Evaluating GPP and respiration estimates over northern mid-latitude ecosystems using solar induced fluorescence and atmospheric CO$_2$ measurements

B. Byrne$^1$, D. Wunch$^1$, D. B. A. Jones$^{1,2}$, K. Strong$^1$, F. Deng$^1$, I. Baker$^3$,
P. Köhler$^4$, C. Frankenberg$^{4,5}$, J. Joiner$^6$, V. K. Arora$^7$, B. Badawy$^{8,12}$,
A. B. Harper$^9$, T. Warneke$^{10}$, C. Petri$^{10}$, R. Kivi$^{11}$, and C. M. Roehl$^4$

Brendan Byrne, bbyrne@physics.utoronto.ca

$^1$Department of Physics, University of Toronto, Toronto, Ontario, Canada
$^2$Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, California, USA
$^3$Atmospheric Science Department, Colorado State University, Fort Collins, Colorado, USA
$^4$Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California, USA

©2018 American Geophysical Union. All Rights Reserved.
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA

NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

Climate Research Division, Environment and Climate Change Canada, Victoria, British Columbia, Canada

Climate Research Division, Environment and Climate Change Canada, Downsview, Ontario, Canada

College of Engineering, Mathematics, and Physical Sciences, University of Exeter, Exeter, UK

Institute of Environmental Physics, University of Bremen, Bremen, Germany

Finnish Meteorological Institute, Sodankylä, Finland

Now at Meteorological Research Division, Environment and Climate Change Canada, Dorval, Quebec, Canada
Abstract. On regional to global scales, few constraints exist on gross primary productivity (GPP) and ecosystem respiration ($R_e$) fluxes. Yet, constraints on these fluxes are critical for evaluating and improving terrestrial biosphere models (TBMs). In this study, we evaluate the seasonal cycle of GPP, $R_e$ and net ecosystem exchange (NEE) produced by four TBMs and FLUXCOM, a data-driven model, over northern mid-latitude ecosystems. We evaluate the seasonal cycle of GPP and NEE using solar induced fluorescence (SIF) retrieved from the Global Ozone Monitoring Experiment-2 (GOME-2) and column-averaged dry-air mole fractions of $CO_2$ ($X_{CO_2}$) from the Total Carbon Column Observing Network (TCCON), respectively. We then infer $R_e$ by combining constraints on GPP with constraints on NEE from two flux inversions. An ensemble of optimized $R_e$ seasonal cycles is generated using five GPP estimates and two NEE estimates. The optimized $R_e$ curves generally show high consistency with each other, with the largest differences due to the magnitude of GPP. We find optimized $R_e$ exhibits a systematically broader summer maximum than modeled $R_e$, with values lower during June–July and higher during the fall than $R_e$. Further analysis suggests that the differences could be due to seasonal variations in the carbon use efficiency (possibly due to an ecosystem scale Kok effect) and to seasonal variations in the leaf litter and fine root carbon pool. The results suggest that the inclusion of variable carbon use efficiency for autotrophic respiration and
carbon pool dependence for heterotrophic respiration are important for accurately simulating $R_e$.

**Keypoints:**

- Top-down constraints on ecosystem respiration are obtained by combining atmospheric CO$_2$ and solar induced fluorescence observations.
- Inferred ecosystem respiration suggests a systematically broader summer maximum than bottom-up estimates over the northern mid-latitudes.
- Inferred ecosystem respiration shows high sensitivity to the magnitude of gross primary productivity.
1. Introduction

The terrestrial biosphere is currently a major sink of carbon dioxide (CO$_2$), taking up roughly one quarter of anthropogenic emissions [Ciais et al., 2013; Le Quéré et al., 2018]. This uptake is largely the result of an imbalance between Gross Primary Productivity (GPP) and ecosystem respiration ($R_e$), the biological processes by which atmospheric CO$_2$ is absorbed from and released to the atmosphere. Although the rate of global net uptake is well constrained, many questions remain. For example, the spatial footprint of the uptake is poorly understood. There remains substantial disagreement on the partitioning of the drawdown between the tropics and northern hemisphere [Ciais et al., 2013; Houweling et al., 2015], and little consensus on smaller scales. Furthermore, little is understood of how the terrestrial carbon sink will change in the future. Projections of the future carbon cycle rely on terrestrial biosphere models (TBMs), which show large disagreements on the relative importance of different processes driving the uptake [Huntzinger et al., 2017].

To refine estimates of uptake, the underlying processes driving imbalances in GPP and $R_e$ need to be better understood. To this end, better constraints on GPP and $R_e$ are required. In this study, we examine regional-to-global scale constraints on fluxes, and how these constraints can be used to evaluate modeled estimates of GPP and $R_e$ in the northern extratropics.

Net ecosystem exchange (NEE), the sum of GPP and $R_e$, can be constrained on large scales using atmospheric CO$_2$ observations. On seasonal timescales, variations of CO$_2$ are primarily driven by NEE fluxes. Therefore, modeled NEE fluxes can be evaluated by comparing simulated atmospheric CO$_2$ using an atmospheric transport model with
observed atmospheric CO$_2$. This method has previously been applied to evaluate TBM fluxes [Messerschmidt et al., 2013; Peng et al., 2015]. However, atmospheric CO$_2$ constraints cannot be used alone to evaluate the component GPP and $R_e$ fluxes. Constraints on NEE can be related to GPP and $R_e$ through the relationship,

$$\text{NEE} = \text{GPP} + R_e,$$

(1)

where fluxes are defined so that negative values represent removal of CO$_2$ from the atmosphere and positive values indicate emission to the atmosphere. Therefore, if independent constraints on either GPP or $R_e$ are used in combination with constraints on NEE, it would be possible to evaluate both GPP and $R_e$. Currently, there are no large-scale observational constraints on $R_e$, but recent advances in remote sensing have provided a new constraint on large-scale GPP.

Within the last few years, satellite observations of solar induced chlorophyll fluorescence (SIF) have become possible [Joiner et al., 2011; Frankenberg et al., 2011a; Guanter et al., 2012]. SIF is the emission of radiation by chlorophyll during photosynthesis and thus provides a proxy for GPP [Papageorgiou and Govindjee, 2007]. Although challenges remain in quantifying GPP from satellite SIF observations, many studies have found linear relationships between satellite retrievals of SIF and GPP from the canopy to ecosystem scale on weekly to monthly time scales [Frankenberg et al., 2011b; Guanter et al., 2012; Yang et al., 2015; Damm et al., 2015; Zhang et al., 2016a, b; Wood et al., 2017; Sun et al., 2017]. SIF is a more direct proxy for GPP than vegetation indices [Walther et al., 2016; Jeong et al., 2017; Luus et al., 2017], because other variables are required to estimate GPP from vegetation indices. For example, photosynthetically active radiation (PAR) and light-use efficiency are required to estimate GPP from normalized differential vege-
tation index (NDVI) [Field et al., 1995]. In this study, we investigate the utility of using atmospheric CO$_2$ and SIF observations to evaluate fluxes of NEE, GPP and $R_e$. First, observations of atmospheric CO$_2$ and SIF are used to evaluate estimates of NEE and GPP, respectively. Then, constraints on NEE and GPP are combined to evaluate $R_e$ estimates.

For atmospheric CO$_2$ observations, we use the retrieved column-averaged dry-air mole fractions of CO$_2$ ($X_{CO_2}$) from the Total Carbon Column Observing Network (TCCON) [Wunch et al., 2011]. Modeled $X_{CO_2}$ is generated by simulating atmospheric CO$_2$ fields with the GEOS-Chem chemical transport model driven with imposed NEE as input surface fluxes. The simulated atmospheric CO$_2$ fields are then integrated in altitude and compared to TCCON. Using this method, the seasonal cycle of the model-NEE-based $X_{CO_2}$ is compared with the seasonal cycle observed at several TCCON sites.

For SIF, the longest record of space-based observations is from the Global Ozone Monitoring Experiment-2 (GOME-2) instrument aboard the Meteorological Operational Satellite-A (MetOp-A), which was launched by the European Space Agency in 2006 [Joiner et al., 2013; Köhler et al., 2015]. Eight years of GOME-2 SIF observations (from 2007–2014) are used to evaluate the mean seasonal behavior of GPP in the TBMs.

We evaluate the constraints that SIF and $X_{CO_2}$ provide on four TBMs and FLUXCOM upscaled fluxes [Tramontana et al., 2016]. Two of the TBMs examined here employ diagnostic phenology: the Carnegie-Ames Stanford Approach [CASA, Potter et al., 1993; Randerson et al., 1996] and the Simple Biosphere model version 3 [SiB3, Baker et al., 2008]. For these TBMs, satellite observations of vegetation indices are assimilated to prescribe phenology. These TBMs are widely used to generate prior fluxes for flux inversion analyses [Schuh et al., 2010, 2013; Gurney et al., 2004], and would likely be employed in flux...
inversions assimilating SIF and atmospheric CO\textsubscript{2}. Thus, it is necessary to understand the level of a priori agreement to expect between the observations and these TBMs. We also evaluate two prognostic TBMs: the Canadian Terrestrial Ecosystem Model [CTEM, Melton and Arora, 2016], and the Joint UK Land Environment Simulator [JULES, Clark et al., 2011; Harper et al., 2018]. In contrast to the diagnostic TBMs, CTEM and JULES model phenology only as a function of the driving meteorology. Prognostic TBMs are used in simulations of future climate, thus it is desirable to understand the agreement of these TBM fluxes with observational constraints. FLUXCOM products are generated using upscaling approaches based on machine learning methods that integrate FLUXNET site level observations of CO\textsubscript{2} fluxes, satellite remote sensing, and meteorological data [Tramontana et al., 2016; Jung et al., 2017]. For this study, we examine upscaled fluxes generated using random forests (RF), multivariate regression splines (MARS) and artificial neural networks (ANN). FLUXCOM GPP and $R_e$ are widely considered to be among the best estimates available, thus it is important to include these fluxes in our comparison. FLUXCOM NEE estimates are known to produce an unrealistic large annual net sink by the biosphere (18–28 Pg/yr) [Tramontana et al., 2016; Jung et al., 2017], thus we do not evaluate FLUXCOM NEE against TCCON.

