Improving Global Gross Primary Productivity Estimates by Computing Optimum Light Use Efficiencies Using Flux Tower Data

Nima Madani

John S. Kimball
University of Montana - Missoula

Steven W. Running
University of Montana - Missoula

Let us know how access to this document benefits you.
Follow this and additional works at: https://scholarworks.umt.edu/ntsg_pubs

Recommended Citation
Improving Global Gross Primary Productivity Estimates by Computing Optimum Light Use Efficiencies Using Flux Tower Data

Nima Madani1,2, John S. Kimball1,2, and Steven W. Running1,2

1Numerical Terradynamic Simulation Group, W.A. Franke College of Forestry and Conservation, University of Montana, Missoula, MT, USA; 2Department of Ecosystem and Conservation Sciences, W.A. Franke College of Forestry and Conservation, University of Montana, Missoula, MT, USA

Abstract In the light use efficiency (LUE) approach of estimating the gross primary productivity (GPP), plant productivity is linearly related to absorbed photosynthetically active radiation assuming that plants absorb and convert solar energy into biomass within a maximum LUE (LUEmax) rate, which is assumed to vary conservatively within a given biome type. However, it has been shown that photosynthetic efficiency can vary within biomes. In this study, we used 149 global CO2 flux towers to derive the optimum LUE (LUEopt) under prevailing climate conditions for each tower location, stratified according to model training and test sites. Unlike LUEmax, LUEopt varies according to heterogeneous landscape characteristics and species traits. The LUEopt data showed large spatial variability within and between biome types, so that a simple biome classification explained only 29% of LUEopt variability over 95 global tower training sites. The use of explanatory variables in a mixed effect regression model explained 62.2% of the spatial variability in tower LUEopt data. The resulting regression model was used for global extrapolation of the LUEopt data and GPP estimation. The GPP estimated using the new LUEopt map showed significant improvement relative to global tower data, including a 15% R² increase and 34% root-mean-square error reduction relative to baseline GPP calculations derived from biome-specific LUEmax constants. The new global LUEopt map is expected to improve the performance of LUE-based GPP algorithms for better assessment and monitoring of global terrestrial productivity and carbon dynamics.

1. Introduction

With the increasing rate of atmospheric CO2 concentrations since the beginning of the industrial age, ecosystem carbon sequestration from vegetation gross primary productivity (GPP) has played a critical role in mitigating the potential climate impacts of global warming (Pan et al., 2011). However, global patterns and variability in GPP, and the continuing status of ecosystems as a sink for atmospheric CO2, are unclear due to large variability and uncertainties in global estimation of GPP (Anav et al., 2013). Among available global GPP models and data products, the NASA Moderate Resolution Imaging Spectroradiometer (MODIS) MOD17 (Running et al., 2004; Zhao & Running, 2010) and Soil Moisture Active Passive Level 4 Carbon (SMAP L4C) (Jones et al., 2017; Kimball et al., 2015) products provide the only operational satellite-based global GPP records. The MOD17 and L4C products utilize a similar light use efficiency (LUE) model framework driven by satellite observations of canopy fraction of photosynthetically active radiation (FPAR) and other geospatial data for consistent global estimation and monitoring of GPP.

