is 13.53 Pg C in 2010. We further compared GPP_{VPM} with GPP from six other models (MODIS PSN, MPI-BGC, LPJ, LPJ-GUESS, ORCHIDEE, and VEGAS) (Fig. 9). The VPM-based GPP estimates are close to the average

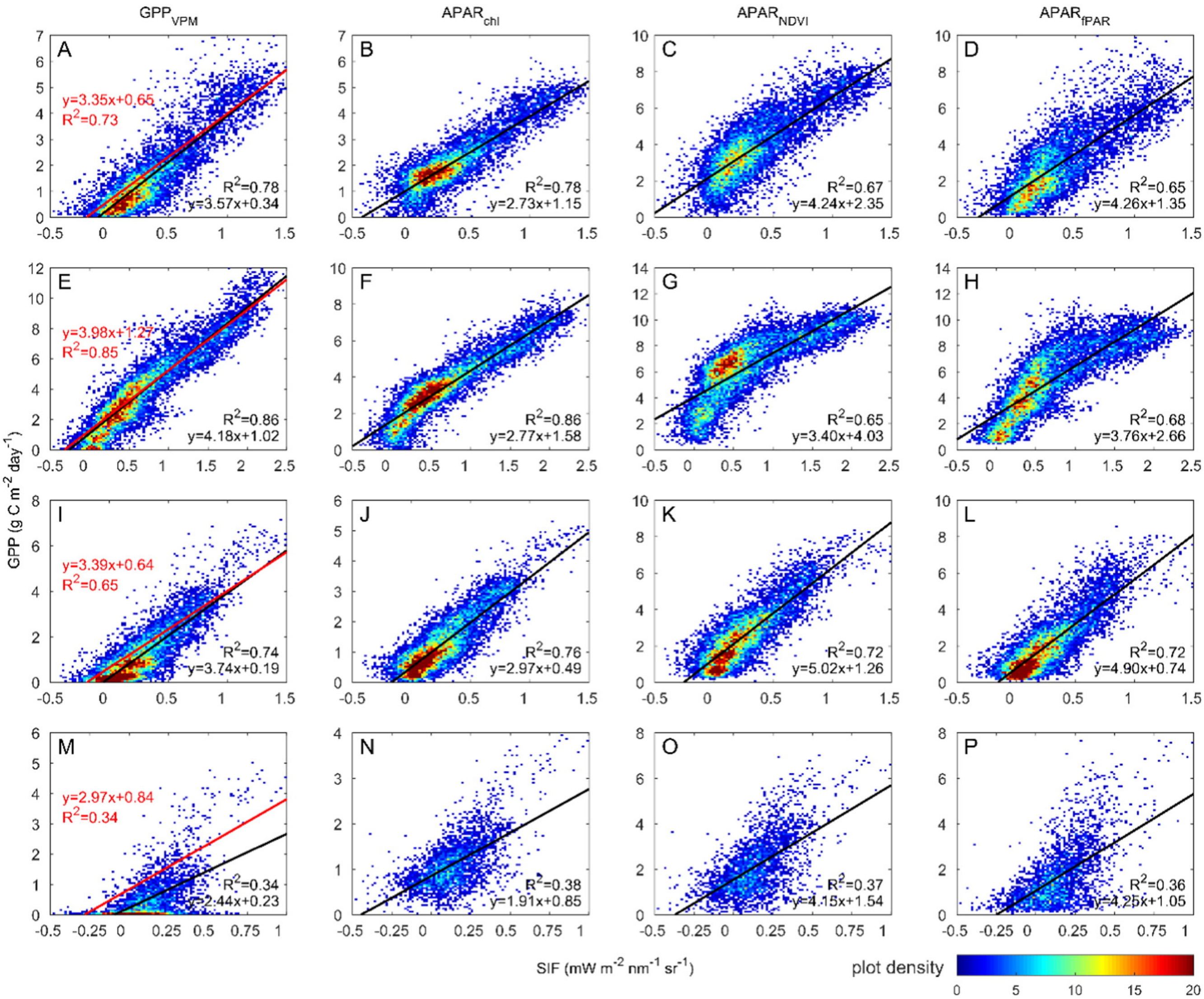

Fig. 8. Relationship between SIF and GPP_{VPM} (A, E, I, M), $APAR_{chl}$ (EVI × PAR) (B, F, J, N), $APAR_{NDVI}$ (NDVI × PAR) (C, G, K, O) and $APAR_{fPAR}$ (fPAR × PAR) (D, H, L, P) for four seasons (by row from first to fourth: spring, summer, autumn, winter) in North America in 2010. EVI and NDVI are from monthly 0.05° MOD13C1 C5, fPAR is from 8-day 1 km MOD15A2 C5, all of which are aggregated to seasonal and 0.5-degree spatial resolution. Black lines are regression for all the points, and the red lines are the regressions between GPP_{VPM} and SIF with $GPP_{VPM} > 1$ g C m^{-2} day^{-1}.

of these six models (15.75 Pg C $year^{-1}$) (Table 3). Three process-based models (LPJ, LPJ-GUESS, and ORCHIDEE) predict very high GPP for the southeastern U.S., which may be caused by different approaches they employed (enzyme kinetic vs. LUE).

Because SIF is directly retrieved from satellite and has a very good correlation with data driven model-based GPP (Frankenberg et al., 2011; Wagle et al., 2016), we use SIF as a reference to compare the spatial variations in GPP of all models. ORCHIDEE, PSN, MPI-BGC, and VPM show high consistency with SIF data. The major difference is the relative underestimation at the Corn-Belt and overestimation in the western coast along the U.S./Canada border in ORCHIDEE, PSN, and MPI-BGC. Recent studies reveal that cropland, especially maize in the U.S., makes a large contribution to the seasonal swing of atmospheric CO_2 concentration (Gray et al., 2014; Zeng et al., 2014). The high GPP values in this region are often underestimated by models (Guanter et al., 2014). Beer et al. (2010) also suggest that given the limited C4 vegetation flux data availability, great uncertainty remains in estimating the contribution of C4 plants while upscaling eddy flux observations. A similar issue is also found in a study focused on the conterminous U.S. (Xiao et al., 2010), which may explain the underestimation of the regional GPP sums. GPP_{VPM} and SIF data show similar spatial patterns for the mid-western Corn Belt ($r = 0.87$, $p < 0.001$) where a previous study showed SIF at a monthly scale has a high correlation with GPP (Guanter et al., 2014); this also supports that the spatial variation of GPP_{VPM} for croplands is to some degree an improvement over the other six models.