After evaluating model GPP and NEE, we examine the possibility of combining GPP and NEE constraints from atmospheric CO\textsubscript{2} and SIF observations to evaluate model $R_e$. An “optimized” $R_e$ seasonal cycle is calculated using NEE fluxes produced by two atmospheric CO\textsubscript{2} flux inversions and GPP fluxes produced by CASA, SiB3 and FLUXCOM, as these GPP fluxes give close agreement with the normalized seasonal cycle of SIF. We further examine the sensitivity of this estimate to uncertainties in GPP and NEE fluxes,
discuss possible reasons for differences between our optimized $R_e$ and TBM $R_e$, and discuss the current limitations of estimating optimized $R_e$ with existing observational constraints.

The area of study is limited to the northern extratropics (39°-65° N). These latitudinal limits were chosen because the seasonal variations in $X_{CO_2}$ and SIF are largest over these latitudes, and thus provide the largest signal in the observations.

This paper is organized as follows. In Section 2, the data and our methods are described. In Section 3, we present the results of our experiments. We first describe the agreement between SIF and GPP, and then $X_{CO_2}$ and NEE. Then the feasibility of evaluating $R_e$ estimates by combining GPP and NEE constraints is examined. In Section 4, we discuss the plausibility of our optimized $R_e$ seasonal cycle and possible sources of error. Then the limitations of calculating optimized $R_e$ given the observational constraints is discussed.

In Section 5, we give our conclusions.

2. Data and Methods

2.1. Terrestrial Biosphere Models

This study examines GPP, $R_e$ and NEE from four TBMs that use a range of input parameters. The TBMs used are: CASA [Potter et al., 1993; Randerson et al., 1996], SiB3 [Baker et al., 2008], CTEM [Melton and Arora, 2016], and JULES [Clark et al., 2011; Harper et al., 2018]. CASA and SiB3 assimilate satellite observations of vegetation indices to produce diagnostic phenology, while CTEM and JULES employ prognostic phenology in which the phenology is a function of the driving meteorology. We use two sets of CTEM fluxes that are driven by different meteorology to examine the impact of the driving meteorology on GPP, $R_e$, and NEE. CTEM-CRU is driven by NCEP-CRU [Wei et al., 2014] and CTEM-GEM is driven by GEM-MACH-GHG [Anselmo et al., 2010;
Robichaud and Ménard, 2014; Makar et al., 2015]. Details of the TBM runs are given in Table 1. A more detailed description of the configuration of the TBMs is given in Appendix A.

2.2. Flux Inversions

In addition to the NEE from the TBMs, posterior NEE fluxes from two flux inversion analyses are examined: one that assimilates boundary layer CO\textsubscript{2} observations (CT2016) and one that assimilates X\textsubscript{CO\textsubscript{2}} observations (GOSAT-Inv) from the Greenhouse Gases Observing Satellite (GOSAT). The motivation for employing two different inversion analyses is that flux estimates from these analyses are often divergent on regional scales. This is partially due to the sparsity of atmospheric CO\textsubscript{2} observations, which results in the observations strongly underconstraining NEE fluxes, although other factors, such as errors in model transport, further confound reliable NEE estimates. Thus, if features in optimized NEE are consistent between inversions, we have increased confidence in the robustness of the results.

For the first inversion analysis, we use optimized NEE from the National Oceanic and Atmospheric Administration’s CarbonTracker, version CT2016 [Peters et al., 2007, with updates documented at http://carbontracker.noaa.gov]. CT2016 optimizes NEE by assimilating in situ observations of boundary layer atmospheric CO\textsubscript{2}. It employs an ensemble Kalman filter approach to assimilate CO\textsubscript{2} with atmospheric chemical transport simulated by the TM5 offline atmospheric model [Krol et al., 2005]. For CT2016, TM5 is driven by ERA-Interim assimilated meteorology from the European Centre for Medium-Range Weather Forecasts (ECMWF), with a horizontal resolution of 3°×2° globally and 1°×1° in a nested grid over North America.

©2018 American Geophysical Union. All Rights Reserved.
In a second inversion, referred to as GOSAT-Inv, NEE fluxes are optimized by assimilating GOSAT $X_{\text{CO}_2}$ observations using the GEOS-Chem four-dimensional variational (4D-Var) data assimilation, with version v35 of the GEOS-Chem adjoint model [Henze et al., 2007]. The GEOS-Chem model (www.geos-chem.org) is a global 3-D chemical transport model driven by assimilated meteorology from the Goddard Earth Observing System (GEOS-5.2.0) of the NASA Global Modeling and Assimilation Office (GMAO). The native resolution of the model is $0.5^\circ \times 0.67^\circ$ with 72 vertical levels from the surface to 0.01 hPa, but here the model is run at a resolution of $4^\circ \times 5^\circ$ resolution with 47 vertical layers. Our model configuration is based on the set-up of Nassar et al. [2011], which was recently employed by Byrne et al. [2017] to examine the sensitivity of CO$_2$ surface flux constraints to observational coverage. Monthly ocean fluxes are from Takahashi et al. [2009], anthropogenic emissions are from Andres et al. [2016], and biomass burning emissions are from the Global Fire Emission Database GFEDv3 [van der Werf et al., 2006]. To optimize surface fluxes, the 4D-Var cost function is minimized as described in Deng et al. [2014] to retrieve monthly scaling factors for prior ocean and terrestrial biosphere fluxes in each grid cell. We use an assimilation window of nine months and keep posterior fluxes from the first six months, then shift the inversion period forward by six months. Using this method, optimized NEE spanning 2010–2013 is generated. Prior NEE fluxes are based on the posterior fluxes from CT2016. We calculate a mean seasonal cycle using 3-hourly fluxes from the period 2010-2013. For error statistics, we assign 16% error to fossil fuels, 38% error to biomass burning, 44% error to ocean fluxes, and 44% error to terrestrial ecosystems, following Deng et al. [2014].
For the GOSAT retrievals, we use version v3.5 of the NASA Atmospheric CO$_2$ Observations from Space (ACOS) GOSAT lite files from the CO$_2$ Virtual Science Data Environment (https://co2.jpl.nasa.gov/#mission=ACOS). Information on the ACOS retrieval algorithm is available in O’Dell et al. [2012] and Crisp et al. [2012]. All bias-corrected measurements from the TANSO-FTS shortwave infrared (SWIR) channel are selected, including ocean glint, high gain and medium gain nadir, which pass the quality flag requirement. We generate “super-obs” from the GOSAT retrievals by aggregating the observations to the grid size of our inversion. Error estimates are generated using the method described by Kulawik et al. [2016]. The reduction in error with aggregation can be calculated using the expression $error^2 = a^2 + b^2/n$, where $a$ represents systematic errors that do not decrease with averaging, $b$ represents random errors that decrease with averaging, and $n$ represents the number of satellite observations that are averaged [Kulawik et al., 2016]. Kulawik et al. [2016] give $a = 0.8$ ppm and $b = 1.6$ ppm as mean Northern Hemisphere geometric (co-located) values for GOSAT, and these are the values that are used in our inversion analyses.

2.3. GOME-2 SIF

We use two different GOME-2 SIF products: NASA Level 2 GOME-2 version 26 (V26) 740 nm terrestrial chlorophyll fluorescence data [NASA-SIF; Joiner et al., 2013, 2016] and the GFZ Postdam product [GFZ-SIF; Köhler et al., 2015]. We have selected observations spanning the period 2007-2014. A “daily correction” is performed to estimate daily average SIF from the instantaneous measurements (see supplementary material of Frankenberg et al. [2011b] for details of this calculation). The observations are then ag-
aggregated spatially to a $2^\circ \times 2.5^\circ$ grid and temporally to week of year by calculating the median value.

The relationship between SIF and GPP has been found to be dependent on vegetation type [Ganter et al., 2012], therefore, we examine different vegetation types separately. We examine GPP over the six northern vegetation regions shown in Figure 1: evergreen needle leaf forests (ENF), deciduous needleleaf forests (DNF), southern mixed forests, northern mixed forests, grasslands and croplands. The vegetation regions are based on the vegetation types from the Moderate Resolution Imaging Spectroradiometer (MODIS) International Geosphere-Biosphere Programme (IGBP) land cover type classification product [Friedl et al., 2010; Channan et al., 2014]. On a $2^\circ \times 2.5^\circ$ grid, each grid cell is assigned a given vegetation type if more than 50% of the grid is made up of a single vegetation type. Mixed forests occur in two distinct spatial regions in Eurasia and North America. For this reason, we split this category into two groups: ”northern mixed forests” in Eurasia and ”southern mixed forests” in North America.

