MODELING SOLAR-INDUCED CHLOROPHYLL FLUORESCENCE (SIF) IN FORESTS: IMPACTS OF MODEL STRUCTURE, PARAMETER UNCERTAINTIES, AND LANDSCAPE HETEROGENEITY USING NEON AS A TESTBED

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Zhenqi Luo

May 2024

© 2024 Zhenqi Luo

ABSTRACT

Accurately quantifying gross primary production (GPP) is crucial for monitoring carbon cycles, assessing the global carbon budget, and predicting vegetation stress in a changing climate. Solar-induced chlorophyll fluorescence (SIF), an optical signal emitted by chlorophyll a during photosynthesis under solar illumination, serves as a remote sensing proxy for mechanistically understanding photosynthesis from space. Realistic model parameterization of SIF is critical for constraining GPP estimations. However, accurate SIF parameterizations and simulations depend on model structure (mathematical formulation of processes) and parameters (coefficients in these mathematical formulations). The degree to which model structure and parameters impact SIF simulations has never been evaluated with independent measurements. In this thesis, I utilized the Soil Canopy Observation Photosynthesis Energy (SCOPE) radiative transfer model to simulate SIF and GPP at three forest sites within the National Ecological Observatory Network (NEON). First, I evaluated the impact of model structure on SIF simulations, through two sets of simulations with different SIF parameterizations: 1) based on an empirical non-photochemical quenching (NPQ) parameterization, and 2) utilizing a mechanistic light reaction (MLR) model by modeling the fraction of open Photosystem II (PSII) reaction centers (qL). Second, to examine the impact of model parameters on SIF (and GPP) simulations, I compared simulations using parameters that are 1) plant functional type (PFT)-specific (denoted as *ctrl*), and 2) foliar trait-specific measurements (denoted as *para*). Further, I investigated the impact of subgrid spatial heterogeneity of coarse-resolution SIF

observations and the resulting scale mismatch with model simulations. To achieve this I compared simulations conducted at tower footprint with a single land cover (denoted as baseline) vs land cover (LC)-specific simulations within a 0.05° satellite grid. My analyses revealed that qL-based SIF parameterization outperformed NPQ-based parameterization in capturing the 16-day SIF time series from Orbiting Carbon Observatory-2 (OCO-2) and TROPOspheric Monitoring Instrument (TROPOMI) satellite observations. In addition, foliar trait-specific parameters significantly improved both the simulated GPP and SIF (when using qL-based scheme). In particular, the choice of SIF model structure impacted the sensitivity of simulated SIF to fundamental photosynthetic parameters; for example, the qL-based SIF model has a higher sensitivity to variations in the maximum carboxylation rate (Vcmax). Finally, accounting for sub-grid spatial heterogeneity, i.e., explicitly prescribing LAI and parameters (foliar trait-based) separately for each end member land cover, significantly improved agreement between the model simulations and satellite observations. In summary, this thesis highlighted the importance of physiologically meaningful combinations of model structure and foliar trait-specific parameters, as well as the consideration of sub-grid spatial heterogeneity in enhancing the realism and accuracy of SIF simulations. These insights are vital for advancing our knowledge in estimating, understanding, and predicting SIF and GPP, and ultimately the global carbon budget in response to climate change.

BIOGRAPHICAL SKETCH

Zhenqi Luo, a third-year MS-PhD student at the Soil and Crop Sciences Section, School of Integrative Plant Science, Cornell University, is engaged in groundbreaking research under the advice of Prof. Ying Sun. Focused on the intersection of eddy-covariance (EC) and remote sensing data, Luo's work delves into the modeling of terrestrial gross primary productivity (GPP) and solar-induced chlorophyll fluorescence (SIF).

Luo earned his Bachelor of Science in Resources and Environmental Science with a minor in International Economy and Trading from Beijing Normal University (BNU) in China. During his undergraduate journey, Luo set the stage for future academic pursuits. In 2019 summer, Luo expanded his horizons as a non-degree Visiting Student in Geophysical Science at the University of Chicago. Under the guidance of Dr. Elisabeth Moyer, he contributed to research on the distribution of convective available potential energy (CAPE), gaining exposure to diverse perspectives in the geophysical field. Due to Covid19 pandemic, Luo deferred his graduate studies to the US and embarked on an Internship Research opportunity at Westlake University in China. As a Visiting Student in the Atmospheric Environmental Research Lab, under the guidance of Dr. Yuzhong Zhang, he explored the global estimation of ammonia emission and its environmental dynamics.

Currently pursuing a PhD at Cornell University, Luo continues to demonstrate a keen interest in advancing biogeochemistry modeling. His dedication to research is underscored by testing three NEON sites, reflecting a strong commitment to knowledge in evaluating, understanding, and predicting SIF and GPP. As he

progresses through his Master's studies towards doctoral studies, his contributions to the understanding of GPP and SIF will hold great promise for the scientific community and beyond.

ACKNOWLEDGMENTS

“Stay hungry, stay foolish.” Three years at Cornell University, the journey has been marked by both sorrows and joys. Although it will become the past, for me, it is far from fleeting; the ups and downs experienced are deeply imprinted in my heart.

To maintain confidence in joyful work with a serious and free-spirited attitude; to gain the courage for a happy life through a spirit of self-improvement and struggle; to fulfill responsibilities to a healthy society with a broad-minded and realistic approach. This is undoubtedly the purest and most valuable wealth bestowed upon me by Professor Ying Sun. From undergraduate studies to graduate studies, the most profound experience is that a sense of belonging differs due to different learning methods — from school to mentorship. My deepest emotional connection to Cornell University belongs to SunLab. In 2021, I enrolled in the Soil and Crop Sciences Section at Cornell and became a student of Professor Sun. I believe it is fate that brought our life paths together. Since then, Professor Sun has been a mentor and friend, trusting me, offering assistance and freedom, and guiding me in the ways of life and learning. Grateful for each teaching and guidance from Professor Sun, satisfied with every understanding and tolerance.

During the past three years, whether in learning or life, I am very grateful to my colleagues at SunLab. Special thanks to Dr. Jiaming Wen for providing constructive opinions during the scientific research process. Thanks to PhD candidate Jiameng Lai for inspiration and companionship in life, as well as guidance and assistance in academics. Thanks to SunLab for enlightening and nurturing me. I am grateful to the teachers, classmates, and friends who brought color to my life: teachers in college, classmates in the class, friends in the soccer club, and more. You have shown me more possibilities in life during my three years of graduate study. Thank you for your presence, and I wish you good health and success in your work! After graduating with

a Master's, I will continue studying in the SunLab to pursue a PhD degree. I believe in the future, our spiritual qualities will continue to walk together. Lastly, I want to express my deepest gratitude to my parents. They patiently accompanied me and tirelessly encouraged me during the confusing moments of my life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The accurate modeling of gross primary production (GPP) is crucial in monitoring carbon cycles, quantifying the global carbon budget, and detecting vegetation stress under changing climate conditions (Bonan et al., 2011; Liu et al., 2017; Sun et al., 2023b). However, measuring photosynthesis beyond the leaf scale is challenging (Porcar-Castell et al., 2014). Solar-induced chlorophyll fluorescence (SIF), from the passive chlorophyll *a* fluorescence (Chl*a*F) emission, has emerged as a mechanistic proxy for remotely sensing vegetation photosynthesis from space since 2007 (Guanter et al., 2007; Mohammed et al., 2019; Sun et al., 2023a). Accurate SIF modeling, facilitated by widely employed leaf/canopy-level radiative transfer models (RTMs), is essential for understanding the potential stressors in carbon cycles and scaling the local measurements to large scales (Thum et al., 2017; Ryu et al., 2019; He et al., 2020; Zhang et al., 2023). Nonetheless, accurately modeling SIF and GPP is also challenging, due to the model structural error and parameter uncertainties (Bonan et al., 2011; Bonan and Doney, 2018).

Several modeling studies have primarily employed the leaf-level formulation of the soil canopy observation photosynthesis energy (SCOPE) model, developed by van der Tol et al., (2009). This model combines radiative transfer and non-radiative energy fluxes into a single representation of the soil and vegetation system. To derive the top-of-canopy (TOC) SIF radiance and reflectance, SCOPE utilizes the leaf-level RTM Fluspect (Vilfan et al., 2016) and the canopy-level RTM FluoSAIL model (van der Tol et al., 2009). The SCOPE model comprises submodules representing leaf photosynthetic processes, incorporating the Farquhar-von Caemmerer-Berry (FvCB)

photosynthesis model and Ball-Berry stomatal conductance model to estimate leaf-level photosynthesis and photochemical quantum yield (ΦP; Farquhar et al., 1980; van der Tol et al., 2014). This photosynthesis estimation is further used to calculate fluorescence yield (ΦF) and consequently, SIF through empirical modeling of non-photochemical quenching (NPQ, denoted as k_N), as an exclusive function of ΦP, based on the relative light saturation of photosynthesis (van der Tol et al., 2014). The initial calibration of the k_N equation utilized experimental data from cotton plants exposed to varying light intensities, CO_2 concentrations, and temperatures, and its performance was evaluated using species of C_3 crops (e.g., tobacco) and temperate woody plants (e.g., grape, deciduous sclerophyll shrub) grown in greenhouses. Subsequently, Lee et al., 2013 proposed a similar formula for k_N, using empirical data from ten Mediterranean species, naturally grown outdoors (Galmés et al., 2007). These ΦP-parameterized empirical formulations of NPQ have been explicitly incorporated into Terrestrial biosphere models (TBMs)/land surface models (LSMs), such as the Community Land Model version 4.5 (CLM4.5; Raczka et al., 2019), CLM5.0 (Li Rong et al., 2022), and have been served as a benchmark for evaluating small-scale complex 3D models like the Fluorescence model with Weighted Photon Spread method (FluorWPS; Zhao et al., 2016), and the Fluorescence Radiative Transfer model based on Escape and Recollision probability (FluorRTER; Zeng et al., 2020). Furthermore, this empirical parameterization has been kept intact in the new generation of the SCOPE model, i.e., SCOPE 2.0 (Yang et al. 2021a).