The LUE model used for estimating GPP is based on the well-known Monteith equation (Monteith, Society, & Ecology, 1972), where plant production is linearly related to photosynthetically active radiation absorbed by the vegetation canopy (APAR) and the efficiency with which solar radiant energy is transformed into vegetation biomass through net photosynthesis, termed the light use efficiency (LUE). The LUE term in the MOD17 and L4C models is derived from a theoretical maximum light use efficiency (LUEmax) prescribed for different plant functional types using a global land cover classification; LUE is derived from LUEmax by accounting for reductions in conversion efficiency due to suboptimal environmental conditions for different biomes as derived from ancillary daily surface meteorology inputs (Jones et al., 2017; Zhao et al., 2005). It has been assumed that LUEmax varies according to vegetation type (Bartlett, Whiting, & Hartman, 1989) and defines the canopy photosynthetic capacity or maximum rate of conversion of APAR to vegetation biomass (g C MJ⁻¹) under optimal (nonlimiting) environmental conditions. LUE is reduced from LUEmax under...
suboptimal conditions defined by simple temperature and water deficit thresholds and linear ramp functions. The initial Monteith approach was adapted for using environmental inputs from Earth Observing System satellites in the 1990s, when there were only a limited number of tower carbon (CO₂) flux sites available for model development, calibration, and validation (Running et al., 1999). The model is based on the assumption that LUEmax variability is conservative within individual biomes characterized using a global land cover classification (Friedl et al., 2010). However, LUEmax can have large variability within individual biomes (Gitelson & Gamon, 2015; Madani et al., 2014; Turner et al., 2002) and is a major source of uncertainty in GPP estimation (Madani et al., 2014). The increasing availability of synergistic geospatial data, including satellite observations of chlorophyll fluorescence (Frankenberg et al., 2014; Joiner et al., 2011, 2013), global plant trait information (Kattge et al., 2011), and consistent long-term carbon flux observations from global tower site networks, can help in understanding variations in global productivity and provide insights to improve model GPP predictions beyond the use of discrete land cover or biome-type classifications and parameterizations.

The tower eddy covariance CO₂ flux measurements from global FLUXNET sites provide useful information for validating ecosystem models and understanding terrestrial carbon budgets (Baldocchi, 2008). Integration of tower observations and satellite remote sensing in context with process-based models and other supporting geospatial data, including key plant traits influencing ecosystem productivity (Reich, 2012) and solar-induced chlorophyll fluorescence (SIF), which is closely related to GPP (Guanter et al., 2014; Joiner, Yoshida, Guanter, et al., 2014 & Joiner, Yoshida, Vasilkov, et al., 2014; Walther et al., 2015; Yang et al., 2015), are expected to improve global GPP estimates and understanding.

Here we extend ecosystem modeling beyond the use of discrete land cover-type categories by developing and exploiting a spatially heterogeneous LUEopt global data layer. The primary objective of this research is to improve satellite-based estimation and understanding of GPP using a refined LUE model framework. We first quantify LUEopt using global carbon flux tower observations representing major vegetation functional types. We use a set of environmental data layers to explain the global variability in tower-derived LUEopt using a regression modeling framework. The resulting model framework is used for spatial extrapolation and prediction of LUEopt over the global domain. The resulting LUEopt map is then used within the LUE model framework for estimating global GPP. Alternative GPP predictions are made using baseline simulations derived using prescribed LUEmax values for different land cover types, consistent with the biome-properties look-up table (BPLUT) used in the MODIS MOD17 operational algorithm (Zhao et al., 2005). For model global GPP, we use the third-generation Global Inventory Modeling and Mapping Studies (GIMMS3g) fraction of photosynthetically active radiation (FPAR) product (Zhu et al., 2013) with daily surface meteorological data from the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) global reanalysis (Bosilovich, Lucchesi, & Suarez, 2015; Molod et al., 2015). This approach provides a pathway to model more than three decades of global GPP using satellite observations. However, our goal in this study is global estimation of LUEopt and assessment of the relative value of spatially explicit LUEopt data on GPP estimation accuracy. For this reason, we model global GPP once using spatially explicit LUEopt inputs and once using prescribed LUEmax constants for different land cover types, and compare the results. Each model GPP simulation uses the same FPAR and surface meteorological inputs so that the resulting model differences are independent of the particular choice of FPAR and meteorological data used. This work is a global extension of a previous study that applied a similar methodology for estimating LUEopt across a North American regional domain (Madani et al., 2014), where the resulting GPP calculations showed significant accuracy improvement against independent regional tower observations and baseline GPP calculations derived using prescribed LUEmax values for different biome types. For the current study, we apply similar methods, while exploiting a larger global tower observation network database and global range of land cover conditions for spatially explicit LUEopt and GPP predictions over a global domain.

2. Methods
2.1. Flux Tower-Based LUEopt Calculations
We extended our previous work of modeling LUEopt (Madani et al., 2014) using a larger global network of 149 tower sites from the La Thuile FLUXNET synthesis database (Baldocchi, 2008) and the more recent FLUXNET2015 (FLUXNET, 2015) global synthesis data record. We selected 54 tower sites from the
FLUXNET2015 record for model testing and used the La Thuile tower data record for model training purposes. The tower eddy covariance CO₂ flux measurement sites were selected for this study on the basis of having at least one full year of gap-filled daily CO₂ flux data and representing a broad range of global biomes (Figure 1 and Table S1 in the supporting information).