To obtain a seasonal cycle for a given vegetation type, the spatial mean is calculated for each region. The seasonal cycle is then smoothed using a 3-week running mean. The mean offset of the SIF seasonal cycle is removed by subtracting the mean SIF value outside of the growing season. We ensure that the period is outside the growing season by checking that the time period is also outside of the growing season for FLUXCOM GPP. Finally, the seasonal cycle of SIF is normalized so that the integrated annual total SIF value equals 1. The normalization is required because the scaling between SIF and GPP is not well known and because TBMs produce a large spread in the magnitude of GPP (see Figure
Thus it is necessary to normalize annual GPP to compare the seasonality. The same scaling and averaging is applied to model GPP.

We construct an error estimate for SIF from the random noise in the seasonal cycle and the error associated with removing the seasonal cycle offset. To estimate the random noise, we calculate the SIF seasonal cycle for 2007-2014 for each year individually, for both NASA and GFZ data products, and take the spread as the uncertainty. In reality, this spread is due to random error combined with interannual variability in SIF, thus it provides an upper bound on uncertainty. We estimate the error due to removing the offset as being the range of SIF values over the dates used to calculate the offset; again this provides an upper estimate of the errors. These errors are summed in quadrature to obtain the total error. See Sec. S1 of the supplementary material for more information on GOME-2 SIF error characterization. Note that these errors do not include systematic errors related to the retrieval or differences between SIF and GPP. Systematic errors are not well characterized and could impact the seasonality of SIF. During the winter months, snow cover and large air masses could introduce biases. Cloud cover is also a potential source of bias, as SIF is not observed under very cloudy conditions. However, Köhler et al. [2015] found that the cloud cover threshold did not have a large impact on the temporal patterns of SIF.

2.4. Atmospheric $X_{CO_2}$

2.4.1. TCCON $X_{CO_2}$

The TCCON is a network of ground-based Fourier transform spectrometers that record spectra of the sun in the near-infrared, from which $X_{CO_2}$ columns are retrieved [Wunch et al., 2011]. We examine $X_{CO_2}$ from Sodankylä [Kivi et al., 2017; Kivi and Heikkinen, 2018 American Geophysical Union. All Rights Reserved.
2016], Białystok [Deutscher et al., 2017], Orléans [Warneke et al., 2017] and Park Falls [Wennberg et al., 2017]. TCCON data were obtained from the TCCON Data Archive, hosted by CaltechDATA [https://tccondata.org].

2.4.2. Simulated $X_{CO_2}$

To simulate $X_{CO_2}$ based on the NEE fluxes from the TBMs, we use the forward model component of the GEOS-Chem adjoint, at a horizontal resolution of $2^\circ \times 2.5^\circ$ with 47 vertical levels. The terrestrial biosphere fluxes are input at 3 hour resolution. We use annually repeating biospheric fluxes by averaging the fluxes over the years given in Table 1. All other fluxes are identical for each simulation, and are the same as those used in the inversion analyses described in Sec. 2.2. We simulate five years of $CO_2$ fields (2008-2012) with a one-year spin up period (2007) for each set of terrestrial biosphere fluxes. For comparison with TCCON, the simulated atmospheric $CO_2$ concentrations are sampled at the time (in 3-hour time intervals) and grid box in which TCCON measurements occur. The TCCON a priori information and averaging kernels are used to generate $X_{CO_2}$ (using the method described in Wunch et al. [2011]).

2.4.3. Seasonal Cycle Fit

The $X_{CO_2}$ time-series is detrended and an annual $X_{CO_2}$ cycle is found by fitting a function of the form:

$$f(x) = a_0 + \sum_{k=1}^{4} a_k \cos(2\pi k x) + b_k \sin(2\pi k x)$$

(2)

where $x$ is the fraction of the year. Equation (2) is a truncated Fourier series and is similar to the NOAA seasonality fitting function that is commonly used to fit the mean annual cycle [Thoning et al., 1989]. To estimate uncertainty in the fit, a Monte Carlo approach is used, in which a set of 50 fits are performed using randomly generated initial parameters.

©2018 American Geophysical Union. All Rights Reserved.
and the standard deviation of the resulting curves is used to estimate the uncertainty. We then take the mean values of the curves as the best fit line. Note that this is only the uncertainty on the fit and does not include other sources of error, such as measurement error.

2.5. FLUXCOM

FLUXCOM products are generated using upscaling approaches based on machine learning methods that integrate FLUXNET site level observations, satellite remote sensing, and meteorological data \cite{Tramontana et al., 2016; Jung et al., 2017}. Jung et al. [2017] generate \( R_e \) products using several machine learning methods. For this study, we downloaded the products generated using random forests (RF), multivariate regression splines (MARS) and artificial neural networks (ANN) at daily resolution from the Data Portal of the Max Planck Institute for Biochemistry (https://www.bgc-jena.mpg.de). The mean seasonal cycle over 2008-2012 is calculated for each product.

2.6. Surface Air and Soil Temperature

Surface air and soil temperatures are required for investigating the heterotrophic respiration (\( R_H \)) produced by TBMs (Sec. 3.4.1). Near-surface air temperature (\( T_{\text{air}} \)) is taken as the lowest atmospheric level of the assimilated meteorology from GEOS-5 after interpolation to \( 2^\circ \times 2.5^\circ \) with 47 vertical levels. This is the same meteorology used to run the GEOS-Chem model. Soil temperature (\( T_{\text{soil}} \)) is taken from the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) \cite{Gelaro et al., 2017}. We take the soil temperature to be the maximum temperature of the top three layers of soil (TSOIL1, TSOIL2 and TSOIL3), which covers a depth down to 0.4 m \cite{Koster et al., 2000}.

©2018 American Geophysical Union. All Rights Reserved.
3. Results

3.1. Model Fluxes

The mean GPP, $R_e$, and NEE fluxes for the TBMs, FLUXCOM, and inversions over all land between 39°N and 65°N are shown in Figure 2. This latitude range is used because, based on GEOS-Chem tagged tracer simulations, NEE over this region produces 85% of the amplitude of the seasonal cycle at Park Falls, 79% at Orléans, 87% at Białystok and 76% at Sodankylä. The mean seasonal cycle is calculated by averaging over the time periods listed in Table 1 to remove interannual variability. CT2016 is averaged over the period 2007–2012, and GOSAT-Inv is averaged over the period 2010–2013. Fluxes are defined so that negative values represent removal of CO$_2$ from the atmosphere and positive values indicate emission to the atmosphere.

The TBMs have large differences in the magnitude of total annual GPP and $R_e$, consistent with the results from Huntzinger et al. [2012]. The start of the growing season is quite variable between models, with JULES beginning earliest and SiB3 latest. For the timing of the end of the growing season, the TBMs are split into two groups: the TBMs with diagnostic phenology (CASA and SiB3) and FLUXCOM algorithms (ANN, MARS and RF) have growing seasons that end earlier, whereas the TBMs with prognostic phenology (CTEM and JULES) end two to four weeks later. For $R_e$, the start of the spring increase begins earliest for CASA and latest for SiB3. The timing of peak $R_e$ is also quite variable between models, ranging from early June (CASA) to late July (SiB3). The diagnostic TBMs and FLUXCOM show similar timing in decreasing $R_e$ throughout the fall, whereas the prognostic TBMs shows significantly higher $R_e$ throughout the fall, mirroring what is seen in GPP. The are also significant differences between models in the magnitude of
winter $R_e$. Prognostic TBMs give the highest winter $R_e$, followed by diagnostic TBMs, whereas FLUXCOM estimates are the lowest. The two sets of CTEM fluxes driven by different meteorology show some marked differences in fluxes, with CTEM-CRU GPP and $R_e$ having larger magnitude than CTEM-GEM.

For NEE, all models produce a positive flux to the atmosphere in the winter months and net drawdown into the terrestrial biosphere during most of the growing season. However, there are differences between models in the timing of the period of net carbon uptake. The start of net uptake ranges between early April and May and the end of net uptake occurs between late August and early October. In contrast, there is close agreement in the seasonality of NEE between CT2016 and GOSAT-Inv. The most notable difference between the inversions is that GOSAT-Inv has somewhat weaker drawdown from June through September, and reaches peak drawdown earlier than CT2016. The inversion NEE fluxes generally fall in the middle of the modeled NEE fluxes from the TBMs and FLUXCOM during the spring and fall, but produce quite strong drawdown during the summer.

All three FLUXCOM NEE products are known to produce unrealistically large annual net sinks [Tramontana et al., 2016; Jung et al., 2017]. However, both MARS and RF NEE show reasonable agreement with the flux inversion NEE through most of the growing season (June–September). This suggests that low $R_e$ fluxes outside the growing season and enhanced drawdown during the early spring could cause the MARS and RF NEE annual bias. In contrast, ANN NEE shows reasonable agreement with the flux inversion NEE during the winter but weaker drawdown during the growing season, suggesting a different source for the NEE annual bias.

©2018 American Geophysical Union. All Rights Reserved.
3.2. Comparing Model GPP and GOME-2 SIF

We compare the normalized seasonal cycle of SIF and model GPP over the six different vegetation regions (Fig. 3). The normalization is required because the scaling between SIF and GPP is not well constrained (see Sec. 2.3), so we are not able to evaluate the magnitude of GPP using SIF observations.

The two GOME-2 SIF products (NASA and GFZ-Potsdam) are in close agreement throughout the year and for all vegetation regions. Differences between the SIF products are always less than the estimated uncertainties, and less than differences between SIF and modeled GPP. For the remainder of this study, comparisons will be performed with the NASA product.