The model of van der Tol et al. (2014) is useful in assisting in interpreting relationships between SIF and GPP, but concerns have been raised about the

uncertainties of model structure and parameters. The complexity involved in modeling NPQ arises from the inconsistent and delayed release of its multiple components at different timescales, and its activation and regulation that occur across different cell compartments (i.e., lumen, thylakoid membrane, and stroma; Han et al., 2022a). Several studies have reported disagreements in modeled SIF with ground and space observations. For instance, Yang et al. (2021b) found that the ΦP-parameterized NPQ approach cannot reproduce nonlinear relationships among ΦP and ΦF from in-situ leaf-level measurements, in which ΦP is much more responsive to photosynthetically active radiation (PAR) than ΦF (Gu et al., 2019). In addition, the SCOPE model may struggle to capture the variability in observed SIF signals under stress conditions (e.g., water limitation; Cannière et al., 2021). Moreover, Parazoo et al. (2020) revealed systematic biases in leaf-to-canopy SIF modeling across wide TBMs. Thus, a mechanistic-based and globally scalable model structure should be developed to improve model performance, hence promoting meaningful integration with satellite SIF observations to constrain TBMs.

The mechanistic light reaction (MLR)-SIF equations, proposed by Gu et al. (2019), establish a theoretical SIF modeling, from modeling the fraction of open Photosystem II (PSII) reaction centers (q_L). Specifically, the FvCB photosynthesis model can be utilized to calculate the actual electron transport rate (J_a), which in turn can be used to determine SIF by considering the q_L which indicates the redox state of PSII (Kramer et al., 2004). This qL-based SIF modeling approach is based on mechanistic principles and has simple intrinsic physical properties, avoiding the complexity of the ΦP-parameterized NPQ approach. The MLR-SIF model has successfully reproduced

observed photosynthesis at both leaf and canopy scales (Han et al., 2022a; Liu et al., 2022; Chen et al., 2023), and under varying soil–water conditions (Shi et al., 2022). However, the applicability of the MLR qL-based SIF modeling approach for estimating SIF using modeled J_a and q_L remains unexplored, and its accuracy and scalability at the canopy scale have not been validated across multiple sites. We term this as SIF model structure uncertainties.

In terms of model parameters, SCOPE simulates leaf physiological processes in a biochemical module as functions of leaf temperature, ambient radiation levels, intercellular CO_2 concentrations (C_i), and other leaf physiological parameters (Collatz et al., 1991). With SCOPE simulations, GPP is shown to be sensitive to plant functional type (PFT)-specific photosynthetic parameters, such as maximum carboxylation rate (Vcmax) and Ball-Berry stomatal conductance slope (Koffi et al., 2015; Pacheco-Labrador et al., 2019; Prikaziuk et al., 2023). In contrast, SIF calculated with the ΦP-parameterized NPQ-based parameterization shows insensitivity to Vcmax values under moderate light conditions and leaf area index (LAI) values (Koffi et al., 2015; Camino et al., 2019), and is more affected by chlorophyll content and canopy structure (e.g., leaf inclination and LAI) (Verrelst et al., 2015; Yang et al., 2019). Furthermore, Han et al. (2022b) demonstrated that the product of the redox state with SIF (q_L ×SIF), rather than SIF alone, is a more powerful predictor of Vcmax sensitivity. We term this as SIF and GPP model parameter uncertainties.

The vast coverage area of satellite-based sensors allows them to collect SIF signals across a variety of PFTs, but the 1-D SIF simulation, which is treated as a

homogeneous stand and usually conducted at coarse spatial resolutions, even with the appropriate model parameters and model structure, may not be comparable with the SIF observations due to the diverse biophysical and biochemical properties of vegetation. Additionally, the spatial resolution of existing native spaceborne SIF products is relatively low (Sun et al., 2023b). Scaling the relationship between SIF and GPP is highly influenced by landscape heterogeneity (i.e., PFTs and environmental conditions), according to empirical analyses. For example, Wang et al. (2021) demonstrated that environmental conditions such as temperature, radiation, and precipitation regulate the spatial distribution of light energy in SIF and GPP. Moreover, the relationship between spatially heterogeneous SIF and GPP can be influenced by different PFTs (Song et al., 2021). However, a mechanistic and quantitative analysis of the degree to which spatial heterogeneity can impact SIF simulations, and how SIF simulations deviate from satellite observations remain unexplored. Thus, a more robust method of SIF modeling should consider diverse landscapes to mitigate the impact of spatial heterogeneity.

Therefore, this study aims to examine the impact of model structure, parameter uncertainties, and landscape heterogeneity on SIF modeling, by conducting simulations in three National Ecological Observatory Network (NEON) forest sites as a testbed, with three main objectives. The first objective is to validate the effectiveness of the mechanistic qL-based model structure in accurately simulating SIF, which relies on the first-order corrections of GPP modeling. Following the validation of the model structure, the second objective is to investigate the impact of key photosynthetic parameters on the simulations of SIF and GPP. Lastly, the third objective is to

explicitly examine the impact of sub-grid heterogeneity on SIF simulations, which will offer critical insights for addressing scale mismatch between satellite and in-situ SIF measurements. These simulations were conducted with the SCOPE radiative transfer framework and were compared against three satellite products (for SIF) and eddy covariance (EC) tower-derived fluxes (for GPP).

2. MATERIALS AND METHOS

We employed the SCOPE framework to simulate forest SIF and GPP fluxes, using three representative sites within the NEON network. Within the SCOPE modeling framework, we implemented a process-based parameterization to calculate q_L and subsequently SIF. This qL-based SIF parameterization was compared with the ΦP-parameterized NPQ-based approach that is widely used in the literature (van der Tol et al., 2014). To drive model simulations, we used a suite of data streams including gap-filled meteorological and EC observations from the National Center for Atmospheric Research (NCAR)-NEON system (Lombardozzi et al. 2023), LAI from Copernicus Global Land Service (CGLS), key photosynthetic parameters from NEON measurements and the global database of plant traits (TRY) that are specific to land cover (LC) types that come from USGS National Land Cover Database (NLCD). To validate our model simulations, we utilized SIF retrievals from TROPOspheric Monitoring Instrument (TROPOMI) and Orbiting Carbon Observatory-2 (OCO-2). Details of our model parameterization, simulation setup, and observational products for model input and validation are provided below.

2.1. Study sites

We chose three forest sites in the NEON network, which operates 47 terrestrial sites across the US and Puerto Rico and covers 20 ecoclimatic domains (https://www.neonscience.org/). These three forest sites have distinct combinations of LC types and are located in different climate regimes, namely Bartlett Experimental Forest (BART), Steigerwaldt-Chequamegon (STEI), and the Jones Center At Ichauway (JERC) (**Table 1**). The dominant LC type within a 0.05° × 0.05° (the pixel

size of satellite SIF utilized in this study, **Sect 2.3.6**) domains is forest (BART: 85 %, STEI: 58 %, and JERC: 50 %; **Table A1**), with some urban area in BART (8 %), wetlands in STEI (22 %), and cultivated crops in JERC (34 %). The LC types within the tower footprint (defined as 90% flux area of the primary and secondary airsheds of each tower) in BART, STEI, and JERC are deciduous forest and mixed forest, deciduous forest, and evergreen forest, respectively (**Fig. 1**).

Table 1 Characteristics of the three selected NEON sites. The footprint LC represents the dominant LC at the flux tower from **Fig. 1**.

Name	Tower footprint[a] LC	Mean canopy height (m)	Mean Elevation (m)	Latitude	Longitude	Number of sampling plots
BART	Deciduous Forest & Mixed Forest	23	274	44.064	-71.287	16
STEI	Deciduous Forest	20	476	45.509	-89.586	18
JERC	Evergreen Forest	27	47	31.195	-84.469	22

[a]90% flux area of the primary and secondary airsheds of each tower

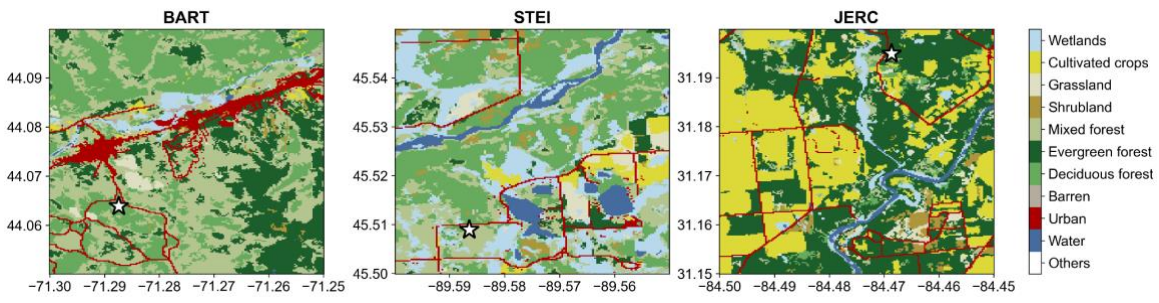

Fig. 1. The land cover (LC) types of our study domains. Each domain corresponds to a single SIF pixel at 0.05° × 0.05° resolution. The locations of the flux towers (BART, JERC, and STEI) are marked with white stars. The LC dataset is from the 2019 National Land Cover Database (NLCD) (Dewitz, J., and U.S. Geological Survey, 2021).

We prioritized these three sites as they had higher quality data during early testing of the NCAR-NEON system, including fewer data gaps during 2018-2021 than other sites. Specifically, for each site over the entire study period (48 months), a minimum of 35 months have at least 50 % valid net ecosystem exchange (NEE) measurements (**Fig. A1**). In addition, these three sites have been documented to have greater data quality and closer agreement with the Community Terrestrial System Model (CTSM) simulations (Lombardozzi et al. 2023).

2.2. Model simulations

2.2.1. SCOPE model

The SCOPE model is an RTM for multi-layer vegetation canopies that is widely used for simulating canopy-level photosynthesis and SIF (van der Tol et al., 2009). In our study, we used the latest version (v2.1; Yang et al., 2021a). At the leaf level, SCOPE employs a unified four-stream radiative transfer scheme among horizontal canopy layers, considering leaf orientations towards the sun and sensors in a stochastic manner (Yang et al., 2021a). Scaling from leaf to canopy level is achieved through leaf-to-canopy integration, incorporating gap fractions and relative contributions of leaf angle classes to account for sunlit and shaded leaves. This model has been used in

studies to estimate GPP and SIF for different PFTs across regional and global scales (e.g., Cui et al., 2020; Li Xiaoyao et al., 2022; Prikaziuk et al., 2023).