We used the GIMMS3g bimonthly FPAR record (Zhu et al., 2013) and applied a temporal linear interpolation of the bimonthly data to produce a continuous daily FPAR record for each global grid cell over the 2000 to 2011 record. The FPAR data for grid cells collocated with selected tower site locations were temporally matched with the tower GPP records. We estimated LUEopt for each selected tower site by selecting the upper 98–99.5% bin of daily gap-filled GPP values throughout the available tower measurement years and using these values to represent the maximum daily GPP (GPPmax) from each tower site. It is assumed that at the upper bin of GPP, plant activity is not restricted by constraining climate factors (Kergoat et al., 2008; Madani et al., 2014). For all days with such criteria, LUE is defined as

$$LUE = \frac{GPP_{\text{max}}}{APAR}$$

In equation (1), APAR is the product of FPAR defined from the GIMMS3g record and daily PAR, which is estimated as half of the global incoming shortwave solar radiation derived from the MERRA-2 global reanalysis (Bosilovich et al., 2015; Molod et al., 2015). For each of the tower sites we averaged the tower-derived daily LUE observations falling in the upper GPP range (98–99.5%) from equation (1) and used these results to represent the LUEopt value of each site. The upper 5% bin of each GPP record was ignored to minimize potential uncertainty contributed by GPP saturation and outlier effects (Kergoat et al., 2008; Turner et al., 2003).

### 2.2. Extrapolating LUEopt From Point to Global Scale

We used multiple explanatory variables derived from other ancillary data as proxies to represent potential landscape features and vegetation factors influencing LUEopt and GPP (Table 1). The climate-related variables included average annual temperature and precipitation, and precipitation and temperature of driest quarter from the WorldClim database (Hijmans et al., 2005) with 1 km resolution; average annual root zone (0–1 m depth) soil moisture, and vapor pressure deficit from MERRA-2 (Bosilovich et al., 2015; Molod et al., 2015) with 0.5 × 0.65° global grid resolution. Topographical related variables included elevation data from a global Shuttle Radar Topography Mission-based digital elevation model (DEM) with 1 km spatial resolution (Farr, Rosen, & Caro, 2007) and the topographic wetness index (TWI) created in the System for Automated...
Geoscientific Analysis environment (Conrad et al., 2015) using the DEM. Vegetation productivity-related variables included extrapolated key plant traits posted to a 5 km resolution global grid representing specific leaf area (SLA; the ratio of leaf area per unit dry mass, $m^2$ kg$^{-1}$) and canopy height (Madani et al., 2014), average annual FPAR derived from the GIMMS3g record posted to a 8 km resolution global grid, composited monthly SIF from the Global Ozone Monitoring Experiment-2 (GOME-2) satellite record posted to a 0.5° resolution global grid (Joiner et al., 2011, 2013), and a MODIS global land cover classification (MCD12C1-Type 2) posted to a 5 km resolution global grid (Friedl et al., 2010).

We used the upper 95–98% quantile of SIF data as a proxy for $LUE_{opt}$ and $SIF_{yield}$ (emitted SIF per absorbed PAR) (Grace et al., 2007; Yang et al., 2015). The nearest neighbor technique was used to resample all of the available data sets into a consistent 8 km resolution global grid as the GIMMS3g FPAR record. The Pearson correlation coefficient in the R programming environment (R Core Team, 2016) was used to select the variables with the highest predictive power and collinearity of less than 70% to build a linear mixed effect model.