Differences between the modeled GPP and NASA GOME-2 SIF are variable, although differences between model GPP and NASA GOME-2 SIF are generally consistent across the vegetation regions. For example, the growing season in NASA GOME-2 SIF ends earlier by several weeks than in the JULES GPP fluxes for all vegetation regions. The two TBMs with diagnostic phenology and FLUXCOM GPP deviate the least from the SIF seasonal cycle. The closer agreement for the diagnostic TBMs relative to the prognostic TBMs suggests that the assimilation of vegetation indices improves GPP fluxes, as expected, although differences in the driving meteorology could play a role. The normalized seasonal cycle of SiB3 GPP is always within the SIF uncertainty for all vegetation types and has a mean root-mean-square (RMS) difference of 0.0031. The normalized seasonal cycle of CASA GPP falls within the SIF uncertainties everywhere except in the fall for DNF and northern mixed forest regions, where CASA shows a more rapid decrease in GPP. CASA GPP has a mean RMS difference of 0.0037 across the vegetation

©2018 American Geophysical Union. All Rights Reserved.
regions. RMS differences for the FLUXCOM are 0.0038, 0.0038, 0.0044 for ANN, MARS and RF respectively. For all FLUXCOM algorithms, GPP is slightly phase shifted earlier in the year for all vegetation types, with the seasonal cycle starting and ending about one week earlier. The TBMs with prognostic phenology, CTEM and JULES, have larger differences between the seasonal cycle of modeled GPP and SIF. These prognostic models suggest growing seasons that are too long by several weeks compared with SIF across the vegetation regions studied here. See Sec. S2 of supplementary material for additional details.

3.3. Comparing Model NEE and TCCON $X_{CO_2}$

The mean seasonal cycles of observed and modeled $X_{CO_2}$ at Sodankylä, Białystok, Orleans, and Park Falls are shown in Figure 4. All of the NEE fluxes reproduce the general shape of the seasonal cycle, with maximum $X_{CO_2}$ in the early spring and minimum in the late summer, although in some cases there are significant differences from the TCCON data in amplitude and phase. The posterior NEE fluxes from the two inversions produce the closest agreement with TCCON. This is expected because the inversions assimilate atmospheric CO$_2$ measurements (which are independent of TCCON observations). Of the two inversions, GOSAT-Inv has a smaller RMS difference of 0.21 ppm across the four sites, compared to 0.40 ppm for CT2016. However, the difference in RMS between GOSAT-Inv and CT2016 is not considered meaningful, because the difference could be due to transport differences between GEOS-Chem and TM5 (see Sec. S3 of supplementary materials). Closer agreement for GOSAT-Inv is expected because the inversion was performed with the same chemical transport model used to simulate TCCON $X_{CO_2}$. The good agreement between the inversions and TCCON is reassuring because

©2018 American Geophysical Union. All Rights Reserved.
both inversions assimilate measurements that are independent of TCCON. These results show good agreement despite uneven observational coverage, errors in model transport, and biases in assimilated observations, all of which can strongly impact inversion analyses [Byrne et al., 2017; Baker et al., 2006; Miller et al., 2018]. Of the TBMs, SiB3 and CASA give the best agreement with TCCON, with RMS differences of 0.50 ppm and 0.58 ppm, respectively. SiB3 gives a smaller seasonal cycle amplitude than TCCON, such that the annual minimum is about 1 ppm higher across the sites examined here. SiB3 also gives an earlier drawdown of X$_{CO_2}$ than suggested by TCCON data at Sodankylä, Białystok and Park Falls. CASA gives a seasonal cycle that lags the TCCON data; the lag is largest at Orléans, where CASA is about two weeks later than the measurements. However, it is unclear whether these differences are significant, as they are approximately the same order of magnitude as transport errors. See Sec. S3 of supplementary materials for more details of transport error quantification, which is based on differences in X$_{CO_2}$ simulated by GEOS-Chem and TM5 as-well as previous studies [Houweling et al., 2010; Basu et al., 2011; Keppel-Aleks et al., 2011, 2012; Barnes et al., 2016].

Of the TBMs with prognostic phenology, CTEM-GEM has the lowest RMS difference relative to TCCON (0.73 ppm), whereas CTEM-CRU and JULES have RMS differences of 1.39 ppm and 1.18 ppm, respectively. Simulated X$_{CO_2}$ using CTEM-CRU and CTEM-GEM NEE show quite different seasonal cycle shapes and phases. In comparison to CTEM-GEM, CTEM-CRU has a more rapid spring drawdown, which is delayed by ten days at all sites. The amplitude of the seasonal cycle is larger for CTEM-CRU than for CTEM-GEM by about 2 ppm at Orléans and Park Falls. In comparison to TCCON, CTEM-GEM produces better agreement in the shape and timing of the X$_{CO_2}$ seasonal
cycle than CTEM-CRU, and is within 2 ppm of the TCCON data at all sites. Simulated \( X_{\text{CO}_2} \) using JULES NEE produces a seasonal cycle which is phase shifted early relative to TCCON at all sites by two to four weeks. JULES underestimates the amplitude of the seasonal cycle by 3 ppm at Park Falls but overestimate the amplitude by 2.5 ppm at Sodankylä.

### 3.4. Comparison of GPP and NEE

Figure 5 shows the mean RMS difference between model GPP and GOME-2 SIF versus the mean RMS difference between model-NEE-based \( X_{\text{CO}_2} \) and the TCCON \( X_{\text{CO}_2} \). The TBMs with diagnostic phenology, SiB3 and CASA, have smaller SIF and TCCON RMS differences than the TBMs with prognostic phenology. This suggests that assimilating phenology improves both GPP and NEE fluxes. However, different driving meteorology was used for the diagnostic TBMs (driven by MERRA) and prognostic TBMs (driven by NCEP-CRU and GEM-MACH-GHG), thus it is unclear how much the differences in driving meteorology could have contributed to the differences between the prognostic and diagnostic TBM fluxes. It is possible that the driving meteorology is partially responsible for the better agreement with SIF and \( X_{\text{CO}_2} \) found with the diagnostic TBMs.

For the prognostic TBMs, the relative agreement between modeled GPP and GOME-2 SIF can be quite different than between model-NEE-based \( X_{\text{CO}_2} \) and TCCON \( X_{\text{CO}_2} \). For example, \( X_{\text{CO}_2} \) simulated with CTEM-GEM NEE has a smaller RMS with respect to TCCON than CTEM-CRU, but CTEM-GEM GPP has a larger RMS with respect to SIF than CTEM-CRU. This shows two things. First, it indicates that a small GPP RMS difference does not necessarily predict small \( X_{\text{CO}_2} \) RMS difference. This suggests that there are compensating discrepancies in GPP and \( R_e \) that improve the NEE fluxes. Second, it
shows that TBM fluxes are highly sensitive to the driving meteorology. High sensitivity to the driving meteorology has previously been reported for other TBMs [Poulter et al., 2011]. Differences between CTEM-GEM and CTEM-CRU fluxes are primarily due to differences in moisture between the NCEP-CRU and GEM-MACH-GHG [Badawy et al., 2018]. NCEP-CRU is wetter than GEM-MACH-GHG, which increases GPP and $R_e$ fluxes in CTEM. These results highlight the difficulty in validating TBMs using only constraints on NEE. This remains a major difficulty in relating optimized NEE from flux inversions to TBM errors or deficiencies.

3.4.1. Comparing SiB3 and CASA

As discussed above, GPP fluxes and model-NEE-based $X_{CO_2}$ from CASA and SiB3 show close agreement with SIF and TCCON data, respectively. Figure 6 compares the seasonality of GPP (times negative one), $R_e$ and NEE between CASA and SiB3 over 39°-65° N. The seasonality of GPP is similar for CASA and SiB3 but the seasonality of $R_e$ and NEE shows significant differences. $R_e$ peaks about a month earlier in CASA than in SiB3, causing CASA NEE to peak about a month later than in SiB3. Similar results are found for individual vegetation regions (see Sec. S4 of supplementary materials). The differences between CASA and SiB3 found here are consistent with the results of Messerschmidt et al. [2013], who found that differences in seasonality of NEE between TBMs were primarily due to the differential phasing of $R_e$ with respect to GPP.

The difference in timing of $R_e$ between SiB3 and CASA could be explained by differences in the parameterizations of heterotrophic respiration ($R_H$) in the TBMs. SiB3 uses a “zero order” parameterization in which

$$R_H \propto f(T) \cdot f(M).$$

(3)

©2018 American Geophysical Union. All Rights Reserved.
So that $R_H$ is only dependent on soil temperature through a function,

$$f(T) = 2^{(T_{soil} - 298)/10},$$

(4)

and soil moisture through a function $f(M)$ [Denning et al., 1996]. This parameterization closely follows the seasonal cycle of soil temperature and peaks at the same time. Figure 6b shows SiB3 $R_e$ and $f(T)$ scaled to have the same annual flux as $R_e$. Clearly, the seasonal cycle of SiB3 $R_e$ and $f(T)$ align closely. In contrast, the CASA $R_e$ curve is phase shifted earlier in the year relative to the $f(T)$ curve (Figure 6a). CASA uses a “first order” parameterization [Randerson et al., 1996] that has an additional dependence on the carbon pool size ($C$) available for $R_H$, resulting in a phase shift in $R_e$ earlier in the season [Randerson et al., 1996]. The reason is that leaf and fine root litter pool grow at the end of the growing season. Low temperatures throughout the winter prevent significant respiration, thus, this pool is near its maximum in the early spring. This results in a larger quantity of available substrate for respiration early in spring [Randerson et al., 1996].