2.2.2. The NPQ-based SIF parameterization

The leaf-level parameterization of SIF emission in SCOPE is based on the regulated heat dissipation process and ΦP (NPQ-based in **Fig. 2**). Specifically, k_N (the rate constant of NPQ) is computed first as an empirical nonlinear function of the degree of light saturation (x_l, defined as the ratio of the actual vs potential electron transport rate (ETR) $J_{a,l}/J_{p,l}$, hereafter with subscript l denoting sunlit or shaded leaves in a canopy):

$$k_{N,l} = k_{N,l}^0 \frac{(1+\beta){x_l}^\alpha}{\beta + {x_l}^\alpha} \quad (1)$$

where $k_{N,l}^0$, α and β are coefficients/parameters fitted with a limited number of experimental data (from cotton, and maize, van der Tol et al., 2014). Here x_l is computed as:

$$x_l = 1 - \frac{\phi_{P,l}}{\phi_{P0,l}} \quad (2)$$

with $\phi_{P,l}$ and $\phi_{P0,l}$ denoting the actual and potential photochemistry yields, respectively. $\phi_{P,l}$ can be generated using:

$$\phi_{P,l} = \phi_{P0,l} \frac{J_{a,l}}{J_{p,l}} \quad (3)$$

where $J_{a,l}$ can be calculated from gross photosynthesis (Ag_l), which is derived from FvCB in SCOPE (**Eq. A1-A2 in Text A1**), and $J_{p,l}$ is derived with the absorbed photosynthetically active radiation ($APAR_l$):

$$J_{a,l} = \begin{cases} Ag_l \dfrac{4(C_{i,l} + 2\Gamma^*{}_l)}{C_{i,i} - \Gamma_l^*}, \ for\ C_3\ speci \\ \dfrac{1}{0.17} Ag_l, \ for\ C_4\ species \end{cases} \quad (4)$$

$$J_{p,l} = 0.5 \cdot APAR_l \cdot \phi_{P0,l} \quad (5)$$

$$\phi_{P0,l} = \frac{k_P}{k_P + k_F + k_{D,l}} \quad (6)$$

where $C_{i,l}$ is the intercellular CO_2 concentration, and Γ_l^* is the CO_2 compensation point in the absence of mitochondrial respiration in the light for C_3 plants. k_P and k_F can be assumed to be constants, representing the rate constants of photochemistry and Chl*a*F emission, respectively. $k_{D,l}$ is modeled using a linear regression of leaf temperature (T_l, k) when above 26.5 °C, to represent the processes of unregulated heat dissipation (van der Tol et al., 2014):

$$k_P = 4.0 \quad (7)$$

$$k_F = 0.05 \quad (8)$$

$$k_{D,l} = max(0.8738,\ 0.0301T_l + 0.0773) \quad (9)$$

Once $J_{a,l}$ is derived, the steady-state fluorescence yield ($\phi_{F,l}$) at the leaf level can be obtained by combining the photochemical yield ($\phi_{P,l}$) and the fluorescence yield at saturating irradiance for a light-adapted leaf ($\phi_{Fm,l}$) at the leaf-level:

$$\phi_{F,l} = \phi_{Fm,l}\left(1 - \phi_{P,l}\right) \quad (10)$$

$$\phi_{Fm,l} = \frac{k_F}{k_F + k_{D,l} + k_{N,l}} \quad (11)$$

Ultimately, the leaf-level ChlaF emission (denoted as $F_{e,l}$) for each layer is generated by multiplying $APAR_l$ and $\phi_{F,l}$. Considering that the total canopy is separated into

sunlit and shaded leaves in the SCOPE, the Fe for the sunlit- ($F_{e,sun}$) and shaded- ($F_{e,sha}$) leaf groups were respectively generated as:

$$F_{e,sun} = APAR_{sun} \cdot \phi_{F,sun} \quad (12)$$

$$F_{e,sha} = APAR_{sha} \cdot \phi_{F,sha} \quad (13)$$

where $APAR_{sun}/APAR_{sha}$ and $\phi_{F,sun}/\phi_{F,sha}$ represent the corresponding sunlit and shaded APAR and ϕ_F, respectively.

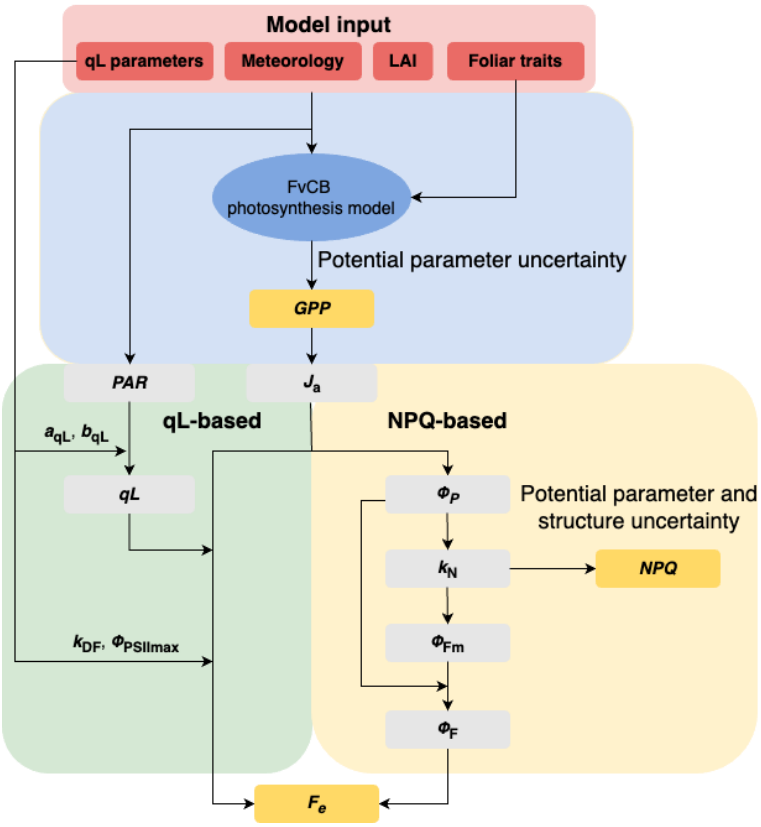

Fig. 2. Diagram of modeling leaf-level GPP and SIF in SCOPE, with NPQ-based (yellow pane; van der Tol et al., 2014) and qL-based (green pane) modeling strategies (this paper), respectively. The symbols are defined in **Table 2**. The qL-based and NPQ-based SIF parameterization share the same meteorological forcing and

vegetation properties (red pane), including LAI and foliar traits. Also, both SIF parameterizations share the same photosynthesis model, i.e., FvCB (blue pane).

Table 2 Notations used in the SCOPE model.

Symbols	Model parameters	Units
Leaf and canopy parameters:		
Cab	Leaf chlorophyll content	$\mu g\ cm^{-2}$
Cca	Leaf carotenoid content	$\mu g\ cm^{-2}$
Cdm	Leaf dry matter content	$\mu g\ cm^{-2}$
Cw	Equivalent water thickness in leaves	cm
$Vcmax_{25}$	Maximum carboxylation capacity (at 25°C)	$\mu mol\ m^{-2}s^{-1}$
LAI	Leaf area index	m^2m^{-2}
Meteorological variables:		
Ta	Air temperature	°C
Rin	Broadband incoming shortwave radiation	$W\ m^{-2}$
Rli	Broadband incoming longwave radiation	$W\ m^{-2}$
P	Air pressure	hPa
ea	Atmospheric vapor pressure	hPa
u	Wind speed	$m\ s^{-1}$
z	Measurement height	m
NPQ-based SIF parameters:		
J_p	Potential electron transport rate	$\mu mol\ m^{-2}s^{-1}$

k_N	Rate constant of NPQ	-
ϕ_P	Actual photochemistry yield	
ϕ_{P0}	Potential photochemistry yield	-
ϕ_F	Steady-state fluorescence yield	-
ϕ_{Fm}	Fluorescence yield at saturating irradiance for a light-adopted leaf	-
qL-based SIF parameters:		
PAR	Photosynthetically active radiation	$\mu mol\ m^{-2}s^{-1}$
a_{qL}, b_{qL}	Empirical parameter for calculating q_L as a function of PAR	-
q_L	The fraction of open PSII reaction centers	-
k_{DF}	The ratio of k_D to k_F	-
$\phi_{PSIImax}$	Maximum quantum efficiency of PSII in dark-adapted leaves	-
Common parameters:		
J_a	Actual electron transport rate	$\mu mol\ m^{-2}s^{-1}$
F_e	leaf-level ChlaF emission	$\mu mol\ m^{-2}s^{-1}$

2.2.3. The $\boldsymbol{q_L}$-based SIF parameterization

The qL-based model is a mechanistic model that simulates redox reactions along the electron transport chain to describe q_L dynamics by PAR (**Fig. 2;** Gu et al., 2019).

This model has been employed to establish the relationship between photosynthesis and SIF at the leaf scale for various PFTs under varying environmental conditions (e.g., Han et al., 2022a; Han et al., 2022b; Shi et al., 2022). Similar to the NPQ-based strategy, it also relies on FvCB to estimate J_a(the blue panel in **Fig. 2**). The difference from NPQ-based modeling is that it computes the leaf-level SIF emission via q_L, the fraction of open PSII reaction centers, which indicates the redox state of PSII acceptors:

$$F_{e,l} = \frac{1 - \phi_{PSIImax}}{\phi_{PSIImax}(1 + k_{DF})} J_{a,l} \frac{1}{q_L} \tag{14}$$

where $\phi_{PSIImax}$ is the maximum photochemical quantum efficiency of PSII in dark-adapted leaves, and k_{DF} is the ratio of k_D and k_F, which is treated as a constant value in this study (Han et al., 2022a):

$$k_{DF} = 10 \tag{15}$$

To model q_L, we utilized the parsimonious equation from Han et al. (2022a). It calculates q_L as a function of incident PAR for each layer:

$$q_L = a_{q_L} e^{-b_{q_L} PAR} \tag{16}$$

where a_{q_L} and b_{q_L} are PFT-specific parameters derived from leaf-level pulse amplitude modulated (PAM) measurements in Han et al. (2022a). The qL-based model parameters (a_{q_L}, b_{q_L}, and $\phi_{PSIImax}$) are PFT-specific for each simulation and are listed in **Table A2**.

Finally, Fe_l in **Eq. (14)** is generated for the sunlit- (Fe_{sun}) and shaded- (Fe_{sha}) leaf groups, as J_a is also separately calculated by sunlit- ($J_{a,sun}$) and shaded- ($J_{a,sha}$) leaves:

$$F_{e,sun} = \frac{1 - \phi_{PSIImax}}{\phi_{PSIImax}(1 + k_{DF})} J_{a,sun} \frac{1}{q_L} \tag{17}$$

$$F_{e,sha} = \frac{1 - \phi_{PSIImax}}{\phi_{PSIImax}(1 + k_{DF})} J_{a,sha} \frac{1}{q_L} \tag{18}$$

2.2.4. Leaf-canopy radiative transfer

In the SCOPE model, a connection between photosynthesis and fluorescence is established through fluorescence emission efficiency (ε), for both sunlit and shaded leaves (ε_{sun} and ε_{sha}), respectively. These efficiencies are derived from corresponding ChlaF emission (Fe_{sun} and Fe_{sha}) and the fluorescence yield for a night-adapted leaf ($\phi_{F0,sun}$ and $\phi_{F0,sha}$) at the leaf-level:

$$\varepsilon_{sun} = \frac{F_{e,sun}}{\phi_{F0,sun}} \tag{19}$$

$$\varepsilon_{sha} = \frac{F_{e,sha}}{\phi_{F0,sha}} \tag{20}$$

$$\phi_{F0,l} = \frac{k_F}{k_F + k_{D,l} + k_P} \tag{21}$$

Within the fluorescence module, TOC SIF (in $\mu mol\ m^{-2}s^{-1}$) is simulated, accomplished by modeling the fluorescence radiative transfer process in a multi-layer canopy (up to 60 layers), as a function of the solar zenith angle and leaf orientation (van der Tol et al., 2009; Yang et al., 2021a).