Using a mixed effect model in the lm4 library (Bates et al., 2014) in R allowed us to use the predictive power of discrete land cover class information in the model. A mixed effect model can use variable intercepts and slopes for continuous environmental variables depending on each category of a random variable. In this model

$$Y = X_i \times \beta + Z_i \times b_i + \epsilon_i$$

where $X_i \times \beta$ is the fixed term, $Z_i \times b_i$ is the random term, and $\epsilon_i$ is the associated error. For modeling $LUE_{opt}$, we separated the global tower sites into two subsets for model training and testing purposes. In this way, we developed the model using 95 tower training sites from the La Thuile record and tested the resulting model against 54 independent tower sites from the FLUXNET2015 database. Due to a limited number of tower sites available for some land cover classes, we merged these classes into coarser set of needleleaf (evergreen needle leaf forest (ENF) + deciduous needle leaf forest (DNF)), shrubland (closed shrubland (CSH) + open shrubland (OSH)), and savanna (woody savannas (WSA) + savanna (SAV)) categories for the regression analysis and kept all other classes consistent with the underlying land cover map.

### 2.3. Modeling the Global Daily GPP

The bimonthly GIMMS3g FPAR data were linearly gap-filled to create a continuous global daily FPAR data record from 2000 to 2011. Meteorology data including daily minimum air temperature and
incoming solar radiation were acquired from MERRA-2 global reanalysis (Bosilovich et al., 2015) and were used with the interpolated daily FPAR as primary LUE model inputs. The daily meteorology data were resampled from a native 0.5 × 0.65° spatial resolution to the same 8 km resolution global grid as the GIMMS3g FPAR inputs. The vapor pressure deficit was estimated using daily surface air temperature and dew point temperature (Murray, 1967), while daily GPP was modeled as

\[
\text{GPPLUEopt} = \frac{\text{FPAR}}{\text{C2}} \times \frac{\text{PAR}}{\text{C2}} \times \frac{\text{LUEopt}}{\text{C2}} \times f_T \times f_{\text{VPD}} \tag{3}\]

where \( f_{\text{VPD}} \) and \( f_T \) represent dimensionless environmental constraint functions ranging between zero (fully constrained) and unity (no effect) that describe the reduction in LUE and GPP due to cold temperatures:

\[
f_T = \begin{cases} 
0, & T_{\text{min}} \leq T_{\text{Mmin}} \\
\frac{T_{\text{Mmin}} - T_{\text{min}}}{T_{\text{Mmax}} - T_{\text{Mmin}}}, & T_{\text{Mmin}} < T_{\text{min}} < T_{\text{Mmax}} \\
1, & T_{\text{min}} \geq T_{\text{Mmax}} 
\end{cases}
\tag{4}\]

and excessive atmosphere moisture deficits:

\[
f_{\text{VPD}} = \begin{cases} 
0, & VPD \geq VPD_{\text{Max}} \\
1 - \frac{VPD}{VPD_{\text{Min}}}, & VPD_{\text{Min}} < VPD < VPD_{\text{Max}} \\
1, & VPD \leq VPD_{\text{Min}} 
\end{cases}
\tag{5}\]

The Min and Max subscripts in equations (4) and (5) represent the minimum and maximum defined thresholds for minimum daily temperature (\( T_{\text{min}} \)) and vapor pressure deficit (VPD) functions (refer to Table S2 of the supporting information for the MOD17 BPLUT). For comparison, we modeled the GPP once again using fixed \( \text{LUEmax} \) values defined for different biome types from the MODIS MOD17 data product (Zhao & Running, 2010). The only difference between two sets of GPP model simulations in this study is the application of \( \text{LUEopt} \) versus the prescribed \( \text{LUEmax} \) constants for different biome types, while all other inputs were consistent between the two model runs. For validation, we compared our estimated average annual GPP for the 2000–2007 record with the average annual flux tower GPP data from the La Thuile record. We used the estimated average annual GPP from the 2007–2011 portion of the model record to compare with tower GPP observations from the 54 independent test sites from the 2015 FLUXNET database. The model GPP results were also compared against other GPP global observational benchmarks, including Model Tree ensemble (MTE)-based GPP derived from upscaled global tower data.
The inset plot shows the same relationship for 54 independent tower test sites. The circles represent different tower sites, while the different colors distinguish the dominant land cover class for each site. The black line denotes the 1:1 relationship. Refer to Table 2 for land cover-type definitions.