Several previous studies indicate that the relationships between GPP and SIF should be different across biomes (Damm et al., 2015; Guanter et al., 2012; Guanter et al., 2014; Parazoo et al., 2014; Verrelst et al., 2015). This ecosystem-dependent GPP-SIF relationship is determined by different SIF contribution from both photosystem I and photosystem II, uncertainty in NPQ, and structural interference of SIF leaving the canopy (Damm et al., 2015; Verrelst et al., 2015). Here we compare SIF with GPP estimates from three diagnostic models (VPM, MPI-BGC, and MODIS PSN) and $APAR_{chl}$, as well as the relationship between SIF_{yield} (SIF/$APAR_{chl}$) and LUE (Fig. 10). Being consistent with a previous study

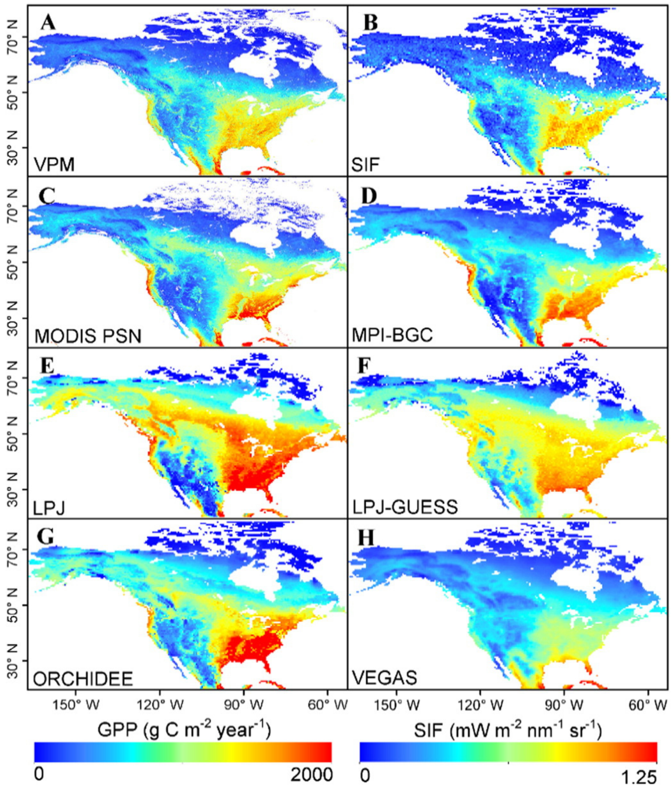

Fig. 9. Comparison of annual gross primary production (GPP) from different LUE-based models (A, C), data-driven model (D), process-based models (E, F, G, H), and with sun-induced fluorescence (SIF) (B). Data are shown for the year 2010.

at site level (Yang et al., 2015), we also find that SIF contains the information of LUE, represented by a high correlation between SIF_{yield} ($SIF/APAR_{chl}$) and LUE_{VPM} (Fig. 10E). This also partially supports the GPP-SIF relationship. However, due to the spatial inconsistency, we did not directly compare GOME-2 SIF_{yield} with LUE_{EC}, more canopy or ecosystem level SIF measurement from *in situ* or airborne spectrometers will enable this kind of comparison in the near future. In terms of inter-model comparison, VPM and MPI-BGC show higher average R^2 (0.86 and 0.89, respectively) for individual biomes than does MODIS PSN (0.83). The data points are also more scattered in the MODIS PSN than in other two models. Different biome types also show distinct differences in slopes (4.03–8.9 for VPM, 3.73–7.83 for MPI-BGC, and 2.76–11.12 for MODIS PSN). For the most highly productive biomes (average SIF > 1 mW m^{-2} nm^{-1} sr^{-1}), the correlations between predicted GPP and SIF are very high ($R^2 > 0.95$) except for EBF; this may be caused by cloud and/or aerosol contamination of the satellite data. The range of slopes for these biomes also shows less variation (4.60–5.55 for VPM, 4.02–5.72 for MPI-BGC, and 3.60–6.02 for MODIS PSN). In contrast, the less productive regions usually have lower regression coefficients and more variable slopes. This may be partially due to the higher relative error for the GOME-2 SIF data (Joiner et al., 2013) and GPP models. SIF retrievals from later satellites (OCO-2, FLEX - Fluorescence Explorer, Sentinel-5 Precursor) will have better accuracy (Frankenberg et al., 2014; Guanter et al., 2015; Kraft et al., 2013) and can be used to improve and benchmark GPP for land models (Lee et al., 2015; Luo et al., 2012; Zhang et al., 2014).

Table 3
Annual gross primary production (GPP) of North America (170°–50°W, 20°–80°N) estimated from different models for year 2010.

Models	Annual GPP (Pg C $year^{-1}$)	Reference
LPJ	22.23	Sitch et al. (2003)
LPJ-GUESS	19.84	Smith, Prentice, and Sykes (2001)
ORCHIDEE	17.52	Krinner et al. (2005)
VEGAS	11.35	Zeng et al. (2005)
MODIS GPP	13.13	Zhao et al. (2005)
MPI-BGC	12.70	Jung et al. (2011)
VPM	13.53	This study

4.3. Sources of uncertainty for VPM simulations in North America

Maps of land cover types affect GPP estimates as the LUE parameter used in the model varies with biomes. In this study, the MOD12 land cover dataset lists croplands as one category and does not distinguish between C3 and C4 crops. Both C3 and C4 crops have different photosynthetic pathways and light use efficiency (Kalfas et al., 2011; Yuan et al., 2015): C4 crops (e.g., maize) have a higher GPP_{EC} than do C3 crops (Fig. 3). Thus, the LUE parameterization of croplands for each year depends upon our knowledge of crop types and rotation. For VPM simulations at the continental scale, there are four options to address this problem in a MODIS cropland pixel: (1) assume 100% C3 plants, (2) assume 100% C4 plants, (3) assume C3 + C4 mixing ratio as 50% each, and (4) use known C3 + C4 mixing ratio from other data sources (*in situ* data, or other maps). Because there is no yearly map of C3/C4 mixing ratio across NA, we simply chose the third option in this study. Therefore, GPP_{VPM} would either overestimate GPP for C3 plants (soybean, wheat, etc.) or underestimate for C4 plants (corn, sugar cane, etc.) in those pure pixels. In those C3/C4 mixed pixels, however, these artifacts (under- or overestimation) can be partially alleviated. For example, both maize and soybean are grown in rotation at the US-Bo1 site within a 50 m radius, but within a 500 m radius of the flux tower site, corn and soybean areas have a mixing ratio of 50% each over the years. The GPP_{VPM}, driven by averaged LUE for C3 and C4 crops, captures both the seasonality and the magnitude at this site (Fig. 11A). For pure pixels, VPM would provide better results if a specific crop type is given and an appropriate LUE value is used. We use the LUE value for C4 plants at the US-Ne1 site where maize is grown throughout the period (Fig. 11B). This modification greatly improves the estimation of GPP, with an RMSE reduces from 3.06 to 2.32 g C m^{-2} day^{-1} and the slope increases from 0.65 to 0.86.

In our study, all cropland flux tower sites are located in the mid-west Corn Belt and altogether we have 16 corn years and 11 soybean years. As we use an average LUE of C3 and C4 for croplands, the model may underestimate GPP at the site scale owing to more corn years (Fig. 4). At a regional scale, the bias mainly depends on the C3 and C4 crop mixing ratios within individual pixels. In the U.S. Midwest where C4 crops (e.g., maize) are dominant, the VPM simulation may underestimate cropland production while in California or the Mississippi River Basin, where C3 crops are dominant, the VPM simulation may overestimate. Therefore, the lack of crop plant functional type (C3 and C4) is likely the largest source of uncertainty in the GPP_{VPM}. This clearly highlights the need to generate annual maps of plant functional types (C3 and C4) in NA in the near future. In addition, the mismatch between the flux tower footprint and the MODIS pixel, and the land cover fragmentation within each MODIS pixel are also critical issues when using EC data for model validation. All flux towers should be evaluated using footprint models and high resolution satellite images to provide the representativeness for the MODIS pixel (Chen et al., 2012).