2.2.5. Simulation setup

The FvCB biochemical model of photosynthesis in SCOPE provides J_a, which serves as an input for both NPQ- and qL-based SIF parameterization methods (**Fig. 2**). To check the first-order corrections of this photosynthesis modeling framework, we first compared the impact of model parameters on GPP simulations (i.e., foliar trait-specific vs. PFT-specific/default parameters). To assess the effects of model structure (i.e., qL-based vs. NPQ-based SIF parameterization), model parameters, and landscape heterogeneity (i.e., heterogeneous vs. homogeneous sub-grid) in simulating SIF, we subsequently conducted a series of simulations to evaluate their impacts.

Regarding the model structural impact, we compared the NPQ-based SIF parameterization (**Eq. (1-13)**) to the qL-based SIF parameterization (**Eq. (14-18)**). For the model parameter impacts, we performed control simulations (denoted as -ctrl) using default SCOPE parameters (except for $Vcmax_{25}$), and parametric simulations (denoted as -para) where we varied five photosynthetic parameters that are foliar trait-specific, including the leaf chlorophyll content (A and B, Cab), leaf carotenoid content (Cca), leaf dry matter content (fresh and dry, Cdm), equivalent water thickness in leaves (Cw), and the maximum carboxylation capacity at 25 °C ($Vcmax_{25}$). Considering that the $Vcmax_{25}$ varies significantly across and within PFTs (Rogers et al., 2014; Prikaziuk et al., 2023), we set the default $Vcmax_{25}$ values in control (-ctrl) simulations based on the PFT defined by Groenendijk et al., 2011. Furthermore, we chose BART and JERC to examine the impact of sub-grid heterogeneity, because these two sites show substantial sub-grid heterogeneity of LC and LAI within a 0.05°

SIF grid (**Fig. 1 and A2**). For instance, at BART, deciduous, evergreen, and mixed forest has area coverage of 27 %, 20 %, and 38 % of the 0.05° grid, respectively (**Table A1**). At JERC, evergreen forest and cultivated crops account for 44 % and 34 % of the 0.05° SIF grid. In terms of LAI, urban areas exhibit lower values than those of forests at BART, while at JERC, LAI of evergreen forests is much higher than that of adjacent croplands (**Fig. 3 and A2**). Compared to the other two sites, STEI tends to have a lesser degree of sub-grid heterogeneity within a 0.05° grid. For example, the two dominant LC types, deciduous forest (33 %) and mixed forest (21 %), show similar LAI values across the growing season (**Fig. 3**). Furthermore, wetlands cover the second-largest area at STEI, occupying 22% of the total LC.

Detailed information about the simulation setup can be found in **Table 3**, and the foliar trait-specific parameters used for simulations are provided in **Table 4**. Additionally, we set the constant coefficients (a_{q_L} and b_{q_L}) of qL-based models for each site following Han et al., 2022a and the Köppen climate classes (Peel et al., 2007; **Table A2**).

Table 3 Model simulation setup in this study.

Simulation experiments	Parameter values in FvCB	Parameterization of SIF
NPQ-ctrl	SCOPE default	**Eq. (1-13)**
qL-ctrl	SCOPE default	**Eq. (14-18)**

Simulation experiments	Parameter values in FvCB	Parameterization of SIF
NPQ-para	trait-specific (**Table 4**)	**Eq. (1-13)**
qL-para	trait-specific (**Table 4**)	**Eq. (14-18)**

Table 4 Foliar trait-specific parameters used for SCOPE simulations. Cab, Cca, Cdm, and Cw were obtained from the NEON plant foliar traits dataset (NEON1). $Vcmax_{25}$ refers to the maximum carboxylation capacity at 25 °C and was from TRY database. Other parameters (including leaf angle distribution, coefficients for Ball-Berry stomatal conductance model, Vcmax canopy extinction coefficient, and other leaf biochemical parameters) were set as the default values in SCOPE and not listed here (refer to Table C.6 in Prikaziuk et al., 2023). For brevity, DF refers to deciduous forest; EF refers to evergreen forest; MF refers to mixed forest; CRO refers to cropland.

Simulation experiments	Cab ($\mu g\ cm^{-2}$)	Cca ($\mu g\ cm^{-2}$)	Cdm ($\mu g\ cm^{-2}$)	Cw (cm)	NEON number of samples[b]	$Vcmax_{25}$ ($\mu mol\ m^{-2}s^{-1}$)

Control (-ctrl)	40	10	0.012	0.009	-	DF: 31; EF: 28; MF: 36; CRO: 49 (Groenendijk et al., 2011)
Parametric (-para):						
STEI:						
Baseline[a]	41.0	8.6	0.008	0.010	19	38
BART:						
Baseline	39.5	9.5	0.011	0.011	15	44
BART-DF	47.1	11.1	0.010	0.010	16	35
BART-EF	45.8	10.8	0.014	0.020	5	59
JERC:						
Baseline	27.1	6.8	0.011	0.008	8	49
JERC-EF	29.8	6.6	0.013	0.012	18	57
JERC-CRO	21.8	4.4	0.005	0.007	3	49

[a]Baseline simulations use photosynthetic parameter values from samples collected within the flux tower footprint, precisely within the dominant vegetation type around

the tower, as defined by NEON. The selection of samples follows a spatially balanced, random design.

[b]The number of samples is determined by NEON1.

In the context of parametric simulations (both NPQ-para and qL-para), we initially ran baseline simulations, i.e., no sub-grid heterogeneity was considered for each site. To achieve this, we used foliar traits (Cab, Cca, Cdm, and Cw) obtained from samples collected around each flux tower, and LAI from the CGLS grid (300 m × 300 m) corresponding to the location of the flux tower. Next, we ran simulations that explicitly account for sub-grid heterogeneity. To achieve this, we conducted end-member LC-specific simulations for a 0.05° grid at BART and JERC respectively. Each simulation has species-specific foliar traits (Cab, Cca, Cdm, and Cw) from samples located within the corresponding LC type (according to the NLCD class in NEON's metadata) and LC-specific LAI based on high-resolution LC datasets. In para simulations, we set the parameters Cab, Cca, Cdm, and Cw to the median values of 3-19 samples from NEON foliar trait datasets (during 2017-2022, **Table 4**, NEON1). $Vcmax_{25}$ values were set as the median values of 3-5 species from TRY database (Kattge et al., 2021), and detailed information about our criteria for selecting species can be found in **Sect. 2.3.3**. Next, we specified the LC-specific $Vcmax_{25}$ values for deciduous forest and evergreen forest, based on the unique species found in TRY database. Specifically, the $Vcmax_{25}$ values from species Pinus strobus (family: Pinaceae), Acer rubrum (family: Aceraceae), and Pinus palustris and Quercus hemisphaerica (family: Pinaceae and Fagaceae) found in TRY database were set for BART-evergreen forest (BART-EF), BART-deciduous forest (BART-DF), and JERC-

evergreen forest (JERC-EF), which are dominant species/families in BART and JERC, according to the NEON sampling (NEON2). The $Vcmax_{25}$ value for JERC-croplands (JERC-CRO), which the TRY database does not contain, was obtained from Groenendijk et al., 2011. While we set the trait-specific parameters as constant with time in SCOPE, we varied them spatially in the parametric simulations, derived from NEON measurements and TRY database (**Table 4**).

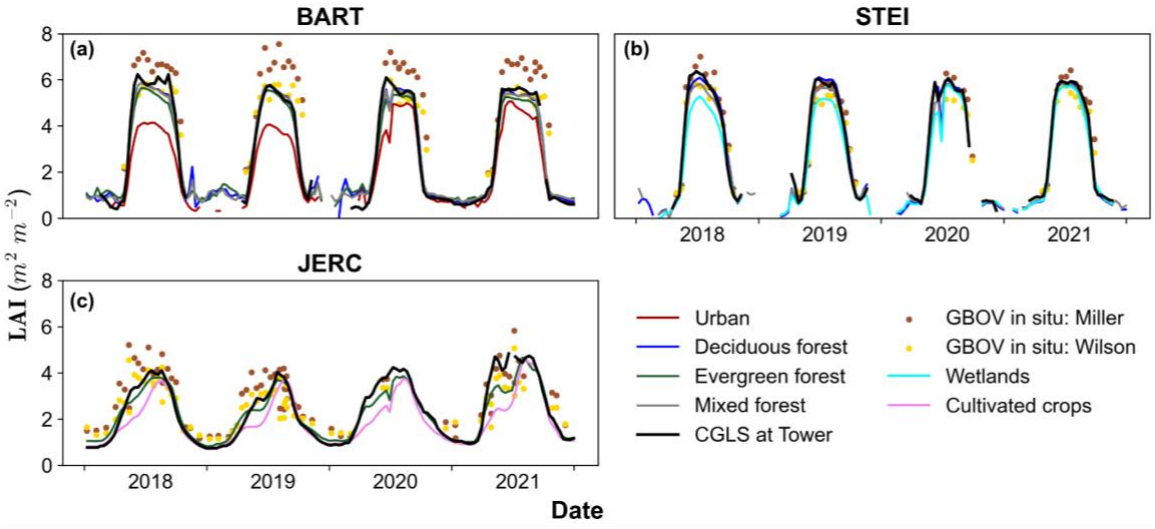

Fig. 3. Time series of LAI for the LC types with > 5 % area percentage (Table A1) in (a) BART, (b) STEI, and (c) JERC. The black line indicates the LAI at one CGLS 300 m × 300 m grid, which corresponds to the location of the flux tower. The brown and yellow dots indicate the in-situ measurements with two different methods (Miller 1967 and Wilson 1963), from the Ground Based Observations for Validation (GBOV) service.