3.5. Estimating $R_e$

Ideally, we would like to optimize $R_e$ by performing a “flux inversion” that assimilates atmospheric CO$_2$ and SIF observations. This would involve first forward modeling SIF and X$_{CO_2}$, and then optimizing GPP and $R_e$ by simultaneously assimilating SIF and X$_{CO_2}$ observations. This is a complicated task, although tools to do this are under development [e.g., Schuh et al., 2016]. Instead, we take a simpler approach that requires additional assumptions. We assume NEE and GPP are known, which allows us to simply calculate $R_e$. We take inferred NEE fluxes, which have been optimized using CO$_2$ observations and produce X$_{CO_2}$ values that agreed closely with TCCON data, to be our “true” NEE.
Modeled GPP from CASA, SiB3, and FLUXCOM are taken as the “true” GPP for our calculation. Given that the normalized seasonal cycles of SiB3, CASA and FLUXCOM GPP were in close agreement with SIF, it is reasonable to assume that the differences in NEE between the TBM s and inversions are primarily due to differences in the seasonal cycle of $R_e$ and the magnitude of GPP. Note that it is not possible to use SIF in place of the modeled GPP because the scaling between SIF and GPP is not well known. The magnitude of GPP from SiB3 and CASA differ by about 45% at the peak of the growing season and FLUXCOM lies between the two TBMs (Figure 2). Any error in the magnitude of GPP will also be projected onto our optimized respiration estimate. Here we examine “optimized” respiration in several steps. First, we calculate an “optimized” $R_e$ over the six vegetation types combined using the equation:

$$optR_{inv-mod} = NEE_{inv} - GPP_{mod}.$$  

Therefore, $optR_{inv-mod}$ is the respiration, which, combined with CASA, SiB3 or FLUXCOM GPP, results in the inversion NEE. Second, we examine the sensitivity of $optR_{inv-mod}$ to the magnitude of GPP by calculating $optR_{inv-mod}$ after scaling GPP over a range of magnitudes, spanning the range of SiB3, CASA and FLUXCOM GPP products. Finally, possible ecological implications for the seasonality of $optR_{inv-mod}$ are discussed.

### 3.5.1. Optimized $R_e$

GPP, $R_e$ and $optR_{inv-mod}$ for SiB3, CASA and FLUXCOM are shown in Figure 7. The $optR_{inv-mod}$ curves show several features consistent across the ensemble. In particular, $optR_{inv-mod}$ generally gives a broader seasonal cycle with a less pronounced summer maximum than $R_e$ produced by SiB3, CASA and FLUXCOM. However, comparing the actual fluxes is difficult due to the differences in magnitude and the fact that the annual net flux...
drawdown is different between TBMs. All FLUXCOM algorithms overestimate the net annual drawdown, such that the total annual $optR_{inv−FLUXCOM}$ is greater than that modeled by the algorithms. In contrast, SiB3 fluxes are generated by assuming the annual net NEE flux is approximately zero at each grid cell, and thus the magnitude of $optR_{inv−SiB3}$ is smaller than $R_e$ produced by the TBM.

To simplify these comparisons, the seasonal cycle of GPP, $R_e$ and $optR_{inv−mod}$ are normalized by the annual total flux. After normalization, more features become clear (Figure 7b,d,f,h). As with the comparisons with SIF, close agreement is found in the seasonality of normalized GPP between the models. In contrast, there are larger differences in normalized $R_e$ between SiB3, CASA and FLUXCOM. In general, differences between $optR_{inv−mod}$ and model $R_e$ are consistent with a broader seasonal cycle in $optR_{inv−mod}$. Normalized $optR_{inv−mod}$ is systematically lower in June–July and higher during October relative to modeled $R_e$. Optimized $R_e$ for CASA ($optR_{inv−CASA}$) shows the largest differences from the mean. This behavior is due, primarily, to the larger magnitude of GPP in CASA relative to SiB3 and FLUXCOM.

### 3.5.2. Sensitivity to GPP magnitude

If Eq. (5) is re-written as:

$$optR_{inv−mod} = -GPP_{mod} \left( 1 - \frac{NEE_{inv}}{GPP_{mod}} \right), \quad (6)$$

it is clear that $optR_{inv−mod}$ will become closer to the shape of GPP as the magnitude of GPP increases. Here we calculate $optR_{inv−mod}$ after scaling SiB3, CASA and FLUXCOM GPP over a range of values to examine the sensitivity of our results. Over the vegetation regions examined here, SiB3 GPP gives an uptake of 23 Pg yr$^{-1}$, ANN gives 23 Pgyr$^{-1}$,
MARS gives 28 Pg yr$^{-1}$, RF gives 29 Pg yr$^{-1}$, and CASA gives 32 Pg yr$^{-1}$. Therefore, we scale GPP to vary over the range 23–32 Pg yr$^{-1}$ and recalculate $optR_{inv-mod}$.

The resulting curves are shown in Figure 8. The $optR_{inv-mod}$ curves are similar when GPP is scaled to the same annual total for SiB3, CASA and FLUXCOM. For the range of annual total GPP examined here, $optR_{inv-mod}$ retains a broad summer maximum. However, a summer peak in respiration becomes more defined as GPP is increased.

### 3.5.3. Implications of $optR_{inv-mod}$

The $optR_{inv-mod}$ curves give systematically broader summer peaks in $R_e$ than are modeled by the diagnostic TBMs: CASA and SiB3. Here we examine how $R_e$ fluxes in these TBMs could be changed to bring the TBMs in agreement with $optR_{inv-mod}$. The objective is to determine whether realistic changes in TBM parameters could produce the seasonality found in $optR_{inv-mod}$, or whether changes to model equations are required. For this analysis, we consider the model equations for SiB3 and CASA,

\[
R_e = R_a + R_H, \tag{7}
\]
\[
R_a = -c_0 \cdot GPP, \tag{8}
\]
\[
R_H = C \cdot f(T) \cdot f(M), \tag{9}
\]
\[
f(T) = Q_{10}^{(T-c_1)/10}, \tag{10}
\]

where $R_a$ is autotrophic respiration, $R_H$ is heterotrophic respiration, and $C$ is the leaf and fine root litter carbon pool size. For CASA, the constants given in Eq. (7)–(10) are: $c_0 = 0.5$, $Q_{10} = 1.5$, $c_1 = 273.15$, and $T = T_{air}$ [Randerson et al., 1996]. For SiB3, the constants given in Eq. (7)–(10) are: $c_0 = 0.6$, $Q_{10} = 2$, $c_1 = 298$, and $T = T_{soil}$ [Denning
et al., 1996]. Note that $R_H$ in SiB3 does not have leaf and fine root litter carbon pool dependence, we have artificially introduced this dependence for this analysis.

In Sec. 3.4.1, it was found that the seasonality of heterotrophic respiration ($R_H$) has a high sensitivity to the leaf and fine root litter carbon pool size. In the following discussion, we show that adding realistic changes to the leaf and fine root litter carbon pool can bring $R_e$ fluxes into closer agreement with $optR_{inv-mod}$, but does not completely resolve the differences. First, $R_H$ is calculated using Eq. (7) and (8), and assuming $R_e = optR_{inv-mod}$, then, $R_H/f(T)$ is calculated. $R_H/f(T)$ is proportional to the leaf and fine root litter carbon pool, neglecting soil moisture dependence. Figure 9 shows $optR_{inv-mod}$, $R_H$, and $R_H/f(T)$ based on CASA and SiB3 model parameterizations. The calculated $R_H$ increases in the spring (Mar–May). At the start of June, there is a rapid decrease in $R_H$ which remains steady until mid-July when $R_H$ increases. $R_H$ remains high until late November when $R_H$ decreases into winter. $R_H/f(T)$ indicates that the carbon pool generally increases from July through October and decreases through the rest of the year. Notably, there is a very rapid decrease in $R_H/f(T)$ during the spring. This rapid decrease corresponds to the decrease in $R_H$. These result suggests that the leaf and fine root litter carbon pool is largely depleted in early June. This lack of substrate then results in a significant decrease in $R_H$ at this time. $R_H$ then remains low until this carbon pool begins to increase in mid July.

Overall, the calculated seasonal cycle of $R_H$ and $R_H/f(T)$ suggests that there is an abundance of substrate for heterotrophs to consume early in the growing season, which is largely depleted by June, resulting in reduced $R_H$ over the summer. In general, the seasonal cycle of $R_H$ and $R_H/f(T)$ calculated here seem plausible, however, it is unlikely
that the carbon pool would begin to increase in July and is more likely this carbon pool would not increase until the fall [Randerson et al., 1996]. Thus, plausible changes in the seasonal cycle of the leaf and fine root litter carbon pool could partially account for the differences found between \( R_e \) and \( optR_{\text{inv-mod}} \), but it seems unlikely that this could be the sole cause of the difference, as some unphysical changes to the carbon pool are required. It is possible that neglecting the soil moisture term impacts the result here, as soil moisture depletion during July could suppress \( R_e \) until the fall, resulting in more realistic seasonality.

The analysis presented here assumes that the model equations (Eq. (7)–(10)) are correct. However, in some cases the true biosphere is known to deviate from these expressions. In many TBMs, including CASA and SiB3, \( R_a \) is assumed to be a constant fraction of GPP implying a constant carbon use efficiency. However, observations suggest substantial variability in the carbon use efficiency over the year [Arneth et al., 1998; DeLucia et al., 2007; Heskel et al., 2013; Wehr et al., 2016; Tcherkez et al., 2017]. Errors in model parameterizations, such as a constant carbon use efficiency, will result in systematic errors in modeled fluxes (discussed in Sec. 4). Therefore, errors in model parameterization could also explain differences between \( R_e \) and \( optR_{\text{inv-mod}} \).