Figure 4. Relationship between LUEopt derived from the generalized mixed effect model (x axis) and tower observation based LUEopt estimates from 95 global tower sites; the modeled results explained 62.2% ($R^2$) of observed tower site LUEopt variability, with an associated RMSE difference of 0.31 g C MJ$^{-1}$.

<table>
<thead>
<tr>
<th>Land cover</th>
<th>Intercept</th>
<th>SIF</th>
<th>Elevation (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENF + DNF</td>
<td>0.491</td>
<td>0.437</td>
<td>$-0.08$</td>
</tr>
<tr>
<td>EBF</td>
<td>0.646</td>
<td>0.351</td>
<td>0.018</td>
</tr>
<tr>
<td>DBF</td>
<td>0.897</td>
<td>0.439</td>
<td>$-0.43$</td>
</tr>
<tr>
<td>MF</td>
<td>0.589</td>
<td>0.364</td>
<td>0.012</td>
</tr>
<tr>
<td>CSH + OSH</td>
<td>0.533</td>
<td>0.372</td>
<td>$-0.1$</td>
</tr>
<tr>
<td>WSA + SAV</td>
<td>0.599</td>
<td>0.369</td>
<td>0.002</td>
</tr>
<tr>
<td>GRA</td>
<td>0.394</td>
<td>0.512</td>
<td>$-0.185$</td>
</tr>
<tr>
<td>CRO</td>
<td>0.409</td>
<td>0.573</td>
<td>$-0.295$</td>
</tr>
</tbody>
</table>

Note. All model parameters are statistically significant ($p < 0.05$).

3. Results

Our results show that LUEopt ranges from 0.3 to 2.62 g C MJ$^{-1}$ and is highly variable both within and among different land cover types, while the prescribed LUE$\text{max}$ values have much less variability among biomes (1.09 ± 0.16 g C MJ$^{-1}$). Based on the tower estimated LUEopt, GPP modeled using prescribed LUE$\text{max}$ values should be significantly lower than GPP recorded by the tower sites for cropland, grassland, deciduous, and mixed forest biomes, as the median LUEopt values for these land cover types are higher than the prescribed LUE$\text{max}$ rates (Figure 2). For cropland sites, LUE$\text{max}$ was approximately 1.5 times lower than the 95th percentile of tower estimated LUEopt (1.044 versus 2.599 MJ m$^{-2}$).

However, the average LUEopt values for Savanna and closed shrubland biomes are respectively about 22.5% and 37.5% lower than LUE$\text{max}$. It should be noted that closed shrublands only cover a small proportion (~0.2%) of the global vegetated land area (Table 2).

Extrapolating the sparse tower LUEopt values over the global domain requires finding meaningful relationships between the in situ tower observations and other spatially explicit explanatory variables. The partial correlation analysis showed that SIF, followed by SLA and the DEM, showed the strongest correlations with tower LUEopt (Figure 3). We used linear regression analysis and found that the global land cover classification (MODIS MCD12C1-Type 2) only explained 29% of the observed variance in LUEopt values among the 95 global tower training sites examined. By comparison, each of the other predictor variables, namely, SIF, SLA, and elevation, respectively, explained 36.3%, 24.3%, and 13.3% of the variance in the flux tower-derived LUEopt values.

By incorporating the extrapolated plant trait information, including SLA, in conjunction with SIF and elevation as predictors in the mixed effect regression model, we explained 62.2% of the spatial variance in global tower observation-based LUEopt with an associated root-mean-square error (RMSE) difference of 0.314 g C MJ$^{-1}$. These results led to a 114% improvement in the LUEopt variance explained by the model over the land cover classification-based estimates. The cross-validation analysis of 54 independent tower sites showed an acceptable model RMSE performance of 0.352 g C MJ$^{-1}$ and $R^2$ of 57.8% in the spatially explicit LUEopt estimates in relation to the tower-derived LUEopt values (Figure 4). The regression analysis shows that LUEopt, as expected, has a positive relationship with the productivity-related environmental variables (SIF and SLA) for all land cover types, but the LUEopt values tend to decline with increasing elevation for needleleaf and deciduous broadleaf forests, shrublands, grasslands, and croplands (Table 3). Increase in elevation not only is associated with a reduction in temperature but also affects plant and soil nutrient supply and availability (Mayor et al., 2017). Because SLA has been shown to have positive correlation with leaf nitrogen content across the biomes (Reich, 2012), and thus productivity and LUE, we used SLA as a fixed term in the mixed effect model (variable intercept and fixed slope of 0.013 ± 0.004). With the aid of global maps of the other explanatory variables, we used the mixed effect model to extrapolate in situ flux tower-estimated LUEopt values over the global domain (Figure 5).