All simulations were carried out with SCOPE v2.1 and driven by meteorological variables (air temperature, shortwave/longwave radiation, air pressure, and wind

speed) and satellite LAI at three NEON sites. **Table 2** lists all parameters and variables used in SCOPE. The meteorological variables are available half-hourly and are identical for all simulations. CGLS LAI data is provided around every 10 days and linearly interpolated to match the time step of the meteorological variables in the SCOPE model (i.e., half-hour in our simulations). For the LC-specific LAI time series, we first regridded the 30 m LC data to a 300 m resolution within the 0.05° grid, based on the dominant LC type in each grid (**Table A1**). Subsequently, we computed the average of 300 m LAI values under the same LC types for each 0.05° grid (**Fig. A2**). The model simulation performance was evaluated with metrics such as Nash-Sutcliffe modeling efficiency (MEF; Nash & Sutcliffe, 1970), root-mean-square error (RMSE), fractional bias (FB), and slope of the linear regression trend (**Text A2**).

2.3 Datasets

We utilized in-situ meteorological observations, LAI from satellite observations, and foliar traits from either NEON measurements (if available) or otherwise TRY database to drive SCOPE simulations. We then compared the simulated GPP with EC-derived GPP from the flux tower, and simulated SIF with satellite SIF at 0.05°. Refer to **Table 5** for all the datasets utilized, with detailed descriptions provided below.

Table 5 Datasets used for driving and evaluating SCOPE model simulations.

Data	Usage	Spatial resolution	Temporal resolution	Data sources	Reference

Meteorological forcing (air temperature, air pressure, relative humidity, wind speed, shortwave/longwave radiation)	Input	In-situ (tower footprint)	Half-hourly	NCAR-NEON v2	Lombardozzi et al., 2023
LAI	Input	300 m	10 days	CGLS v1.1	Fuster et al., 2021
Cab, Cca, Cdm, Cw	Input	In-situ (tower footprint)	Time-invariant	NEON DP1.10026.001	NEON1
$Vcmax_{25}$	Input	Within 300 km × 300 km radius	Time-invariant	TRY	Kattge et al., 2021
GPP from NEE-partitioning (daytime-based (Lasslop et al,	Model Evaluation	In-situ (tower footprint)	Half-hourly	NCAR-NEON v2	Lombardozzi et al., 2023

2010; Pastorello et al., 2021))

OCO-2 SIF	Model Evaluation	0.05°	16-day	SIF_{OCO-2_005}	Yu et al., 2018
TROPOMI SIF (Caltech)	Model Evaluation	0.05°	16-day	Caltech	Köhler et al. (2018)
TROPOMI SIF (ESA)	Model Evaluation	0.05°	16-day	European Space Agency	Guanter et al., 2021

2.3.1. Meteorological data

At terrestrial NEON sites, carbon, water, and energy fluxes are measured with the EC technique (Baldocchi 2003), in conjunction with observations of various meteorological variables. To address gaps in the original NEON data, ranging from hours to months, we employed gap-filled meteorological data from the NCAR-NEON dataset (Lombardozzi et al. 2023). Additionally, we incorporated ERA5-Land reanalysis as meteorological inputs for ONEFlux data pipeline (see **Sect. 2.3.5;** Muñoz Sabater, J., 2019). Specifically:

a) NCAR-NEON quality-controlled dataset, which provides half-hourly measurements of meteorological variables

(https://storage.googleapis.com/neon-ncar/listing.csv; Lombardozzi et al. 2023). These include air temperature, incident longwave/shortwave radiation, air pressure, precipitation, relative humidity, and wind speed.

b) ERA5-Land (Muñoz Sabater, J., 2019), with a spatial resolution of ~11 km. This dataset provides meteorological variables such as air temperature, dew point temperature of air, incoming shortwave/longwave radiation, total precipitation, surface pressure, and wind speed.

To correct and downscale the ERA5-Land meteorological data, from the original hourly to the half-hourly time step for each site, we followed the gap-filling procedure of Vuichard and Papale, 2015, separately for each month of a year. The resulting de-biased and gap-filled meteorological data were utilized as inputs for both the ONEFlux processing pipeline for NEE-partitioning and SCOPE model simulations.

2.3.2. Leaf area index (LAI) data

The LAI data comes from CGLS 300 m V1.1 (Fuster et al., 2021). This dataset was derived from PROBA-V and Sentinel-3 OLCI, covering the time period of January 2014 to the present and available around every 10 days. We used both RT5 (50 days consolidation of the Real-Time products (RT0), Jan. 2018 to June 2020) and RT6 (July 2020 to Dec. 2021) products (**Fig. 3**). To verify the CGLS LAI, we compared it with in-situ LAI (Brown et al., 2020), from the existing environmental monitoring networks Ground Based Observations for Validation (GBOV) service. GBOV estimates LAI using digital hemispherical photography (DHP) with two different methods ("Miller 1967" and "Wilson 1963"). Our results confirm the overall consistency between CGLS and GBOV LAI (**Fig. 3**), except for some differences in

BART (for the entire period) and in JERC during the spring of 2018 and 2019. Such differences may arise from an overestimation of the "Miller 1967" approach due to unstable estimates of Plant Area Index (PAI) under canopies with different leaf angle distributions (LAD) (Leblanc et al., 2014), as well as mismatch in spatial resolutions between the two datasets (ie., 300 m for CGLS and 20 m plots for GBOV).

2.3.3. Foliar traits data

We obtained foliar traits data from two sources: 1) NEON measurements of sampled foliage traits, including Cab, Cca, Cdm, and Cw (**Table 4,** NEON1). NEON established 16-22 plots (with a size of 20 m × 20 m) to sample foliage from dominant species at a 5-year cycle, either within the tower footprint or at stratified-random distributed plots nearby. The data we used consisted of 19-36 samples at each site gathered from dominant species at a 5-year cycle (2017 and 2022). The $Vcmax_{25}$ values come from the TRY database (Kattge et al., 2021). We selected records that belong to the same species as reported in the NEON sampling by plots (NEON2), within an area of 9×10^4 km^2 around the flux tower. For all species at each site, we collected a total of 73 records for baseline simulations, which represent 12 different species.

2.3.4. Land cover data

We utilized the USGS NLCD 2019 product at 30 m resolution, assuming no major LC changes within our study period (Dewitz, J., and U.S. Geological Survey, 2021) (**Fig. 1**). To derive LC-specific LAI, NLCD LC types were spatially aggregated to match the 300 m resolution of LAI (see **Sect. 2.3.2**). To achieve this, we calculated the

frequency of each LC type within a 300 m × 300 m grid and assigned the most-occurring LC type as the dominant LC type.

2.3.5. EC data from NEON towers

We obtained EC data from NEON flux towers to evaluate GPP simulations. These EC data are available half-hourly from 2018 to 2021, and are quality-controlled and gap-filled by the NCAR-NEON team (Lombardozzi et al. 2023). In order to obtain the standardized GPP that aligns with the FLUXNET2015 dataset, we partitioned the NEE utilizing the ONEFlux processing pipeline (Pastorello et al., 2021). ONEFlux employs both the daytime (Lasslop et al., 2010) and nighttime NEE-partitioning approaches (Reichstein et al. 2003). We compared GPP from these two methods and found negligible differences on the daily scale. Therefore, we used the daytime approach for our model evaluation. Here daily GPP estimations that lack the original NEE input for any single half-hour in a day were excluded from our analysis.

2.3.6. Satellite SIF products

We utilized two satellite SIF datasets for our model evaluation: National Aeronautics and Space Administration (NASA)'s OCO-2 and the TROPOMI onboard the Copernicus Sentinel-5P mission.

We used a spatially contiguous OCO-2 SIF product with a 16-day temporal and 0.05° spatial resolution (Yu et al., 2018). This product employed machine learning algorithms to fill the spatial gaps of the native OCO-2 SIF with the Moderate Resolution Imaging Spectrometer (MODIS) Bidirectional Reflectance Distribution Function (BRDF)-corrected reflectance. The OCO-2 SIF was evaluated at 757 nm and

further converted to 740 nm, at which TROPOMI SIF is retrieved, by multiplying with a wavelength scaling factor (1.69) following Wen et al. (2020).

TROPOMI SIF, with a daily global coverage at a resolution of 7 km × 3.5 km (at nadir) in the near-infrared range. In this study, we used two TROPOMI SIF retrievals from Caltech (Köhler et al. 2018) and ESA (Guanter et al., 2021), respectively. To match the resolutions with the OCO-2 data, we aggregated the TROPOMI data to the same 16-day 0.05° resolutions, based on the Generic Gridding Routine (https://github.com/cfranken/gridding). Prior to regridding to 0.05°, we applied multiple criteria to screen both TROPOMI SIF retrievals following Guanter et al. (2021), which exclude the data with cloud fraction greater than 80 %, view zenith angle (VZA) greater than 60°, sun zenith angle (SZA) greater than 70°, and absolute SIF values larger than 10 $W\ m^{-2}\ \mu m^{-1}\ sr^{-1}$. In addition, TROPOMI has a wide swath (across 2600 km), with VZA varying from nadir to ~60° (Veefkind et al., 2012; Guanter et al., 2021). Joiner et al. (2020) reported that this large VZA variation would produce systematic differences in SIF; fortunately, such VZA-introduced SIF variation can be mitigated if sufficient samples in a grid are included for averaging. To ensure consistency in sun-sensor-viewing geometry between level 3 SIF data and model simulations, we set VZA to 0 in SCOPE.

3. RESULTS

3.1. Impact of photosynthetic parameters on simulations of GPP

First, we assessed the degree to which the five photosynthetic parameters (Cab, Cca, Cdm, Cw, and $Vcmax_{25}$) impact the NPQ-based simulation of GPP for three forest sites over diurnal, daily, and seasonal scales, from 2018 to 2021 (**Fig. 4-6**). Our analysis indicated that the NPQ-based GPP simulations with adjusted photosynthetic parameters (NPQ-para) displayed a more robust agreement than the simulations with default photosynthetic parameters (NPQ-ctrl), when compared with EC-derived GPP, which exhibited more day-to-day variation than the model simulations, across all three sites during peak seasons. For instance, during the period from the 150th to the 300th day of the year, the mean differences between NPQ-para and EC-derived GPP in BART, STEI, and JERC were 8.5 %, 5.7 %, and 6.9 %, respectively (**Fig. 4c, 5c, 6c**). In contrast, NPQ-ctrl simulations exhibited mean differences of -12.8 %, -9.8 %, and -26.7 % for the same sites. Note that EC-derived GPP in STEI showed slightly lower values during the dormant season (e.g., before the 120th day, **Fig. 5c**), compared to both NPQ-ctrl and NPQ-para simulations. This variation might be attributed to a data quality issue of the observed NEE, i.e., the unrealistically rapid rises in spring during 2018 and 2019, which did not occur in 2021 (**Fig. 5a**). Analyzing diurnal variations, as anticipated, the NPQ-para GPP showed higher consistency with EC-derived GPP in STEI (3.7 %) and JERC (1.1 %) during the growing seasons (dates between average green increasing and average green minimum day), when compared to NPQ-ctrl (with differences of -11.5 % and -31.2 % in **Fig. 5b and 6b**). While the GPP diurnal difference during growing seasons in BART NPQ-para (11.7 %) was slightly higher

than NPQ-ctrl (-10.4 %). Meanwhile, the diurnal differences for the whole year between NPQ-para (10.8 %) and EC-derived GPP were much more pronounced than that for NPQ-ctrl (-5.7 %) in STEI (**Fig. 5b**), also likely due to the generally lower EC-derived GPP values during the dormant seasons of 2018 and 2019.