### 3.5.4. Continental Scales

So far, the analysis for \( optR_{\text{inv-mod}} \) has focused on all land between 39°–65° N. In this section, \( optR_{\text{inv-mod}} \) is calculated separately over North America (51°–167° W), Europe (12° W–41° E), and Asia (41°–180° E) between 39°–65° N to evaluate the continental differences in \( optR_{\text{inv-mod}} \). Figure 10 shows the normalized seasonal cycle of GPP, \( R_e \), and \( optR_{\text{inv-mod}} \) for each continent (absolute fluxes are given in Sec. S5 of the supplementary material).
material), revealing significant differences in the seasonality of model GPP between the continents (Fig. 10a–c). Europe has the longest growing season: normalized model GPP increases earlier in spring than for the other continents, peaks in June and then slowly decreases from July to November. For North America and Asia, modeled GPP peaks in early July. Asia has the shortest growing season of the three continents.

Normalized model $R_e$ indicates that Europe has the broadest seasonal cycle of the three continents, while North America and Asia have narrower summer peaks in modeled $R_e$ (Fig. 10d–f). Similarly, normalized $optR_{inv-mod}$ indicates a broader seasonal cycle for Europe relative to North America and Asia (Fig. 10g–i). Comparing $optR_{inv-mod}$ and modeled $R_e$, we find that there is consistency between modeled $R_e$ and $optR_{inv-mod}$ throughout the spring and summer for Europe relative to North America and Asia, but, $optR_{inv-mod}$ suggests enhanced $R_e$ in the late fall (Fig. 10k). For North America and Asia there are large differences between $optR_{inv-mod}$ and modeled $R_e$ (Fig. 10j,l). For both continents, differences between $optR_{inv-mod}$ and modeled $R_e$ are similar to those seen for the entire northern extratropics, with $optR_{inv-mod}$ suggesting reduced $R_e$ in the early summer and enhanced $R_e$ in the fall. Systematic differences between $optR_{inv-mod}$ and modeled $R_e$ are largest for Asia, for which $optR_{inv-mod}$ suggests systematically lower $R_e$ throughout June and into early July, and systematically higher $R_e$ from mid-September to early November.

Why do $optR_{inv-mod}$ and model $R_e$ generally show consistency for Europe, but large systematic differences for Asia? The answer may be linked to the differences in vegetation and climate between the continents. Europe has a milder climate than Asia due to the influence of the Gulf Stream. Furthermore, Europe has a large fraction of croplands...
while Asia has a high fraction of evergreen and deciduous needleleaf forests. The simplest explanation for the difference is that Europe has a milder climate and thus the TBMs and FLUXCOM suggest a broader season in $R_e$ than for Asia. Since $optR_{inv-mod}$ generally indicates a broader seasonal cycle in $R_e$ across the northern extratropics, this will result in a smaller difference for Europe since the seasonality of $R_e$ is already broad.

4. Discussion

The differences between $optR_{inv-mod}$ and FLUXCOM $R_e$ may indicate biases in the methods used to model GPP and $R_e$ at FLUXNET sites. At these sites, partitioning methods are required to decompose observed NEE fluxes into GPP and $R_e$ components. Standard methods perform the partitioning using hypothesized responses of GPP and/or $R_e$ to light, water, and/or temperature [Wehr et al., 2016]. These methods are also applied to generate FLUXCOM products [Tramontana et al., 2016]. Recently, Wehr et al. [2016] used isotopic measurements to determine daytime $R_e$ in a temperate deciduous forest (Harvard forest) and found that daytime $R_e$ was only about half as large as night-time NEE during June–July but roughly equal to night-time NEE during August–September. Standard partitioning methods do not account for this variability in daytime $R_e$, thus Wehr et al. [2016] suggest that FLUXNET $R_e$ fluxes are over-estimated in June–July relative to August–September. Consistent with this result, reduced fluxes during June–July are obtained for $optR_{inv-mod}$ relative to FLUXCOM $R_e$. This suggests that reduced daytime $R_e$ fluxes during June–July could be present across much of the northern mid-latitudes, particularly in North America and Asia. The ecological explanation for reduced June–July daytime $R_e$ suggested by Wehr et al. [2016] is the “Kok effect”, wherein leaf respiration is inhibited by light [Heskel et al., 2013]. Therefore, this effect is largest during
June–July when insolation is greatest. However, it should be noted that Wehr et al. [2016] also found reduced GPP during June–July (to balance NEE). In contrast, we find strong agreement in the seasonality of FLUXCOM GPP and SIF.

The continental-scale differences between $optR_{inv-mod}$ and model $R_e$ found in this study are consistent with those expected from the Kok effect. The magnitude of the Kok effect has been found to depend on the plant species and ecosystem [Heskel et al., 2013; Tcherkez et al., 2017]. As pointed out by Tcherkez et al. [2017], the inhibition of daytime respiration was found to be small for a multi-site study of European grasslands [Gilmanov et al., 2007], but significant in a North American forest [Jassal et al., 2007]. Furthermore, the impact of the Kok effect has been suggested to be larger for evergreen vegetation than for deciduous vegetation [Wohlfahrt et al., 2005; Heskel et al., 2013]. Therefore, one could expect that the Kok effect would be larger in temperate Asia than Europe, which has more evergreen forests but less cropland. Consistent with this, $optR_{inv-mod}$ suggests greater reductions in $R_e$ in June–July in Asia than in Europe (Sec. 3.5.4).

The inhibition of daytime $R_a$ due to the Kok effect also has implications for TBMs, as it implies variability in the carbon use efficiency throughout the year. In a recent review, Tcherkez et al. [2017] concluded that leaf day respiration should be regarded as a central actor of plant carbon-use efficiency. Furthermore, He et al. [2018] argue that understanding the mechanisms behind spatiotemporal changes in $R_a$ is critical for better quantifying global carbon use efficiency. However, as discussed in Sec. 3.5.3, many TBMs assume constant carbon use efficiency throughout the year. These results suggest that this assumption may introduce a systematic bias of high $R_a$ in June–July relative to August–September in TBMs, and may explain why differences in the leaf and fine root
litter carbon pool between SiB3 and CASA could not fully explain the differences between model \( R_e \) and \( \text{opt} R_{\text{inv-mod}} \).

Enhanced fall \( R_e \) recovered in \( \text{opt} R_{\text{inv-mod}} \) provides an interesting parallel with a recent study of the Alaskan carbon cycle by Commane et al. [2017], who used aircraft and tower atmospheric CO\(_2\) observations and GOME-2 SIF observations to show that \( R_e \) fluxes from Alaskan tundra are significant during October–December. We obtain similar enhanced fall \( R_e \) over our much larger study region, with the largest enhancement of \( R_e \) in Asia from mid-September to mid-November. These results suggest that fall \( R_e \) fluxes are larger across boreal and northern regions than has previously been appreciated. Precisely why we obtain enhanced fall \( R_e \) relative to SiB3, CASA and FLUXCOM is unclear. Commane et al. [2017] showed that TBMs do not represent fall respiration well, especially when soil temperatures are near 0 °C. They suggested that during the zero curtain period, when the active layer is freezing from above and below, microbes continue to metabolize in the subsurface as long as liquid water is present [Zona et al., 2016], and that this process can persist for months after the surface is frozen and snow covered. It is unclear if a similar process could explain the enhanced fall \( R_e \) in our more southern domain (39°-65° N). Monson et al. [2006] showed that snow cover during the fall can insulate the soils, producing enhanced \( R_e \). This mechanism could provide enhanced \( R_e \) fluxes in the northern parts of our domain. A second possibility is that the size of the leaf and fine root litter carbon pool has an impact on the fall \( R_e \). The comparison presented in Sec. 3.5.3 suggests that this carbon pool may increase rapidly over the fall and peak in mid-October. If this is the case, it would provide a large quantity of substrate for heterotrophic respiration in
the mid to late fall, and provide a possible explanation for the enhanced rates of fall $R_e$ found in this study.

4.1. Remaining Challenges

One challenge with exploiting SIF and $X_{CO_2}$ data is the differences in the scales on which the two types of observations provide information about CO$_2$ surface fluxes. The footprints of SIF observations are highly localized; the observed SIF is representative of the footprint of the satellite. In contrast, an $X_{CO_2}$ observation has a large surface NEE footprint. On seasonal timescales, variations in $X_{CO_2}$ are driven by the meridional flux distribution [Keppel-Aleks et al., 2011]. Therefore, the scales that can be examined by combining SIF and $X_{CO_2}$ observations are limited by the scales on which $X_{CO_2}$ observations can inform surface fluxes.

Differences between inversions in regional net annual fluxes have been well documented, as annual net fluxes have been the primary focus of the majority of CO$_2$ flux inversion studies. To a lesser extent, regional-scale differences in the seasonal cycle of NEE between inversions have also been documented in the literature, particularly when comparing inversions which assimilate in-situ versus GOSAT observations [Chevallier et al., 2014; Ishizawa et al., 2016]. As a demonstration of these differences, we compare the NEE fluxes from CT2016 and GOSAT-Inv at 2° × 2.5° resolution. Figure 11 shows the maximum rate of NEE drawdown during the growing season for CT2016 and GOSAT-Inv for each grid cell. There are substantial differences between the inversions which have structure on the scales of the biomes examined in this study. Thus, it is unlikely that reliable NEE seasonal cycle estimates are possible on these scales, however, more research is needed to quantify the scales that can be constrained.