Image data quality is always an important issue for the application of remote sensing. In this study, we use the vegetation indices calculated directly from the MODIS surface reflectance product. These indices are subject to atmospheric contamination (i.e., clouds, aerosols) and often result in a lower-than-normal value for EVI, especially in those regions where cloud and aerosol are persistent

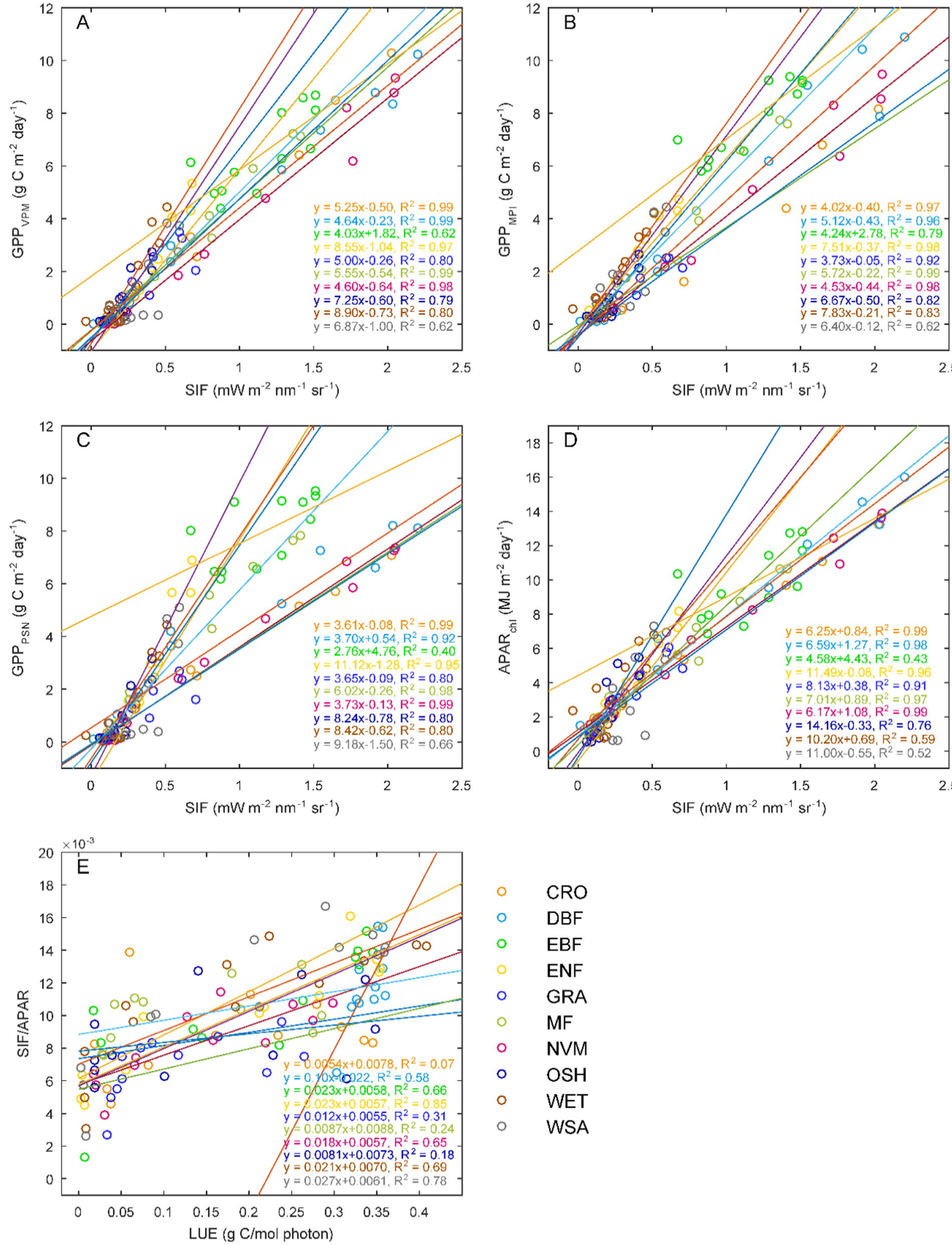

Fig. 10. A comparison for relationship between GPP_{VPM} and SIF (A), GPP_{MPI} and SIF (B), GPP_{PSN} and SIF (C), $APAR_{chl}$ (EVI × PAR) and SIF (D), SIF_{yield} (SIF/APAR) and LUE_{VPM} (E) for different biome types in North America in 2010. For each month each biome type, a value is given by spatially averaging all the grid cells with in this biome type.

(boreal and tropical regions in our study). The effect of the atmospheric contamination can be partially eliminated through a gap-fill method. Fig. 12 shows the comparison between the gap-filled and no gap-filled results. Obvious cloud contamination is marked in the black ellipse in Fig. 12A, C. The gap-fill method used in our study not only temporally interpolates the low values that are marked as cloud or aerosol contaminated by the quality control layer, but also removes the noises caused by other factors. Some

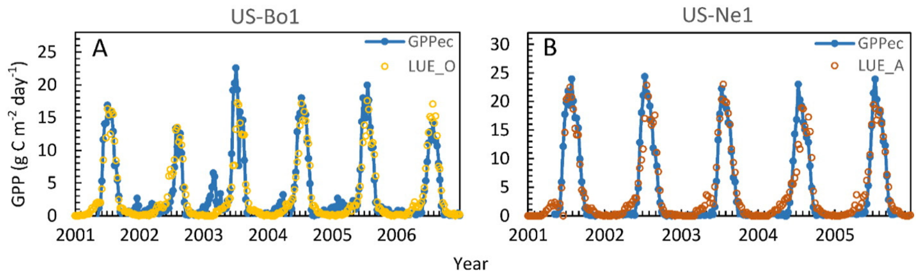

Fig. 11. Seasonal dynamics and interannual variations of the tower-based (GPP_{EC}) and the modeled (GPP_{VPM}) gross primary production at two flux tower sites at 8-day intervals at a maize/soybean rotation site (US-Bo1) (A) and a continuous maize site (US-Ne1) (B). Blue lines represent estimated GPP from flux tower, yellow circles represent the present simulation result using the original LUE (LUE_O) and brown circles represent improved simulation result using an alternative LUE (LUE_A) for C4 plant. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

extremely high value data (dark green dots) in Fig. 12A are also temporally smoothed, as shown in Fig. 12B. The use of this gap-fill method also results in different regional GPP estimates. The GPP estimate without the gap-fill method shows a total GPP of NA in 2010 as 13.23 Pg C, while the gap-filled method leads to an annual GPP estimation of 13.53 Pg C. In addition, the GPP simulations with the gap-filled processing are more stable when conducting interannual comparisons or trend analyses.

Climate data input is another potential uncertainty source for VPM simulation. Previous studies show that VPM accurately simulates GPP at flux tower sites, when driven by *in situ* (site-specific) meteorological data and parameters (Jin et al., 2013; Kalfas et al., 2011; Wagle et al., 2014; Xiao et al., 2004a; Xiao et al., 2004b). As radiation is one of the direct inputs to model GPP, the accuracy of radiation directly influences GPP simulation. Recent studies which employ different models (MODIS PSN, EC-LUE) to investigate the performance of multiple meteorological datasets in estimating regional GPP report that the NCEP product overestimates radiation as compared with meteorological stations in U.S. and China (Cai et al., 2014; Zhao et al., 2006). Jin et al. (2015) assesses the feasibility of using large scale reanalysis meteorological data (NCEP-NARR) to drive VPM at cropland flux tower sites, and the resultant GPP_{VPM} agrees well with GPP_{EC} at those sites. Our validation at the site level shows that VPM accurately simulates GPP across