We used the new refined LUE model to calculate mean annual GPP (2000–2007) across the model training tower sites. The new model...
explained 84.3% ($R^2$) of the spatial variance in the tower annual GPP observations, with associated RMSE differences of 313 g C m$^{-2}$. In contrast, using the $LUE_{\text{max}}$ constants with the same meteorology and FPAR inputs explained a lower proportion (75.4%) of the variance in tower GPP and degraded the associated RMSE (508 g C m$^{-2}$) (Figure 6). Our new GPP model ($LUE_{\text{opt}}$-GPP) also showed higher performance ($R^2 = 62.5\%$, RMSE = 488 g C m$^{-2}$ yr$^{-1}$) than GPP modeled using land cover prescribed $LUE_{\text{max}}$ constants for the independent tower test sites ($LUE_{\text{max}}$-GPP: $R^2 = 53.2\%$, RMSE = 705 g C m$^{-2}$ yr$^{-1}$).

We extended the analysis and compared the sum of annual GPP (g C m$^{-2}$ yr$^{-1}$) averaged from 2007 to 2011 between the new model GPP calculations derived using spatially explicit $LUE_{\text{opt}}$ values versus baseline GPP
calculations derived using fixed biome-specific LUEmax inputs and other GPP estimates derived using a tower observation upscaled MTE approach (Jung et al., 2011). Figure 7 shows the difference in GPP estimates derived from the different approaches. The average total global GPP from 2007 to 2011 derived using the new LUEopt data was 121.8 ± 0.9 (SD) Pg C yr\(^{-1}\), which is 16.5% higher than the estimated GPP derived using fixed LUEmax values from this study, and 10.4% higher than the MOD17 operational GPP product for the same period. On the other hand, the new GPP product is only 2.1% higher than the MTE-based GPP record.

We further compared the three GPP products with the GOME-2 SIF record, which extends from 2007. We compared the half-degree binned latitudinal averaged daily GPP data from 2007 to 2011 based on LUEopt with similarly averaged GPP derived from LUEmax and the MTE-GPP record. Our results show that LUEopt-GPP has higher correspondence (\(R^2 = 94\%\)) with the average monthly SIF record than LUEmax-GPP (\(R^2 = 86.6\%\)) and MTE-GPP (\(R^2 = 89\%\)) for the same time period (Figure 8). The LUEopt-GPP values are higher than LUEmax-GPP and MTE-GPP between 40°N and 60°N. This region covers 42% of the global cropland area, where LUEopt-GPP generally shows higher productivity than LUEmax-GPP and MTE-GPP (Figure 7). In contrast, MTE-GPP is generally higher than LUEmax-GPP and LUEopt-GPP in tropical regions. The annual LUEopt-GPP is 3.2 Pg (8.5%) higher than LUEmax-GPP for evergreen broadleaf biomes and close to the MTE-GPP estimate for these areas.

A comparison of the GPP climatology trends modeled using LUEopt versus LUEmax in relation to the tower GPP observations for selected validation sites representing major global land cover types is presented in Figure 9. These results show that the LUEopt- and LUEmax-derived GPP results show similar seasonal phenology as the tower observations and with each other due to the models using the same FPAR inputs and bioclimatic factors. However, the use of land cover independent LUEopt data over the LUEmax constants generally lowers the associated GPP estimation error (RMSE) relative to the tower observations so that among the test sites the RMSE difference was approximately 30% lower when we used LUEopt to model the GPP.