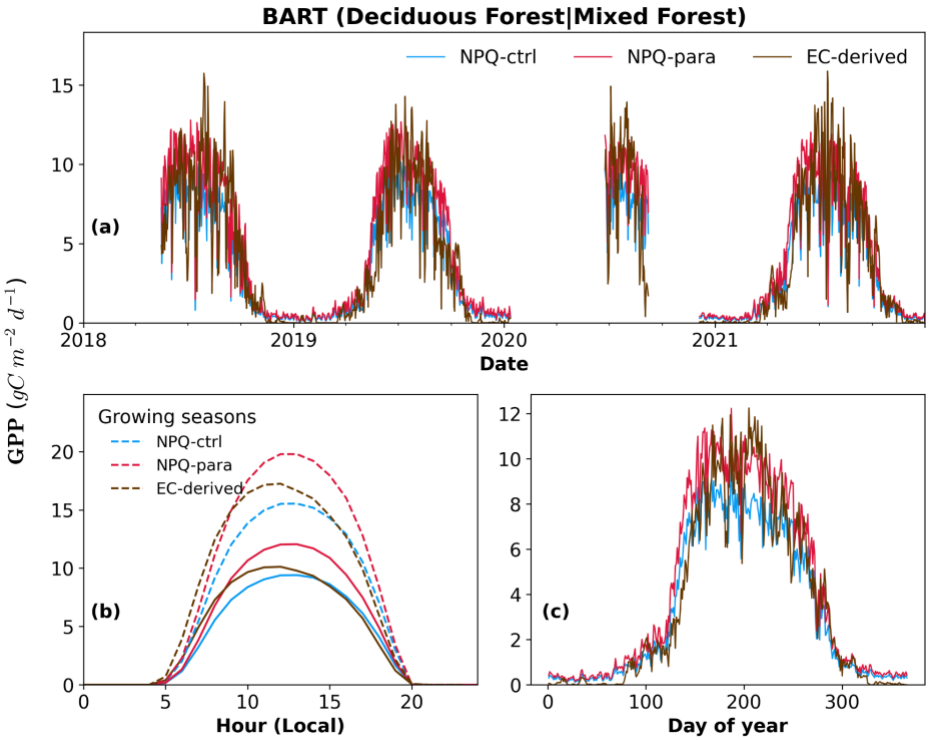

Fig. 4. Temporal dynamics of GPP from NPQ-ctrl, NPQ-para, and EC-derived, respectively, at BART from 2018 to 2021: (a) daily GPP; (b) mean diurnal cycle from Jan. 2018 to Dec. In **(b)**, the dashed lines represent the mean diurnal cycle during the growing seasons. The growing seasons are referred to as dates between average green increasing and average green minimum day according to NEON metadata. Note that we excluded the daily values that were averaged from all gap-filling timestamps.

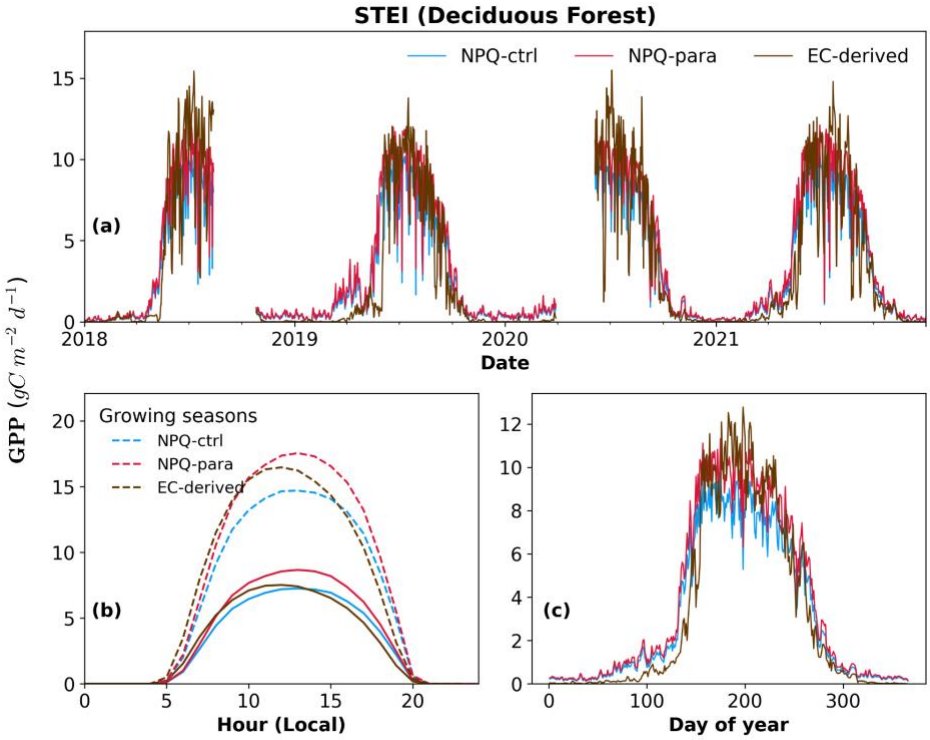

Fig. 5. Same as **Fig. 4**, but for STEI.

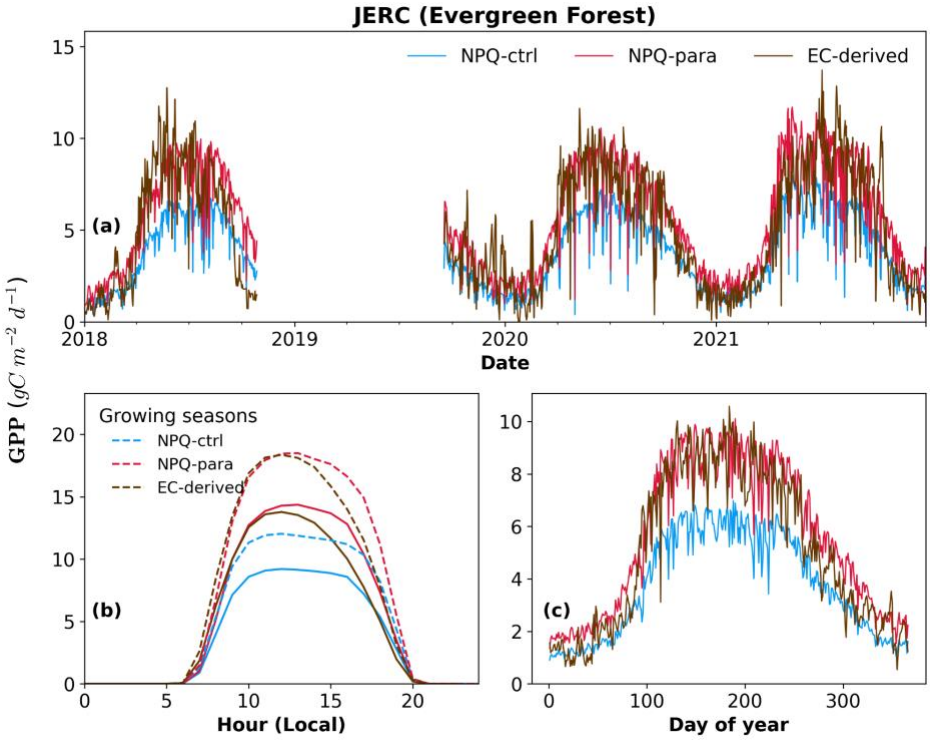

Fig. 6. Same as **Fig. 4**, but for JERC.

Furthermore, we compared the performance of NPQ-para and NPQ-ctrl simulations, in replicating daily GPP values with EC-derived GPP. While the agreement between NPQ-ctrl GPP simulations and EC-derived GPP appeared reasonable (**Fig. 7a, c, e,** yielding MEF > 0.47 (higher values mean greater consistency); regression slope = 0.54-0.70), NPQ-para simulations demonstrated better performance (**Fig. 7b, d, f**, MEF > 0.67, slope = 0.77-0.82), which indicated the positive impact of adjusting model parameters on GPP simulations.