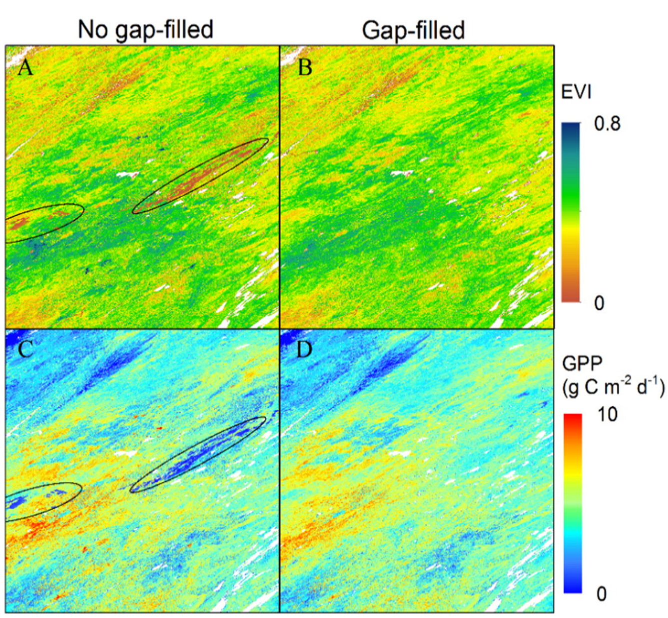

Fig. 12. Comparison between no gap-filled and gap-filled enhanced vegetation index (EVI) and the corresponding modeled gross primary production (GPP_{VPM}). The low value in (A) and (C) are marked out using ellipses. The scene is from the tile h11v03 during the mid-growing season on August 13th, 2010.

different natural biome types in NA using the regional reanalysis meteorological data and biome specific parameters, suggesting that the recalibrated NCEP-NARR radiation product can be used to estimate regional GPP effectively in NA.

5. Conclusions

In this study, we use VPM, climate reanalysis data, and MODIS products (vegetation indices, land cover, and LST) to simulate GPP of North America. GPP_{VPM} agrees well with GPP_{EC} at individual flux tower sites and the GOME-2 SIF data across North America. The comparison between SIF and GPP_{VPM} showed very high spatial-temporal consistency during the growing season, mostly due to the close relationship between SIF and $APAR_{chl}$. The quality of GOME-2 SIF data may limit its application for evaluating the seasonal variation of GPP for very low productive biome types. The results from this study clearly demonstrate the potential of VPM for estimating GPP at the continental scale, and highlights the value of GOME-2 SIF data for evaluation of various LUE-based and process-based GPP models. The resultant high spatial and temporal resolution GPP_{VPM} dataset in North America will be provided to the public, which can be further used in a wide variety of applications, especially in those studies related to trend analysis, regional disturbance evaluation, model comparison, and the carbon cycle under global climate change.

Acknowledgement

We acknowledge M. Reichstein for providing the MPI-BGC dataset, and the Numerical Terradynamic Simulation Group at the University of Montana for providing the improved MOD17 GPP dataset. We thank the TRENDY modelers for contributing model output: B. Poulter (LPJ), A. Ahlström (LPJ-GUESS), N. Viovy (ORCHIDEE), and N. Zeng (VEGAS). This study is supported in part by a research grant (Project No. 2013-69002) through the USDA National Institute for Food and Agriculture (NIFA)'s Agriculture and Food Research Initiative (AFRI), Regional Approaches for Adaptation to and Mitigation of Climate Variability and Change, and a research grant (IIA-1301789) from the National Science Foundation EPSCoR. Flux data were obtained from the AmeriFlux database (http://ameriflux.ornl.gov/). Funding for AmeriFlux data resources is provided by the U.S. Department of Energy's Office of Science. US-UMB site is supported by the Department of Energy [Award No. DE-SC0006708] and by an Ameriflux Core Site award; US-NC1 site is supported by USDA FS EFETAC cooperative agreements [03-CA-11330147-073] and [04-CA-11330147-238]; US-Ro1 and US-Ro3 data courtesy of TJ Griffis and JM Baker, funding provided by US DOE Ameriflux program and USDA-ARS; US-IVO site is supported by National Science Foundation [Award No. OPP 0421588/ARC-1204263] and DOE grant [DE-FC02-06ER64159]. We thank Ms. Sarah Xiao at Yale University for the English editing of the manuscript.

Appendix A

Table A1
Biome specific lookup-table (LUT) used in the VPM model.

IGBP class	ENF[1]	EBF[2]	DNF	DBF[1]	MF[2]	CSH[2]	OSH[2]	WSA[2]	SAV[2]	GRA[2]	WET	CRO[3]	NVM
T_{min} (°C)	−1	2	−1	−1	−1	−1	1	−1	1	0	−1	−1	0
T_{opt} (°C)	20	28	20	20	19	25	31	24	30	27	20	30	27
T_{max} (°C)	40	48	40	40	48	48	48	48	48	48	40	48	48
ε_0 (g C m^{-2} day^{-1}/W m^{-2})	0.078	0.078	0.078	0.078	0.078	0.078	0.078	0.078	0.078	0.078	0.078	0.108	0.078

ENF: evergreen needleleaf forest; EBF: evergreen broadleaf forest; DNF: deciduous needleleaf forest; DBF: deciduous broadleaf forests; MF: mixed forest; CSH: closed shrublands; OSH: open shrublands; WSA: woody savannas; SAV: savannas; GRA: grassland; WET: wetland; CRO: cropland; NVM: cropland/natural vegetation mosaic.
We use a similar temperature limitation from the Terrestrial Ecosystem Model and the T_{min}, T_{opt}, T_{max} used in this table are given by [1]Aber, Reich, and Goulden (1996), [2]McGuire et al. (1992) and [3]Wagle et al. (2015) and Kalfas et al. (2011). For some biome types (DNF, WET, NVM) which we did not find reference for temperature parameters, we use parameters from similar ecosystems (e.g. ENF for DNF and WET, GRA for NVM). ε_0 for C3 plants are estimated from the Wagle et al. (2014), ε_0 for C4 crops is from Kalfas et al. (2011). Cropland is regarded as the half-half C3/C4 therefore uses an average value.

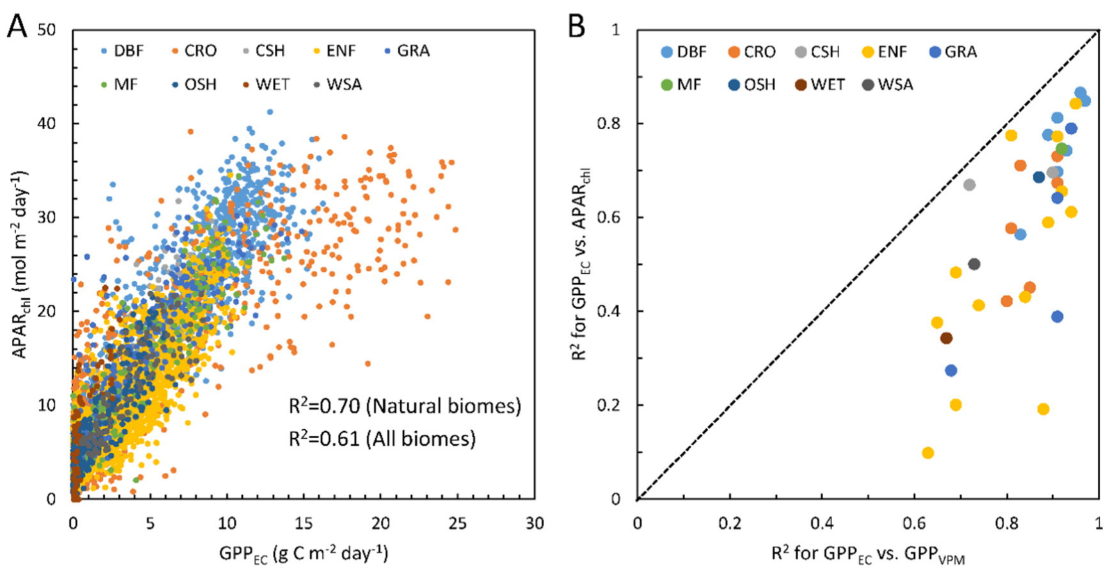

Fig. A1. (A) A comparison between GPP_{EC} and $APAR_{chl}$ for all 39 sites using 8-day data. (B) comparison between the coefficient of determination (R^2) between GPP_{EC} vs. GPP_{VPM}, and GPP_{EC} vs. $APAR_{chl}$ for individual sites.