4. Discussion

The light use efficiency approach in modeling GPP can exploit relatively high spatial and temporal resolution remote sensing-based vegetation information for modeling GPP over the entire globe. Spatially explicit global FPAR data from operational satellites have been used to continuously model global GPP for almost two decades (Running et al., 2004). While these operational GPP records have been shown to have favorable correspondence with in situ tower measurement-based GPP observations and other model GPP products (e.g., Li et al., 2013; Running et al., 2004; Zhao et al., 2005), our results indicate that the LUE algorithm approach can
Figure 7. Comparison between modeled GPP for the 2007–2011 record derived using spatially variable LUE_{opt} values in relation to alternative GPP estimates derived using biome-specific LUE_{max} constants and tower observation upscaled (MTE) GPP (Jung et al., 2011). (a) Annual GPP difference derived from LUE_{opt} and LUE_{max} approaches. (b) Annual GPP difference between LUE_{opt} and MTE GPP data. (c) The bar plot showing the sum of annual GPP and the temporal standard deviations (vertical error bars) between LUE_{opt}, LUE_{max}, and MTE GPP for different land cover types.
be further refined to produce better GPP performance. The LUE\textsubscript{opt} parameter defines the maximum rate of conversion of APAR into vegetation biomass in the LUE model framework and is highly variable within and across different biomes. Our results indicate that LUE\textsubscript{opt} variability is the highest for cropland and grassland biomes that cover about a quarter of the total global vegetated land area.

Our biome-level analysis of differences between the two LUE-based GPP estimation approaches examined in this study shows that the use of fixed LUE\textsubscript{max} constants for different biome types, similar to the method used in the MODIS MOD17 operational product, generally results in underestimated GPP for croplands. There is a 7.5 ± 0.6 Pg CO\textsubscript{2} difference in average annual GPP derived using spatially variable LUE\textsubscript{opt} inputs versus land cover class-specific LUE\textsubscript{max} constants for the 2007–2011 record. We showed that the estimated global annual GPP derived using the spatially heterogeneous LUE\textsubscript{opt} inputs is 121.8 ± 0.9 Pg C and that using prescribed LUE\textsubscript{max} constants results in a 16.5% reduction in this value. Our results are similar to other global GPP estimates from the literature ranging between 106 and 175 Pg C yr\textsuperscript{-1} (Anav et al., 2013); our GPP results are also similar to values reported from two previous flux tower observation global upscaling assessments (Beer et al., 2010; Jung et al., 2011).

With respect to the uncertainties in our LUE\textsubscript{opt} modeling approach, we showed that SIF alone can explain 36.3% of the variance in LUE\textsubscript{opt}, which represents a 7.3% improvement in performance over using a discrete global land cover-type classification to explain LUE\textsubscript{opt} variability. Limitations of the SIF data used in this study include coarse spatial resolution and temporal compositing used to enhance the satellite signal to noise. The relationship between SIF and GPP may vary also under environmental stress (Kouril et al., 2004), which was not accounted for in this investigation. However, our use of the global satellite SIF observations for LUE model parameterization can offset limitations imposed by sparse in situ tower sites. The LUE approach used in this study provides much finer spatial resolution and temporal fidelity than the global SIF record and is similar to the local landscape-level sampling footprint from tower eddy covariance CO\textsubscript{2} flux measurements. The LUE model also provides a direct link to underlying environmental controls influencing productivity based on ecological principles (Madani et al., 2017). SIF is related to APAR based on SIF\textsubscript{yield} (Grace et al., 2007), which may provide a more direct estimate of GPP from SIF observations. However, a general constraint to this approach is the coarse spatial and temporal resolution of current global satellite-based SIF records, which may not effectively capture subgrid-scale heterogeneity. These limitations are being addressed through planned improvements in satellite remote sensing and finer landscape-level SIF observations (Coppo et al., 2017). These data are contributing to the development of a new generation of physical models and LUE algorithms for estimation and extrapolation of SIF and GPP calculations that are more directly comparable with both satellite SIF observations and tower carbon flux measurements (Thum et al., 2017).