©2018 American Geophysical Union. All Rights Reserved.
As demonstrated in Sect 3.5, the magnitude of GPP has a large influence on the \( \text{opt} R_{\text{inv-mod}} \) seasonal cycle. Thus, the fact that there is no consensus on the global total GPP \([\text{Anav et al.}, 2015]\) remains a major limitation on inferring \( R_e \). Furthermore, we do not have a forward model that can relate GPP to SIF observed from space. Previous studies, which assimilated SIF to optimize GPP, have relied on independent GPP estimates \([\text{Parazoo et al.}, 2014; \text{Bowman et al.}, 2017; \text{Liu et al.}, 2017]\). \text{Parazoo et al.} [2014] prescribed the annual magnitude of GPP (using MPI-BGC), but optimized the temporal-spatial structure redistributed by the assimilation of SIF. Therefore, the development of a forward model relating GPP to SIF would greatly improve our ability to isolate GPP and \( R_e \) fluxes. These estimates could then be directly compared to FLUXCOM and TBM GPP and \( R_e \).

5. Conclusions

In the first part of this study (Sec. 3.1–3.4), GOME-2 SIF and TCCON \( X_{\text{CO}_2} \) data were employed to evaluate carbon fluxes produced by three FLUXCOM products (ANN, MARS, and RF) and four TBMs (CTEM, JULES, CASA, and SiB3). In general, the normalized seasonal cycle of GPP for the TBMs with diagnostic phenology (SiB3 and CASA) and FLUXCOM were in close agreement with SIF (with RMS differences less than 0.0045). TBMs with prognostic phenology (CTEM and JULES) showed comparatively worse agreement with SIF (with RMS differences greater than 0.006). The closer agreement between the seasonality of GPP and SIF for the diagnostic TBMs relative to prognostic TBMs suggests that the assimilation of vegetation indices improves GPP fluxes. However, we did not control for the driving meteorology, which could be partially responsible for the differences. Comparisons of simulated \( X_{\text{CO}_2} \) with TCCON showed close
agreement for the diagnostic TBMs (with RMS differences less than 0.6 ppm) and worse agreement for the prognostic TBMs (with RMS differences greater than 0.7 ppm). Differences in the driving meteorology for CTEM resulted in larges differences in simulated fluxes and, consequently, in the agreement with observations.

In the second part of this study (Sec. 3.5), a simple method for estimating $R_e$ fluxes using constraints on GPP and NEE was described and tested. In this method, $optR_{inv-mod}$ is calculated by making idealized assumptions about NEE and GPP. Strong agreement between the seasonality of the normalized seasonal cycle of SiB3, CASA, and FLUXCOM GPP with SIF suggested that differences between model-NEE-based $X_{CO_2}$ and TCCON $X_{CO_2}$ seasonality were driven by differences in $R_e$. Thus, we assumed that GPP from SiB3, CASA, and FLUXCOM were correct. To generate constraints on NEE, we used optimized NEE from two flux inversions, which produced a posteriori CO$_2$ fields that were in close agreement with TCCON data. Assuming that GPP and NEE were known, $optR_{inv-mod}$ was calculated as the difference between the optimized NEE and the TBM and FLUXCOM GPP (Eq. (5)).

Using this approach, we calculated $optR_{inv-mod}$ for all possible combinations of GPP (ANN, MARS, RF, SiB3, and CASA) and NEE (CT2016, GOSAT-Inv). This ensemble of GPP and NEE was found to produce $optR_{inv-mod}$ with reasonable precision. The largest differences in the seasonality of $optR_{inv-mod}$ curves were due to the magnitude of GPP, which is variable among models but not well constrained by observations [Anav et al., 2015]. $optR_{inv-mod}$ exhibited a broader summer peak than $R_e$ modeled by SiB3, CASA, and FLUXCOM. The seasonality of $optR_{inv-mod}$ suggested reduced $R_e$ in the summer but enhanced $R_e$ in the spring and fall. Differences were systematic from FLUXCOM.
and the TBMs during June–July, when $optR_{\text{inv-mod}}$ was reduced, and during October, when $optR_{\text{inv-mod}}$ was enhanced. Reduced $R_e$ during the early summer is consistent with the results of Wehr et al. [2016], and could be explained by the Kok effect (inhibition of leaf respiration by light). Enhanced fall $R_e$ is consistent with Commane et al. [2017], who found significant fall $R_e$ in Alaska, and suggests that fall $R_e$ may be greater than previously appreciated.

The seasonality of $optR_{\text{inv-mod}}$ has significant implications on $R_e$ calculations in TBMs. We demonstrated that carbon pool dependence for $R_H$ is important for recovering $R_e$ consistent with $optR_{\text{inv-mod}}$, however, it was also shown that this carbon pool dependence could not solely explain systematic differences in $R_e$ (Sec. 3.5.3). Instead, the results suggest that using a constant carbon use efficiency throughout the year introduces biases into $R_e$ fluxes. Overall, the results suggest that the inclusion of variable carbon use efficiency for $R_a$ and carbon pool dependence for $R_H$ are important for accurately simulating $R_e$.

Our results highlight the utility of the SIF data for informing CO$_2$ flux inversions. The significant differences found between bottom-up and top-down estimates of $R_e$ motivate further development of inversion methods to assimilate both atmospheric CO$_2$ and SIF observations. Based on this analysis, both CASA and SiB3 produce realistic prior GPP and NEE fluxes and can therefore provide useful prior fluxes for future analysis. A current limitation is that only large-scale $optR_{\text{inv-mod}}$ was investigated due to the fact that the accuracy of the seasonal cycle of NEE from flux inversions on smaller scales is uncertain. The scales over which the mean seasonal cycle of NEE is consistent between inversions is not well documented in the literature, but needs to be further investigated to provide GPP and $R_e$ estimates on smaller scales.
Appendix A: Model Description


The version of the model used here, CASA-GFED3, was modified from Potter et al. [1993] as described in Randerson et al. [1996] and van der Werf et al. [2006]. It is driven by MERRA reanalysis and satellite-observed Normalized Difference Vegetation Index (NDVI) to track plant phenology. We use the same fluxes as are used for the CarbonTracker 2016 (http://carbontracker.noaa.gov) prior. CASA outputs monthly fluxes of Net Primary Productivity (NPP) and heterotrophic respiration ($R_H$). From these fluxes, GPP and $R_e$ are estimated to be GPP = 2NPP and $R_e = R_H - NPP$. Temporal downscaling and smoothing was performed from monthly CASA fluxes to 90 minute fluxes using temperature and shortwave radiation from the ECMWF ERA-interim reanalysis (note this method differs from Olsen and Randerson [2004]). GFED_CMS is used for global fire emissions (http://nacp-files.nacarbon.org/nacp-kawa-01/). We use average model fluxes by averaging the fluxes for 2007-2012.

A2. Simple Biosphere Model (SiB3)

SiB3 was originally designed as a lower boundary for General Circulation Models with explicit treatment of biophysical processes. The ability to ingest satellite phenology was introduced [Sellers et al., 1996a, b], and further refinements included a prognostic canopy air space [Vidale and Stöckli, 2005], more realistic soil and snow [Baker et al., 2003] and modifications to calculations of root water uptake and soil water stress [Baker et al., 2008]. The current version is called SiB3. Simulations used in this analysis use phenology (Leaf Area Index, LAI; fraction of Photosynthetically Active Radiation, fPAR) from MODIS. Modern Era Retrospective-analysis for Research and Applications (MERRA) reanalysis

©2018 American Geophysical Union. All Rights Reserved.
is used as model inputs, with precipitation scaled to Global Precipitation Climatology Project [GCP: Adler et al., 2003] following Baker et al. [2010].

A3. Canadian Terrestrial Ecosystem Model (CTEM)

CTEM is a dynamic vegetation model developed for inclusion in the Canadian Center for Climate Modeling and Analysis (CCCma) coupled general circulation model. Because CTEM is designed to model ecosystems under climate change, the phenology parametrization has to be independent of current climatic factors. Thus, a “carbon-gain” approach is used to determine phenology, which is based on local environmental conditions. In this approach, leaf onset is initiated when it is beneficial for the plant, in carbon terms, to produce new leaves. Leaf offset is initiated by unfavorable environmental conditions that incur carbon losses and these include shorter day length, cooler temperatures, and dry soil moisture [Melton and Arora, 2016; Arora and Boer, 2005]. We use two sets of CTEM fluxes which are driven by different meteorology. One set is generated using CRU NCEP (merged product of NCEP reanalysis and CRU observations) which we refer to as CTEM-CRU. Another set is generated using the Global Environmental Multi-scale - Modeling Air Quality and CHemistry - Greenhouse Gas (GEM-MACH-GHG) operational weather prediction model [Polavarapu et al., 2016] (run is refer to as CTEM-GEM). We use average model fluxes by averaging 2009-2010 fluxes, the only two years available.

A4. Joint UK Land Environment Simulator (JULES)

JULES is a community land surface model that has evolved from the Met Office Surface Exchange Scheme. Phenology in JULES affects leaf growth rates and timing of leaf growth/senescence based on temperature alone [Clark et al., 2011]. Vegetation cover

©2018 American Geophysical Union. All Rights Reserved.
is predicted based on nine plant functional types that compete for space based on their relative productivity and height but are excluded from growing on agricultural land, based on a fraction of agriculture in each grid cell [Harper et al., 2018]. CRU-NCEP was used as model forcing data.

Acknowledgments.