References

Aber, J. D., Reich, P. B., & Goulden, M. L. (1996). Extrapolating leaf CO2 exchange to the canopy: A generalized model of forest photosynthesis compared with measurements by eddy correlation. *Oecologia*, *106*, 257–265.

Baker, N. R. (2008). *Chlorophyll fluorescence: A probe of photosynthesis in vivo.* Plant Biol: Annu. Rev.

Baldocchi, D. (2014). Measuring fluxes of trace gases and energy between ecosystems and the atmosphere - The state and future of the eddy covariance method. *Global Change Biology*, *20*, 3600–3609.

Baldocchi, D., Falge, E., Gu, L. H., Olson, R., Hollinger, D., Running, S., ... Wofsy, S. (2001). FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, *82*, 2415–2434.

Ballantyne, A. P., Alden, C. B., Miller, J. B., Tans, P. P., & White, J. W. (2012). Increase in observed net carbon dioxide uptake by land and oceans during the past 50 years. *Nature*, *488*, 70–72.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., ... Bonan, G. B. (2010). Terrestrial gross carbon dioxide uptake: Global distribution and covariation with climate. *Science*, *329*, 834–838.

Booth, B. B. B., Jones, C. D., Collins, M., Totterdell, I. J., Cox, P. M., Sitch, S., ... Lloyd, J. (2012). High sensitivity of future global warming to land carbon cycle processes. *Environmental Research Letters*, *7*, 024002.

Cai, W., Yuan, W., Liang, S., Zhang, X., Dong, W., Xia, J., ... Zhang, Q. (2014). Improved estimations of gross primary production using satellite-derived photosynthetically active radiation. *Journal of Geophysical Research: Biogeosciences*, *119*, 110–123.

Chen, B. Z., Coops, N. C., Fu, D., Margolis, H. A., Amiro, B. D., Black, T. A., ... Wofsy, S. C. (2012). Characterizing spatial representativeness of flux tower eddy-covariance measurements across the Canadian Carbon Program Network using remote sensing and footprint analysis. *Remote Sensing of Environment*, *124*, 742–755.

Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., & Eklundh, L. (2004). A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sensing of Environment*, *91*, 332–344.

Chen, J., Yan, H., Wang, S., Gao, Y., Huang, M., Wang, J., & Xiao, X. (2014). Estimation of gross primary productivity in Chinese terrestrial ecosystems by using VPM model. *Quaternary Sciences*, *34*.

Collatz, G. J., Ribas-Carbo, M., & Berry, J. A. (1992). Coupled photosynthesis-stomatal conductance model for leaves of C4 plants. *Australian Journal of Plant Physiology*, *19*, 519–538.

Cook, B. D., Davis, K. J., Wang, W. G., Desai, A., Berger, B. W., Teclaw, R. M., ... Heilman, W. (2004). Carbon exchange and venting anomalies in an upland deciduous forest in northern Wisconsin, USA. *Agricultural and Forest Meteorology*, *126*, 271–295.

Coops, N. C., Ferster, C. J., Waring, R. H., & Nightingale, J. (2009). Comparison of three models for predicting gross primary production across and within forested ecoregions in the contiguous United States. *Remote Sensing of Environment*, *113*, 680–690.

Coulter, R. L., Pekour, M. S., Cook, D. R., Klazura, G. E., Martin, T. J., & Lucas, J. D. (2006). Surface energy and carbon dioxide fluxes above different vegetation types within ABLE. *Agricultural and Forest Meteorology*, *136*, 147–158.

Damm, A., Guanter, L., Verhoef, W., Schläpfer, D., Garbari, S., & Schaepman, M. E. (2015). Impact of varying irradiance on vegetation indices and chlorophyll fluorescence derived from spectroscopy data. *Remote Sensing of Environment*, *156*, 202–215.

Desai, A. R., Bolstad, P. V., Cook, B. D., Davis, K. J., & Carey, E. V. (2005). Comparing net ecosystem exchange of carbon dioxide between an old-growth and mature forest in the upper Midwest, USA. *Agricultural and Forest Meteorology*, *128*, 33–55.

Dijkstra, P., Hymus, G., Colavito, D., Vieglais, D. A., Cundari, C. M., Johnson, D. P., ... Drake, B. G. (2002). Elevated atmospheric CO2 stimulates aboveground biomass in a fire-regenerated scrub-oak ecosystem. *Global Change Biology*, *8*, 90–103.

Dong, J. W., Xiao, X. M., Kou, W. L., Qin, Y. W., Zhang, G. L., Li, L., ... Moore, B. (2015b). Tracking the dynamics of paddy rice planting area in 1986–2010 through time series Landsat images and phenology-based algorithms. *Remote Sensing of Environment*, *160*, 99–113.

Dong, J., Xiao, X., Wagle, P., Zhang, G., Zhou, Y., Jin, C., ... Moore Iii, B. (2015a). Comparison of four EVI-based models for estimating gross primary production of maize and soybean croplands and tallgrass prairie under severe drought. *Remote Sensing of Environment*, *162*, 154–168.

Dore, S., Kolb, T. E., Montes-Helu, M., Sullivan, B. W., Winslow, W. D., Hart, S. C., ... Hungate, B. A. (2008). Long-term impact of a stand-replacing fire on ecosystem CO(2) exchange of a ponderosa pine forest. *Global Change Biology*, *14*, 1801–1820.

Epstein, H. E., Calef, M. P., Walker, M. D., Chapin, F. S., & Starfield, A. M. (2004). Detecting changes in arctic tundra plant communities in response to warming over decadal time scales. *Global Change Biology*, *10*, 1325–1334.

Farquhar, G. D., Caemmerer, S. V., & Berry, J. A. (1980). A biochemical-model of photosynthetic Co2 assimilation in leaves of C-3 species. *Planta*, *149*, 78–90.

Fischer, M. L., Billesbach, D. P., Berry, J. A., Riley, W. J., & Torn, M. S. (2007). Spatiotemporal variations in growing season exchanges of CO2, H2O, and sensible heat in agricultural fields of the southern Great Plains. *Earth Interactions*, *11*.

Flexas, J., Briantais, J. M., Cerovic, Z., & Medrano, H. (2000). Steady-state and maximum chlorophyll fluorescence responses to water stress in grapevine leaves: A new remote sensing system. *Remote Sensing of Environment*, *73*, 283–297.

Forkel, M., Carvalhais, N., Rodenbeck, C., Keeling, R., Heimann, M., Thonicke, K., ... Reichstein, M. (2016). Enhanced seasonal CO2 exchange caused by amplified plant productivity in northern ecosystems. *Science*, *351*, 696–699.

Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J. E., ... Yokota, T. (2011). New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophysical Research Letters*, *38*.

Frankenberg, C., O'Dell, C., Berry, J., Guanter, L., Joiner, J., Köhler, P., ... Taylor, T. E. (2014). Prospects for chlorophyll fluorescence remote sensing from the Orbiting Carbon Observatory-2. *Remote Sensing of Environment*, *147*, 1–12.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. M. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, *114*, 168–182.

Gitelson, A. A., Vina, A., Verma, S. B., Rundquist, D. C., Arkebauer, T. J., Keydan, G., ... Suyker, A. E. (2006). Relationship between gross primary production and chlorophyll content in crops: Implications for the synoptic monitoring of vegetation productivity. *Journal of Geophysical Research-Atmospheres*, *111*.