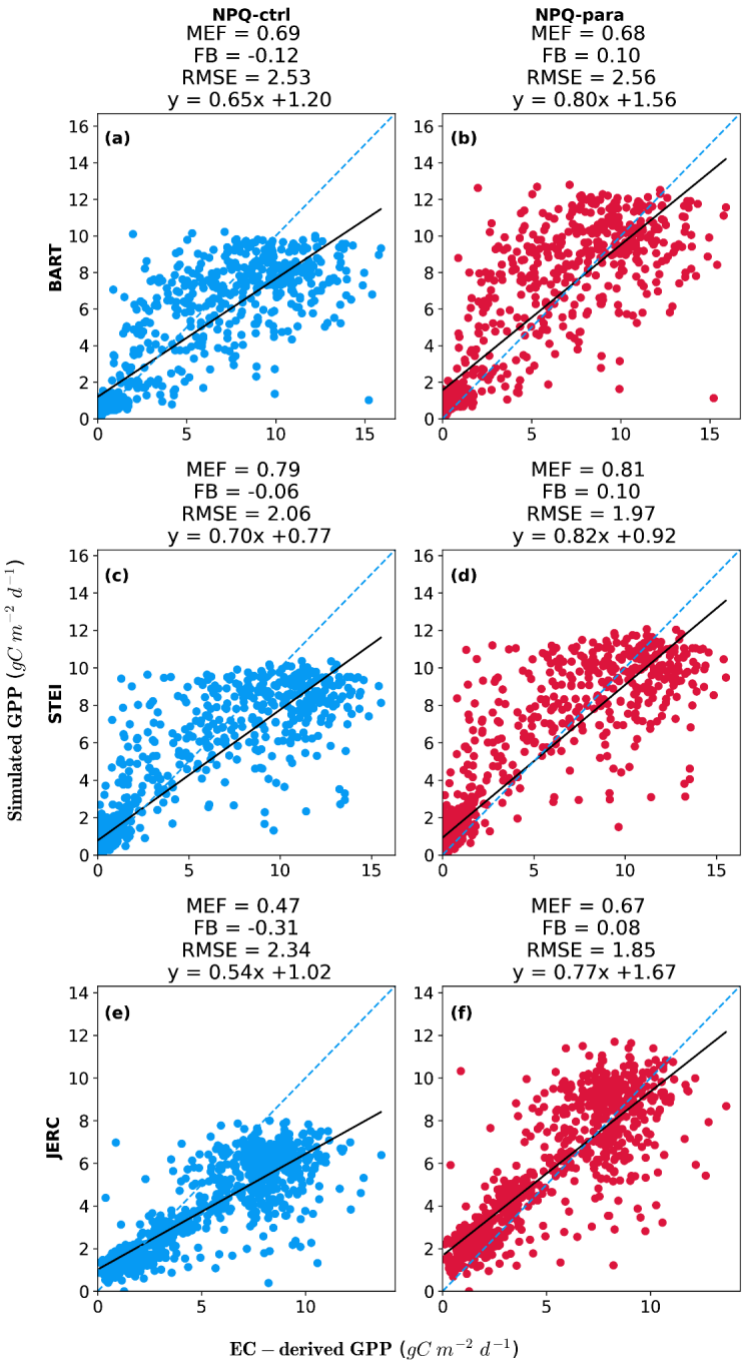

Fig. 7. Comparison of (a, c, e) NPQ-ctrl and (b, d, f) NPQ-para simulations of daily mean GPP against EC derived observations at (a, b) BART, (c, d) STEI, and (e, f) JERC during 2018–2021. The blue dashed line represents a 1:1 relationship, and the black line represents the regression relationship of simulations with observations. The

values of MEF (ranging from −Inf to 1), FB, RMSE ($gC\ m^{-2}\ d^{-1}$), and the regression equations are on top of each panel.

Overall, these results suggest that foliar trait-specific parameters could potentially enhance the performance of GPP simulations, at the diurnal, daily, and seasonal scales. Moreover, by comparing the impact of these parameters on GPP simulations, we have demonstrated the first-order corrections inherent in SCOPE's photosynthesis modeling.

3.2. Effect of model structure and parameters on SIF simulation

Next, we examined the impact of model structure (i.e., NPQ-based vs qL-based) on the realism of SIF simulations at three NEON sites. **Fig. 8-10** depict comparisons of SIF temporal dynamics over diurnal, 16-day, and seasonal scales. Our results indicated that the NPQ-based SIF parameterization generated better agreement with OCO-2 and TROPOMI SIF during peak seasons at all three sites; conversely, qL-based SIF parameterization was superior in capturing the temporal dynamics of non-peak and shoulder seasons. Notably, the NPQ-based SIF parameterization consistently yielded higher SIF values than qL-based parameterization throughout the day-time and TROPOMI observations during the overpass time, except for qL-para in JERC (**Fig. 8b, 9b, 10b**). These findings demonstrated the significant impact of model structure on SIF simulations. The cause of the incapability of qL-based SIF parameterization in reproducing the observed peak-season and daytime SIF magnitude is examined in-depth in **Sect. 3.3**.

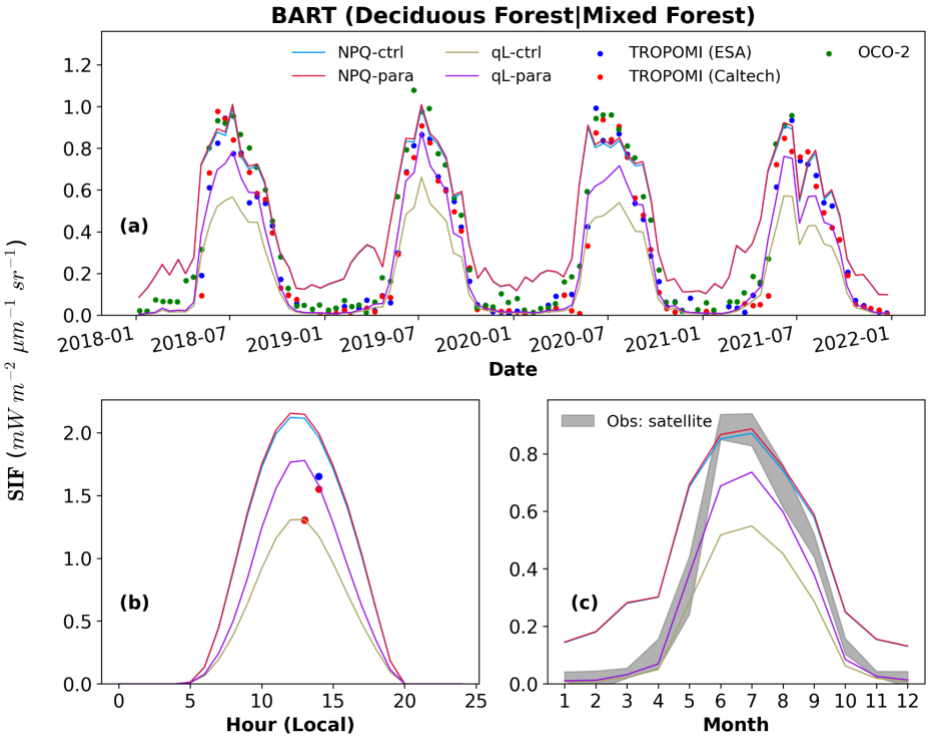

Fig. 8. Temporal dynamics of SIF from NPQ-ctrl, NPQ-para, qL-ctrl, qL-para, and satellites, respectively, at BART from 2018 to 2021: (a) 16-day time series; (b) mean diurnal cycle from June 2018 to June 2021; and (c) mean seasonal cycle from June 2018 to June. The shaded regions in **(c)** represent the upper and lower bounds of TROPOMI and OCO-2 observations. Note that in **(b)** we compared with instantaneous TROPOMI SIF and excluded dates without satellite observations.

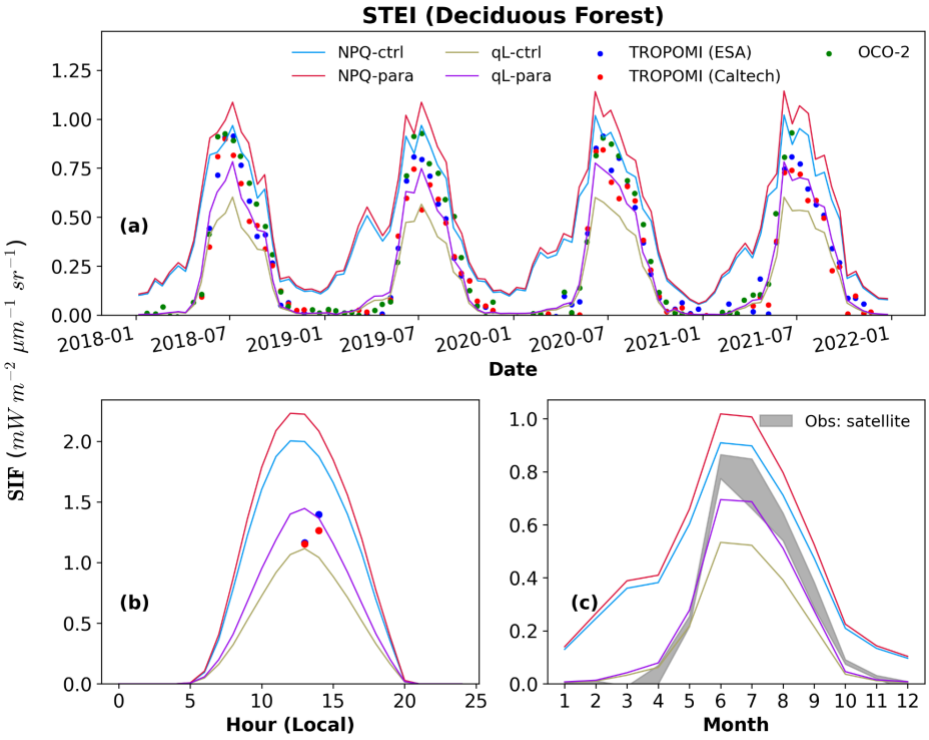

Fig. 9. Same as **Fig. 8**, but for the site STEI.

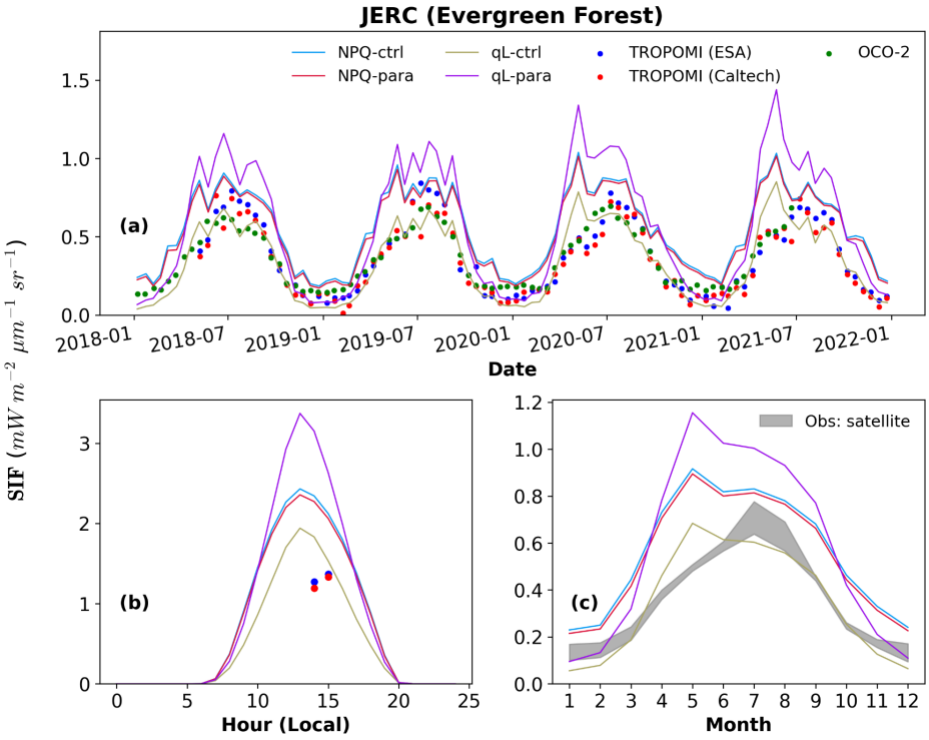

Fig. 10. Same as **Fig. 8**, but for the site JERC.