Funding for this work was provided by the Canadian Space Agency, NSERC, and Environment and Climate Change Canada. I.B.’s contribution was sponsored by the National Science Foundation Science and Technology Center for Multi-Scale Modeling of Atmospheric Processes, managed by Colorado State University under cooperative agreement No. ATM-04252467. CarbonTracker CT2016 results were provided by NOAA ESRL, Boulder, Colorado, USA from the website at http://carbontracker.noaa.gov. TCCON data were obtained from the TCCON Data Archive, hosted by CaltechDATA [http://tccondata.org]. NASA and GFZ Potsdam GOME-2 SIF products were obtained from Aura Validation Data Center [http://avdc.gsfc.nasa.gov] and GFZ-Potsdam FTP [ftp://ftp.gfz-potsdam.de], respectively. FLUXCOM products were obtained from the Data Portal of the Max Planck Institute for Biochemistry [https://www.bgc-jena.mpg.de]. MERRA-2 products were downloaded from MDISC [https://disc.sci.gsfc.nasa.gov], managed by the NASA Goddard Earth Sciences (GES) Data and Information Services Center (DISC). The GEOS-Chem forward and adjoint models are freely available to the public. Instructions for downloading and running the models can be found at http://wiki.geos-chem.org/. ACOS GOSAT lite files were obtained from the CO2 Virtual Science Data Environment [https://co2.jpl.nasa.gov/#mission=ACOS]. We thank Kenneth Schuldt for providing TM5 \(X_{\text{CO}_2}\) at TCCON sites. We thank Andy Jacobson, Saroja Polavarapu, Paul

©2018 American Geophysical Union. All Rights Reserved.
Wennberg, Randy Kawa, Jim Collatz and two anonymous referees for helpful comments on this manuscript.

References


©2018 American Geophysical Union. All Rights Reserved.


©2018 American Geophysical Union. All Rights Reserved.


©2018 American Geophysical Union. All Rights Reserved.


Friedl, M. A., D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, and X. Huang (2010), MODIS collection 5 global land cover: Algorithm refinements

©2018 American Geophysical Union. All Rights Reserved.


©2018 American Geophysical Union. All Rights Reserved.


©2018 American Geophysical Union. All Rights Reserved.


©2018 American Geophysical Union. All Rights Reserved.


van der Werf, G. R., J. T. Randerson, L. Giglio, G. J. Collatz, P. S. Kasibhatla, and A. F. Arellano Jr (2006), Interannual variability in global biomass burning emissions from


©2018 American Geophysical Union. All Rights Reserved.


©2018 American Geophysical Union. All Rights Reserved.
<table>
<thead>
<tr>
<th>Model</th>
<th>Meteorology</th>
<th>Phenology</th>
<th>Respiration variables</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>CASA</td>
<td>MERRA&lt;sup&gt;a&lt;/sup&gt;</td>
<td>NDVI</td>
<td>carbon pool size (C), temperature (T) and soil moisture (M)</td>
<td>2007-2012</td>
</tr>
<tr>
<td>SiB3</td>
<td>MERRA&lt;sup&gt;b&lt;/sup&gt;</td>
<td>MODIS FPAR and LAI&lt;sup&gt;c&lt;/sup&gt;</td>
<td>T and M</td>
<td>2007-2012</td>
</tr>
<tr>
<td>CTEM-CRU</td>
<td>NCEP-CRU&lt;sup&gt;c&lt;/sup&gt;</td>
<td>carbon-gain&lt;sup&gt;f&lt;/sup&gt;</td>
<td>C, T and M</td>
<td>2009-2010</td>
</tr>
<tr>
<td>CTEM-GEM</td>
<td>GEM-MACH-GHG&lt;sup&gt;d&lt;/sup&gt;</td>
<td>carbon-gain</td>
<td>C, T and M</td>
<td>2009-2010</td>
</tr>
</tbody>
</table>

<sup>a</sup>Modern-Era Retrospective Analysis for Research and Applications [Rienecker et al., 2011]

<sup>b</sup>Precipitation scaled to Global Precipitation Climatology Project [GCP: Adler et al., 2003] following Baker et al. [2010]

<sup>c</sup>Wei et al. [2014]

<sup>d</sup>Anselmo et al. [2010]; Robichaud and Ménard [2014]; Makar et al. [2015]

<sup>e</sup>Stöckli et al. [2008]

<sup>f</sup>see Appendix A
Figure 1. MODIS IGBP vegetation classification (at a horizontal resolution of $2^\circ \times 2.5^\circ$). Coloring indicates that the given vegetation type makes up more than 50% of the vegetation type in the gridcell. The vegetation regions are: evergreen needle leaf forests (ENF), deciduous needleleaf forests (DNF), southern mixed forests, northern mixed forests, grasslands and croplands. Red circles show the locations of the four TCCON sites examined in this study: Park Falls (45.9°N, 90.3°W), Orléans (48.0°N, 2.1°E), Bialystok (53.2°N, 23.0°E), and Sodankylä (67.4°N, 26.6°E).
Figure 2. Mean seasonal cycles of (a) GPP, (b) $R_e$, and (c) NEE simulated by SiB3 (green), CASA (blue), CTEM-CRU (dashed salmon), CTEM-GEM (dashed orange), JULES (dashed purple), ANN (dotted gray), MARS (dotted purple-gray), and RF (dotted cyan) between 39°N and 65°N. CT2016 (solid red) and GOSAT-Inv NEE (dash-dotted red) are also plotted.

©2018 American Geophysical Union. All Rights Reserved.
Figure 3. Normalized seasonal cycles of GOME-2 SIF (NASA and GFZ), and model GPP (SiB3, CASA, CTEM-CRU, CTEM-GEM, JULES and FLUXCOM ANN, FLUXCOM MARS, and FLUXCOM RF) for six vegetation regions. For each panel, the upper plot shows the seasonal cycle of GPP and SIF scaled so that the integral over the season equals one. The lower plot shows the difference between scaled TBM or FLUXCOM GPP and scaled NASA GOME-2 SIF. Grey shaded regions show the uncertainty estimate of the SIF seasonal cycle.

©2018 American Geophysical Union. All Rights Reserved.
Figure 4. Five-year mean $X_{\text{CO}_2}$ seasonal cycle at Sodankylä (top row), Białystok (second row), Orléans (third row) and Park Falls (bottom row). TCCON $X_{\text{CO}_2}$ is shown in black. The columns show, from left to right: GOSAT-Inv, CT2016, CASA, SiB3, CTEM-CRU, CTEM-GEM, and JULES. Shaded regions indicate the uncertainty in the functional fit.
Figure 5. The RMS difference between normalized GPP and NASA GOME-2 SIF (averaged across all vegetation regions) versus the RMS difference between simulated X\(\text{CO}_2\) and TCCON (averaged across the four TCCON sites).
Figure 6. GPP·(−1), $R_e$, and NEE fluxes over 39°–65° N. (a) CASA GPP·(−1), $R_e$ and scaled $f(T)$ (Eq. 4). (b) SiB3 GPP·(−1), $R_e$, and scaled $f(T)$ (Eq. 4). (c) SiB3 and CASA NEE.
Figure 7. Left column shows (a) model GPP·(-1), (c) model $R_e$, (e) $optR_{\text{inv-mod}}$, and (g) the difference between optimized and model $R_e$ for SiB3, CASA, and FLUXCOM (ANN, MARS, and RF). $optR_{\text{CT2016-mod}}$ (based on CT2016 NEE) is represented by solid lines, whereas $optR_{\text{GOSAT-Inv-mod}}$ (based on GOSAT-Inv) is indicated by dashed lines. Right column shows the normalized seasonal cycles of (b) model GPP, (d) model $R_e$, (f) $optR_{\text{inv-mod}}$, and (h) the difference between optimized and model $R_e$. In all panels the solid, heavy black line represents the mean of all the curves shown.

©2018 American Geophysical Union. All Rights Reserved.
Figure 8. $optR_{inv-mod}$ (solid lines), after scaling GPP over a range of values, for SiB3, CASA and FLUXCOM. $optR_{inv-mod}$ are calculated after first scaling GPP to (a) 23 Pg yr$^{-1}$, (b) 26 Pg yr$^{-1}$, (c) 29 Pg yr$^{-1}$, and (d) 32 Pg yr$^{-1}$. $R_e$ produced by SiB3, CASA, and FLUXCOM without scaling is indicated by the dotted lines. Colors are as in Figure 7.
Figure 9.  (a, b) optR_{inv-mod}, (c, d) optimized $R_H$, and (e, f) optimized $R_H/f(T)$ using parameterizations from Randerson et al. [1996] (left column) and from [Denning et al., 1996] (right column). The solid black line shows the mean seasonal cycle and shaded gray region shows the range of optimized $R_e$ for all NEE and GPP. The dotted black line shows $f(T)$ for CASA (a) and for SiB3 (b) scaled to fit plot area.
Figure 10. Normalized seasonal cycles of (a–c) model GPP, (d–f) model $R_e$, (g–i) $optR_{inv-mod}$, and (j–l) the difference between $optR_{inv-mod}$ and model $R_e$ for (left column) North America, (middle column) Europe, and (right column) Asia. For subplots g–l, $optR_{CT2016-mod}$ is represented by solid lines, whereas $optR_{GOSATinv-mod}$ is indicated by dashed lines. In all panels the solid, heavy black line represent the mean of all the curves shown.
Figure 11. Maximum rate of drawdown (gC m$^{-2}$ day$^{-1}$) during the growing season for (a) CT2016, (b) GOSAT-Inv, and (c) GOSAT-Inv minus CT2016.