Goldstein, A. H., Hultman, N. E., Fracheboud, J. M., Bauer, M. R., Panek, J. A., Xu, M., ... Baugh, W. (2000). Effects of climate variability on the carbon dioxide, water, and sensible heat fluxes above a ponderosa pine plantation in the Sierra Nevada (CA). *Agricultural and Forest Meteorology*, *101*, 113–129.

Gough, C. M., Vogel, C. S., Schmid, H. P., Su, H. B., & Curtis, P. S. (2008). Multi-year convergence of biometric and meteorological estimates of forest carbon storage. *Agricultural and Forest Meteorology*, *148*, 158–170.

Goulden, M. L., Winston, G. C., McMillan, A. M. S., Litvak, M. E., Read, E. L., Rocha, A. V., & Elliot, J. R. (2006). An eddy covariance mesonet to measure the effect of forest age on land-atmosphere exchange. *Global Change Biology*, *12*, 2146–2162.

Graven, H. D., Keeling, R. F., Piper, S. C., Patra, P. K., Stephens, B. B., Wofsy, S. C., ... Bent, J. D. (2013). Enhanced seasonal exchange of CO2 by northern ecosystems since 1960. *Science*, *341*, 1085–1089.

Gray, J. M., Frolking, S., Kort, E. A., Ray, D. K., Kucharik, C. J., Ramankutty, N., & Friedl, M. A. (2014). Direct human influence on atmospheric CO2 seasonality from increased cropland productivity. *Nature*, *515*, 398–401.

Griffis, T. J., Baker, J. M., & Zhang, J. (2005). Seasonal dynamics and partitioning of isotopic CO2 exchange in C-3/C-4 managed ecosystem. *Agricultural and Forest Meteorology*, *132*, 1–19.

Gu, L. H., Meyers, T., Pallardy, S. G., Hanson, P. J., Yang, B., Heuer, M., ... Wullschleger, S. D. (2006). Direct and indirect effects of atmospheric conditions and soil moisture on surface energy partitioning revealed by a prolonged drought at a temperate forest site. *Journal of Geophysical Research-Atmospheres*, 111.

Guan, K., Berry, J. A., Zhang, Y., Joiner, J., Guanter, L., Badgley, G., & Lobell, D. B. (2015). Improving the monitoring of crop productivity using spaceborne solar-induced fluorescence. *Global Change Biology*. http://dx.doi.org/10.1111/gcb.13136.

Guanter, L., Aben, I., Tol, P., Krijger, J. M., Hollstein, A., Kohler, P., ... Landgraf, J. (2015). Potential of the TROPOspheric Monitoring Instrument (TROPOMI) onboard the Sentinel-5 Precursor for the monitoring of terrestrial chlorophyll fluorescence. *Atmospheric Measurement Techniques*, *8*, 1337–1352.

Guanter, L., Frankenberg, C., Dudhia, A., Lewis, P. E., Gomez-Dans, J., Kuze, A., ... Grainger, R. G. (2012). Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements. *Remote Sensing of Environment*, *121*, 236–251.

Guanter, L., Rossini, M., Colombo, R., Meroni, M., Frankenberg, C., Lee, J. E., & Joiner, J. (2013). Using field spectroscopy to assess the potential of statistical approaches for the retrieval of sun-induced chlorophyll fluorescence from ground and space. *Remote Sensing of Environment*, *133*, 52–61.

Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J. A., ... Griffis, T. J. (2014). Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, E1327–E1333.

He, H. L., Liu, M., Xiao, X. M., Ren, X. L., Zhang, L., Sun, X. M., ... Yu, G. R. (2014). Large-scale estimation and uncertainty analysis of gross primary production in Tibetan alpine grasslands. *Journal of Geophysical Research-Biogeosciences*, *119*, 466–486.

Heinsch, F. A., Heilman, J. L., McInnes, K. J., Cobos, D. R., Zuberer, D. A., & Roelke, D. L. (2004). Carbon dioxide exchange in a high marsh on the Texas Gulf Coast: Effects of freshwater availability. *Agricultural and Forest Meteorology*, *125*, 159–172.

Hollinger, D. Y., Aber, J., Dail, B., Davidson, E. A., Goltz, S. M., Hughes, H., ... Walsh, J. (2004). Spatial and temporal variability in forest-atmosphere CO2 exchange. *Global Change Biology*, *10*, 1689–1706.

Hollinger, S. E., Bernacchi, C. J., & Meyers, T. P. (2005). Carbon budget of mature no-till ecosystem in North Central Region of the United States. *Agricultural and Forest Meteorology*, *130*, 59–69.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*, 195–213.

Huntzinger, D. N., Post, W. M., Wei, Y., Michalak, A. M., West, T. O., Jacobson, A. R., ... Cook, R. (2012). North American Carbon Program (NACP) regional interim synthesis: Terrestrial biospheric model intercomparison. *Ecological Modelling*, *232*, 144–157.

Jenkins, J. P., Richardson, A. D., Braswell, B. H., Ollinger, S. V., Hollinger, D. Y., & Smith, M. L. (2007). Refining light-use efficiency calculations for a deciduous forest canopy using simultaneous tower-based carbon flux and radiometric measurements. *Agricultural and Forest Meteorology*, *143*, 64–79.

Jin, C., Xiao, X. M., Merbold, L., Arneth, A., Veenendaal, E., & Kutsch, W. L. (2013). Phenology and gross primary production of two dominant savanna woodland ecosystems in Southern Africa. *Remote Sensing of Environment*, *135*, 189–201.

Jin, C., Xiao, X., Wagle, P., Griffis, T., Dong, J., Wu, C., & Qin, Y. (2015). Effects of in-situ and reanalysis climate data on estimation of cropland gross primary production using the Vegetation Photosynthesis Model. *Agricultural and Forest Meteorology*, *213*, 240.

Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A. P., Middleton, E. M., ... Frankenberg, C. (2013). Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: Methodology, simulations, and application to GOME-2. *Atmospheric Measurement Techniques*, *6*, 2803–2823.

Joiner, J., Yoshida, Y., Vasilkov, A. P., Middleton, E. M., Campbell, P. K. E., Yoshida, Y., ... Corp, L. A. (2012). Filling-in of near-infrared solar lines by terrestrial fluorescence and other geophysical effects: Simulations and space-based observations from SCIAMACHY and GOSAT. *Atmospheric Measurement Techniques*, *5*, 809–829.

Joiner, J., Yoshida, Y., Vasilkov, A. P., Schaefer, K., Jung, M., Guanter, L., ... Belelli Marchesini, L. (2014). The seasonal cycle of satellite chlorophyll fluorescence observations and its relationship to vegetation phenology and ecosystem atmosphere carbon exchange. *Remote Sensing of Environment*, *152*, 375–391.

Joiner, J., Yoshida, Y., Vasilkov, A. P., Yoshida, Y., Corp, L. A., & Middleton, E. M. (2011). First observations of global and seasonal terrestrial chlorophyll fluorescence from space. *Biogeosciences*, *8*, 637–651.

Jung, M., Reichstein, M., & Bondeau, A. (2009). Towards global empirical upscaling of FLUXNET eddy covariance observations: Validation of a model tree ensemble approach using a biosphere model. *Biogeosciences*, *6*, 2001–2013.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., ... Williams, C. (2011). Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research-Biogeosciences*, *116*.

Kalfas, J. L., Xiao, X., Vanegas, D. X., Verma, S. B., & Suyker, A. E. (2011). Modeling gross primary production of irrigated and rain-fed maize using MODIS imagery and CO2 flux tower data. *Agricultural and Forest Meteorology*, *151*, 1514–1528.