Some of the limitations of our approach in defining ecosystem LUE\textsubscript{opt} are related to the way we estimated LUE\textsubscript{opt} for each tower location using tower measurement-based GPP observations, which are subject to uncertainty. However, we minimized potential errors by ignoring the upper 5th percentile distribution of the GPP record at each site and used the 98–99.5% bin of daily GPP observations to derive LUE\textsubscript{opt}. Spatial resampling of the 5 km land cover map to the coarser 8 km resolution of the FPAR record caused some tower site differences between the dominant land cover conditions represented within the local tower footprint and the overlying land cover type represented by the global satellite observations and model simulations. These differences included four ENF towers classified as MF, and one MF tower classified as ENF; one CSH tower classified as WSA; one SAV tower was classified as GRA; and two GRA towers classified as CRO types. These classification differences represented only ~9% of the tower sites represented, while land cover class was a relatively minor element of the statistical model used for predicting LUE\textsubscript{opt}. Overall, our model
indicated favorable GPP performance using the LUE\textsubscript{opt} inputs ($R^2 = 84.3\%$, RMSE = 313 g C m$^{-2}$). These results represent 9.3\% $R^2$ and 30.6\% RMSE improvement in GPP performance relative to baseline model simulations derived using fixed LUE\textsubscript{max} inputs (e.g., Figures 6 and 9).

Previous vegetation productivity and validation studies have generally used a limited number of sparsely distributed tower sites, while our results indicate that satellite SIF observations and other plant trait information can improve LUE model-based predictions and global monitoring of ecosystem productivity. The use of satellite SIF observations as an independent proxy for GPP and LUE can also offset potential errors imposed by sparse tower measurement sites on the estimation of LUE\textsubscript{opt} across different biomes. Our latitudinal GPP analysis showed a higher correlation between SIF and LUE\textsubscript{opt}-GPP, than the LUE\textsubscript{max}-GPP and MTE-GPP products, indicating that the LUE\textsubscript{opt}-GPP data may provide more effective representation of regional-scale variations in ecosystem productivity.

Three decades of FPAR observations and long-term tower observations offer the potential to analyze LUE\textsubscript{opt} changes and underlying physical drivers, and their impact on GPP trends. Availability of three

**Figure 9.** Comparison between monthly climatology (2007–2011) in GPP modeled using the extrapolated global LUE\textsubscript{opt} map and the baseline GPP modeled using biome-specific LUE\textsubscript{max} constants and selected tower GPP observations used for model validation. Refer to Table 2 for land cover-type definitions.
decades of remote sensing FPAR observations, potentially allow new assessments of global GPP trends using variable LUEopt inputs. Moreover, current operational satellite-based GPP records derived using the LUE modeling approach, including the MODIS-MOD17 and SMAP-L4C products, may benefit from spatially explicit LUEopt inputs.

5. Conclusion

The results of this study revealed large spatial variability in optimal LUE levels both within and among global biomes that is related to heterogeneous landscape and plant trait characteristics. We defined LUEopt across a global network of flux tower measurement sites representing major biome types and explained the observed LUEopt spatial variability using a set of predictor variables, including vegetation characteristics represented by a global land cover classification, satellite-based SIF observations, SLA from a physical plant traits global data record, and landscape characteristics represented by a digital terrain map. We modeled global GPP using a light use efficiency (LUE) model and the new LUEopt map as a primary ancillary input. Our results showed significant improvement over alternative GPP simulations derived using prescribed LUEmax Constants for different biome types. The LUEopt modeled GPP also performed better than the LUEmax-GPP simulations for a set of independent global tower validation sites. Our approach and the output LUEopt data, combining synergistic information from spatially explicit land cover, topography, and SIF observations with extrapolated plant trait information, can lead to better LUE model-based global GPP predictions and understanding.

Acknowledgments
This study was supported by funding from the NASA Earth Science program (NNX14AI50G and NNX15AS59G). This work used eddy covariance data acquired and shared by the FLUXNET community, including these networks: AmeriFlux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada, GreenGrass, ICOS, KoFlux, LBA, NECC, OzFlux-TERN, TCOS-Siberia, and USCC. The FLUXNET eddy covariance data processing and harmonization were carried out by the ICOS Ecosystem Thematic Center, AmeriFlux Management Project, and Fluxdata project of FLUXNET, with the support of CDIAC, and the OzFlux, ChinaFlux and AsiaFlux offices. All data sets used in this paper are available from the cited literature and publicly accessible records. The data products developed from this research are available through the following link: http://files.ntsg.umt.edu/data/Optimum_LUE.

References