Keenan, T. F., Baker, I., Barr, A., Ciais, P., Davis, K., Dietze, M., ... Richardson, A. D. (2012). Terrestrial biosphere model performance for inter-annual variability of land-atmosphere CO2 exchange. *Global Change Biology*, *18*, 1971–1987.

Kraft, S., Bézy, J. L., Del Bello, U., Berlich, R., Drusch, M., Franco, R., ... Silvestrin, P. (2013). FLORIS: Phase A status of the fluorescence imaging spectrometer of the Earth Explorer mission candidate FLEX. *Proc. SPIE 8889, Sensors, Systems, and Next-Generation Satellites XVII. 8889*. (pp. 88890T–888812).

Krinner, G., Viovy, N., de Noblet-Ducoudre, N., Ogee, J., Polcher, J., Friedlingstein, P., ... Prentice, I. C. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, *19*.

Law, B. E., Turner, D., Campbell, J., Sun, O. J., Van Tuyl, S., Ritts, W. D., & Cohen, W. B. (2004). Disturbance and climate effects on carbon stocks and fluxes across Western Oregon USA. *Global Change Biology*, *10*, 1429–1444.

Law, B. E., Williams, M., Anthoni, P. M., Baldocchi, D. D., & Unsworth, M. H. (2000). Measuring and modelling seasonal variation of carbon dioxide and water vapour exchange of a *Pinus ponderosa* forest subject to soil water deficit. *Global Change Biology*, *6*, 613–630.

Lee, J. E., Berry, J. A., van der Tol, C., Yang, X., Guanter, L., Damm, A., ... Frankenberg, C. (2015). Simulations of chlorophyll fluorescence incorporated into the Community Land Model version 4. *Global Change Biology*, *21*, 3469–3477. http://dx.doi.org/10.1111/gcb.12948.

Luo, Y. Q., Randerson, J. T., Abramowitz, G., Bacour, C., Blyth, E., Carvalhais, N., ... Zhou, X. H. (2012). A framework for benchmarking land models. *Biogeosciences*, *9*, 3857–3874.

Ma, S. Y., Baldocchi, D. D., Xu, L. K., & Hehn, T. (2007). Inter-annual variability in carbon dioxide exchange of an oak/grass savanna and open grassland in California. *Agricultural and Forest Meteorology*, *147*, 157–171.

McGuire, A. D., Melillo, J., Joyce, L., Kicklighter, D., Grace, A., Moore, B., & Vorosmarty, C. (1992). Interactions between carbon and nitrogen dynamics in estimating net primary productivity for potential vegetation in North America. *Global Biogeochemical Cycles*, *6*, 101–124.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., ... Shi, W. (2006). North American regional reanalysis. *Bulletin of the American Meteorological Society*, *87*, 343.

Monteith, J. L. (1972). Solar-radiation and productivity in tropical ecosystems. *Journal of Applied Ecology*, *9*, 747–766.

Noormets, A., Gavazzi, M. J., Mcnulty, S. G., Domec, J. C., Sun, G., King, J. S., & Chen, J. Q. (2010). Response of carbon fluxes to drought in a coastal plain loblolly pine forest. *Global Change Biology*, *16*, 272–287.

Parazoo, N. C., Bowman, K., Fisher, J. B., Frankenberg, C., Jones, D. B., Cescatti, A., ... Montagnani, L. (2014). Terrestrial gross primary production inferred from satellite fluorescence and vegetation models. *Global Change Biology*, *20*, 3103–3121.

Piao, S., Sitch, S., Ciais, P., Friedlingstein, P., Peylin, P., Wang, X., ... Zeng, N. (2013). Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO2 trends. *Global Change Biology*, *19*, 2117–2132.

Porcar-Castell, A., Bäck, J., Juurola, E., & Hari, P. (2006). Dynamics of the energy flow through photosystem II under changing light conditions: A model approach. *Functional Plant Biology*, *33*, 229–239. http://dx.doi.org/10.1071/FP05133.

Porcar-Castell, A., Tyystjarvi, E., Atherton, J., van der Tol, C., Flexas, J., Pfundel, E. E., ... Berry, J. A. (2014). Linking chlorophyll *a* fluorescence to photosynthesis for remote sensing applications: Mechanisms and challenges. *Journal of Experimental Botany*, *65*, 4065–4095.

Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., & Klooster, S. A. (1993). Terrestrial ecosystem production - A process model-based on global satellite and surface data. *Global Biogeochemical Cycles*, *7*, 811–841.

Poulter, B., Frank, D., Ciais, P., Myneni, R. B., Andela, N., Bi, J., ... van der Werf, G. R. (2014). Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle. *Nature*, *509*, 600–603.

Raich, J., Rastetter, E., Melillo, J., Kicklighter, D., Steudler, P., Peterson, B., ... Vorosmarty, C. (1991). Potential net primary productivity in South America: Application of a global model. *Ecological Applications*, *1*, 399–429.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., ... Valentini, R. (2005). On the separation of net ecosystem exchange into assimilation and ecosystem respiration: Review and improved algorithm. *Global Change Biology*, *11*, 1424–1439.

Rossini, M., Nedbal, L., Guanter, L., Ač, A., Alonso, L., Burkart, A., ... Rascher, U. (2015). Red and far-red sun-induced chlorophyll fluorescence as a measure of plant photosynthesis. *Geophysical Research Letters* n/a-n/a.

Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M. S., Reeves, M., & Hashimoto, H. (2004). A continuous satellite-derived measure of global terrestrial primary production. *Bioscience*, *54*, 547–560.

Schmid, H. P., Grimmond, C. S. B., Cropley, F., Offerle, B., & Su, H. B. (2000). Measurements of CO2 and energy fluxes over a mixed hardwood forest in the mid-western United States. *Agricultural and Forest Meteorology*, *103*, 357–374.

Sitch, S., Huntingford, C., Gedney, N., Levy, P. E., Lomas, M., Piao, S. L., ... Woodward, F. I. (2008). Evaluation of the terrestrial carbon cycle, future plant geography and climate–carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs). *Global Change Biology*, *14*, 2015–2039.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., ... Venevsky, S. (2003). Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, *9*, 161–185.

Smith, B., Prentice, I. C., & Sykes, M. T. (2001). Representation of vegetation dynamics in the modelling of terrestrial ecosystems: Comparing two contrasting approaches within European climate space. *Global Ecology and Biogeography*, *10*, 621–637.

Soukupová, J., Cséfalvay, L., Urban, O., Košvancová, M., Marek, M., Rascher, U., ... Nedbal, L. (2008). Annual variation of the steady-state chlorophyll fluorescence emission of evergreen plants in temperate zone. *Functional Plant Biology*, *35*, 63–76. http://dx.doi.org/10.1071/FP07158.

Sulman, B. N., Desai, A. R., Cook, B. D., Saliendra, N., & Mackay, D. S. (2009). Contrasting carbon dioxide fluxes between a drying shrub wetland in Northern Wisconsin, USA, and nearby forests. *Biogeosciences*, *6*, 1115–1126.

Sun, G., Noormets, A., Chen, J., & McNulty, S. G. (2008). Evapotranspiration estimates from eddy covariance towers and hydrologic modeling in managed forests in Northern Wisconsin, USA. *Agricultural and Forest Meteorology*, *148*, 257–267.

Suyker, A. E., Verma, S. B., Burba, G. G., & Arkebauer, T. J. (2005). Gross primary production and ecosystem respiration of irrigated maize and irrigated soybean during a growing season. *Agricultural and Forest Meteorology*, *131*, 180–190.

Urbanski, S., Barford, C., Wofsy, S., Kucharik, C., Pyle, E., Budney, J., ... Munger, J. W. (2007). Factors controlling CO2 exchange on timescales from hourly to decadal at Harvard Forest. *Journal of Geophysical Research-Biogeosciences*, *112*.

Vanderhoof, M., Williams, C., Pasay, M., & Ghimire, B. (2013). Controls on the rate of CO2 emission from woody debris in clearcut and coniferous forest environments. *Biogeochemistry*, *114*, 299–311.

Verrelst, J., Rivera, J. P., van der Tol, C., Magnani, F., Mohammed, G., & Moreno, J. (2015). Global sensitivity analysis of the SCOPE model: What drives simulated canopy-leaving sun-induced fluorescence? *Remote Sensing of Environment*, *166*, 8–21.

Wagle, P., Xiao, X. M., & Suyker, A. E. (2015). Estimation and analysis of gross primary production of soybean under various management practices and drought conditions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *99*, 70–83.

Wagle, P., Xiao, X., Torn, M. S., Cook, D. R., Matamala, R., Fischer, M. L., ... Biradar, C. (2014). Sensitivity of vegetation indices and gross primary production of tallgrass prairie to severe drought. *Remote Sensing of Environment*, *152*, 1–14.

Wagle, P., Zhang, Y., Jin, C., & Xiao, X. (2016). Comparison of solar-induced chlorophyll fluorescence, light-use efficiency, and process-based GPP models in maize. *Ecological Applications.*. http://dx.doi.org/10.1890/15-1434.1.

Wilson, T. B., & Meyers, T. P. (2007). Determining vegetation indices from solar and photosynthetically active radiation fluxes. *Agricultural and Forest Meteorology*, *144*, 160–179.

Wu, C. Y., Munger, J. W., Niu, Z., & Kuang, D. (2010a). Comparison of multiple models for estimating gross primary production using MODIS and eddy covariance data in Harvard Forest. *Remote Sensing of Environment*, *114*, 2925–2939.

Wu, C. Y., Niu, Z., & Gao, S. A. (2010b). Gross primary production estimation from MODIS data with vegetation index and photosynthetically active radiation in maize. *Journal of Geophysical Research-Atmospheres*, *115*.

Xia, J., Chen, J., Piao, S., Ciais, P., Luo, Y., & Wan, S. (2014). Terrestrial carbon cycle affected by non-uniform climate warming. *Nature Geoscience*, *7*, 173–180.

Xiao, X., Boles, S., Liu, J., Zhuang, D., & Liu, M. (2002). Characterization of forest types in Northeastern China, using multi-temporal SPOT-4 VEGETATION sensor data. *Remote Sensing of Environment*, *82*, 335–348.

Xiao, X., Hollinger, D., Aber, J., Goltz, M., Davidson, E. A., Zhang, Q., & Moore, B. (2004a). Satellite-based modeling of gross primary production in an evergreen needleleaf forest. *Remote Sensing of Environment*, *89*, 519–534.

Xiao, J. F., Ollinger, S. V., Frolking, S., Hurtt, G. C., Hollinger, D. Y., Davis, K. J., ... Suyker, A. E. (2014). Data-driven diagnostics of terrestrial carbon dynamics over North America. *Agricultural and Forest Meteorology*, *197*, 142–157.

Xiao, X., Zhang, Q., Braswell, B., Urbanski, S., Boles, S., Wofsy, S., ... Ojima, D. (2004b). Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data. *Remote Sensing of Environment*, *91*, 256–270.

Xiao, X. M., Zhang, Q. Y., Saleska, S., Hutyra, L., De Camargo, P., Wofsy, S., ... Moore, B. (2005). Satellite-based modeling of gross primary production in a seasonally moist tropical evergreen forest. *Remote Sensing of Environment*, *94*, 105–122.

Xiao, J., Zhuang, Q., Law, B. E., Chen, J., Baldocchi, D. D., Cook, D. R., ... Ma, S. (2010). A continuous measure of gross primary production for the conterminous United States derived from MODIS and AmeriFlux data. *Remote Sensing of Environment*, *114*, 576–591.

Xu, L. K., & Baldocchi, D. D. (2004). Seasonal variation in carbon dioxide exchange over a Mediterranean annual grassland in California. *Agricultural and Forest Meteorology*, *123*, 79–96.

Yang, F. H., Ichii, K., White, M. A., Hashimoto, H., Michaelis, A. R., Votava, P., ... Nemani, R. R. (2007). Developing a continental-scale measure of gross primary production by combining MODIS and AmeriFlux data through Support Vector Machine approach. *Remote Sensing of Environment*, *110*, 109–122.

Yang, X., Tang, J., Mustard, J. F., Lee, J. -E., Rossini, M., Joiner, J., ... Richardson, A. D. (2015). Solar-induced chlorophyll fluorescence correlates with canopy photosynthesis on diurnal and seasonal scales in a temperate deciduous forest. *Geophysical Research Letters* n/a-n/a.

Yu, G. R., Zhu, X. J., Fu, Y. L., He, H. L., Wang, Q. F., Wen, X. F., ... Tong, C. L. (2013). Spatial patterns and climate drivers of carbon fluxes in terrestrial ecosystems of China. *Global Change Biology*, *19*, 798–810.

Yuan, W., Cai, W., Nguy-Robertson, A. L., Fang, H., Suyker, A. E., Chen, Y., ... Zhang, H. (2015). Uncertainty in simulating gross primary production of cropland ecosystem from satellite-based models. *Agricultural and Forest Meteorology*, *207*, 48–57.

Yuan, W., Cai, W., Xia, J., Chen, J., Liu, S., Dong, W., ... Wohlfahrt, G. (2014). Global comparison of light use efficiency models for simulating terrestrial vegetation gross primary production based on the LaThuile database. *Agricultural and Forest Meteorology*, *192-193*, 108–120.

Yuan, W., Liu, S., Zhou, G., Zhou, G., Tieszen, L. L., Baldocchi, D., ... Wofsy, S. C. (2007). Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. *Agricultural and Forest Meteorology*, *143*, 189–207.

Zeng, N., Mariotti, A., & Wetzel, P. (2005). Terrestrial mechanisms of interannual CO(2) variability. *Global Biogeochemical Cycles*, *19*.

Zeng, N., Zhao, F., Collatz, G. J., Kalnay, E., Salawitch, R. J., West, T. O., & Guanter, L. (2014). Agricultural Green Revolution as a driver of increasing atmospheric CO2 seasonal amplitude. *Nature*, *515*, 394–397.

Zhang, Y., Guanter, L., Berry, J. A., Joiner, J., van der Tol, C., Huete, A., ... Kohler, P. (2014). Estimation of vegetation photosynthetic capacity from space-based measurements of chlorophyll fluorescence for terrestrial biosphere models. *Global Change Biology*, *20*, 3727–3742.

Zhang, G., Xiao, X., Dong, J., Kou, W., Jin, C., Qin, Y., ... Biradar, C. (2015). Mapping paddy rice planting areas through time series analysis of MODIS land surface temperature and vegetation index data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *106*, 157–171.

Zhang, Y., Xiao, X., Zhou, S., Ciais, P., McCarthy, H., & Luo, Y. (2016). Canopy and physiological controls of GPP during drought and heat wave. *Geophysical Research Letters*, *43*, 3325–3333.

Zhao, M., Heinsch, F. A., Nemani, R. R., & Running, S. W. (2005). Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sensing of Environment*, *95*, 164–176.

Zhao, M., Running, S. W., & Nemani, R. R. (2006). Sensitivity of Moderate Resolution Imaging Spectroradiometer (MODIS) terrestrial primary production to the accuracy of meteorological reanalyses. *Journal of Geophysical Research-Biogeosciences*, *111*.