Furthermore, we evaluated the effect of model parameters on SIF simulations. The qL-para simulations, which used foliar trait-specific photosynthetic parameters, showed improved consistency with OCO-2 and TROPOMI observations, especially in the SIF magnitude during the peak seasons (May to Sept.) compared to the qL-ctrl simulations (**Fig. 8c and 9c**). In BART and STEI, the qL-para simulations exhibited substantially better overall performance than qL-ctrl (i.e., higher MEF, lower Bias and RMSE, slope values closer to 1, **Fig. A3-A5**).

On the other hand, the differences between NPQ-para and NPQ-ctrl across the diurnal and seasonal cycles for all three sites were relatively small (less than 10 % in **Fig. 8-10**). This suggests that the simulated SIF is relatively insensitive to the model

parameters when using NPQ-based SIF parameterization. For example, NPQ-para and NPQ-ctrl exhibited almost identical simulated diurnal and seasonal dynamics, while there are substantial differences in simulated SIF temporal dynamics between qL-para and qL-ctrl, with discrepancies ranging from 30 % to 70 % in these sites. These results indicate that simulated SIF is more sensitive to photosynthetic parameters when using qL-based SIF parameterization.

3.3. Spatial heterogeneity effects on SIF simulation

In both BART and JERC, there are still discrepancies in SIF simulation between qL-para experiments and satellite SIF, including both underestimation and overestimation of SIF. For example, there is a low bias in the peak season of 2020 in BART (**Fig. 8**), and large positive biases in the early summers of 2020 and 2021 in JERC (**Fig. 9**). To pinpoint the cause of these remaining discrepancies, we explicitly conducted LC-specific SIF simulations, which have LC-dependent model parameters (trait-specific) and LAI (**Table 4, and Fig. 3**). We denoted these simulations as BART-EF, BART-DF, JERC-EF, and JERC-CRO.

We first aggregated the CGLS LAI based on the LC type in each 0.05° × 0.05° study domain and compared it with the LAI grid at the tower (**Fig. A2**). There were only minor LAI differences between deciduous and evergreen forests in BART, and the aggregated LAI of these two forests was similar to LAI at the tower site (LC: deciduous forest; **Fig. 3**). However, there are major discrepancies between the evergreen forest and cultivated crops in JERC (~20 %), and both were lower than the LAI at the tower (LC: evergreen forest; **Fig. 3**). In addition, the growing season of the evergreen forest is earlier and longer than the cultivated crops in JERC. These LC-

dependent differences likely contributed to the overestimation and incorrect seasonal cycle of SIF simulations in JERC.

Furthermore, there are substantial variations in model parameters (Cab, Cca, Cdm, Cw) obtained from different LC plots around the same tower at both sites (**Fig. A6**). These parameter differences range up to 80 % for BART and 55 % for JERC. All parameters of BART evergreen forest were higher than those around the tower (14 to 81 %), while those for JERC cultivated crops were lower than those around the tower (-6 to -55 %).

Fig. 11-12 show the comparisons of area-weighted LC-specific SIF simulations for the 0.05° grid with TROPOMI and OCO-2 observations. Our results revealed much enhanced agreement with observations relative to the original baseline SIF simulations. Specifically, LC-specific qL-para simulations can capture the seasonal dynamics in both sites, not only capturing the dormant and shoulder season SIF, but also the peak season SIF magnitude. These results demonstrated that spatial heterogeneity may lead to huge biases in model simulated SIF.

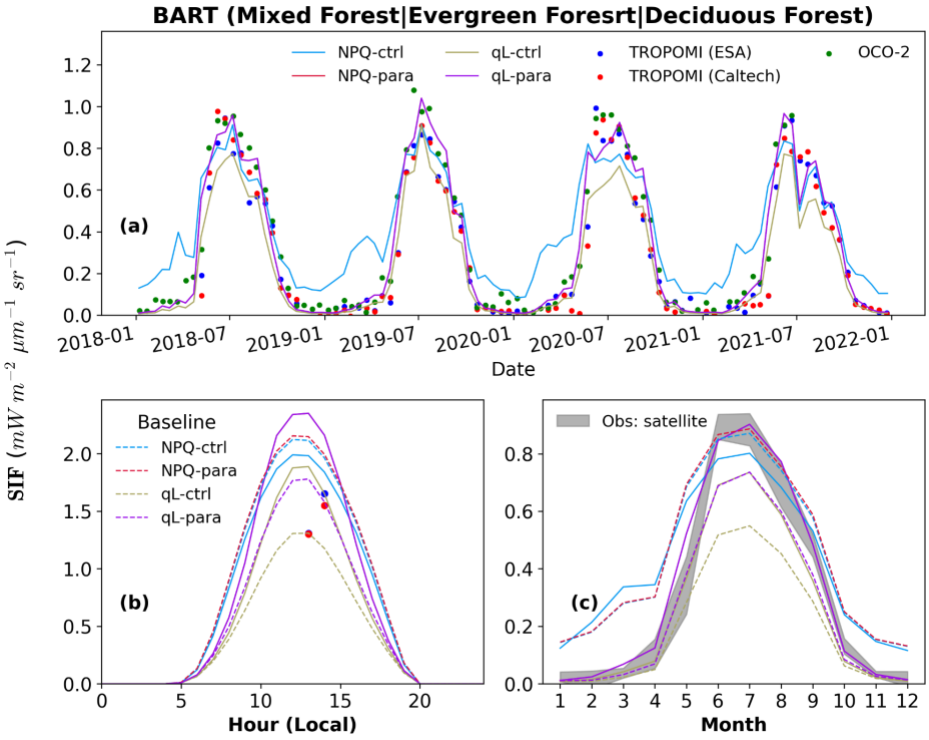

Fig. 11. Same as **Fig. 8**, but for using the LC-specific LAI and foliar trait values in simulations, with a weighting of 35.5 % for evergreen forests and 56.5 % for the deciduous forests. The weighting values are the percentage of aggregated LC area within the 0.05° × 0.05° domain (**Table A1**). Dashed lines represent **(b)** mean diurnal cycle and **(c)** mean seasonal cycle of the baseline simulations.

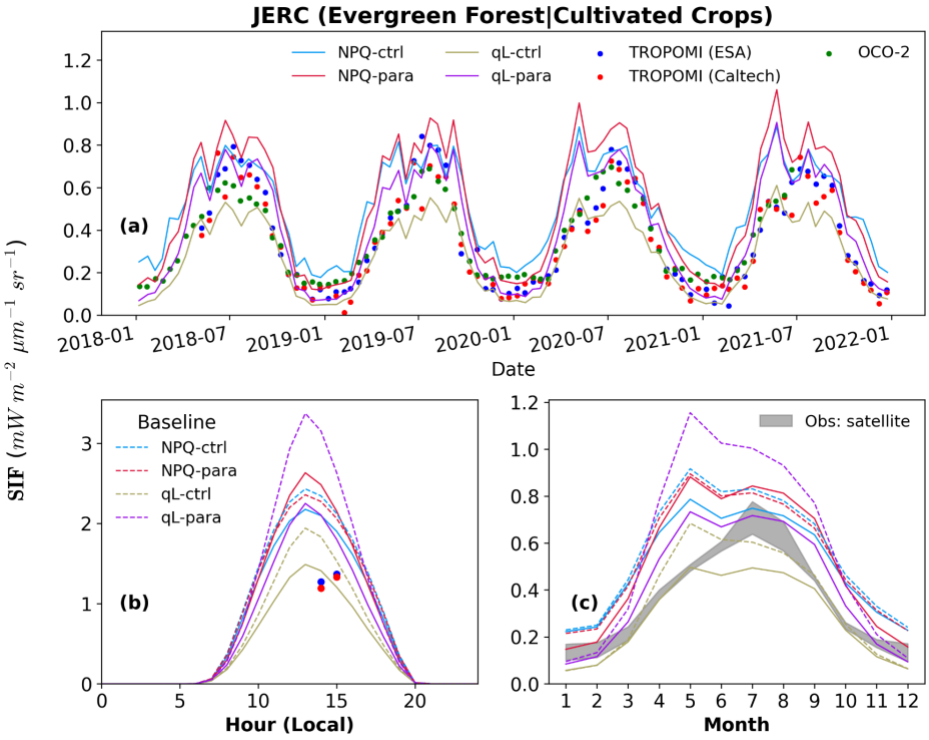

Fig. 12. Same as **Fig. 11**, but for JERC, with a weighting of 56 % for evergreen forests and 38 % for cultivated crops.

We also compared the LC-specific SIF simulations with the baseline simulations at the diurnal and seasonal scales (**Fig. 11b-c, 12b-c**). qL-para exhibited substantial differences/improvements over qL-ctrl (**Fig. A7-A8**), indicating that foliar trait-specific parameters are necessary for augmenting SIF simulations when sub-grid heterogeneity was explicitly accounted for. In contrast, NPQ-based SIF parameterization did not exhibit much improvement compared to baseline, indicating its insensitivity to sub-grid heterogeneity treatment. Specifically, the differences between LC-specific and baseline in NPQ-ctrl and NPQ-para remained below 15 % in both BART and JERC over diurnal cycles during the growing seasons and seasonal

cycles throughout the seasons. Conversely, notable improvements were observed in the case of qL-ctrl and qL-para, with differences reaching approximately 32 % and 30 % in BART, and -18 % and -29 % in JERC, respectively.

4. DISCUSSION

In this study, we demonstrated the efficacy of qL-based model structure, foliar trait-specific parameters, and landscape heterogeneity to address the complexities and uncertainties of SIF modeling, using three NEON forest sites as a testbed. Through achieving these objectives, our study contributes to the development of more accurate models of SIF and GPP, which are essential for improving our understanding of the global carbon budget and climate change. In the following sections, we further highlight the benefits of our MLR qL-based SIF parameterization in simulating SIF with foliar trait-specific photosynthetic parameters (Sect. 4.1). Following that, we compare our model parameterization regime and its impact with other studies (Sect. 4.2). Then, we also discuss the impact of parameter k_{DF} in SIF simulations and several caveats in this study, which warrant future research (Sect. 4.3, 4.4).

4.1. Advantages of mechanistic light reaction (MLR) model

We have validated that the SIF simulation can be substantially improved by utilizing the MLR-SIF qL-based model and trait-specific photosynthetic parameters. Our evaluation metrics consistently outperform NPQ-based SIF parameterization and PFT-specific/default parameters, even before considering sub-grid heterogeneity. For BART and STEI, our metrics show improved performance with MEF = 0.86-0.90/0.91-0.95, |FB| < 31/17 %, RMSE = 0.10-0.13/0.07-0.10 $mW\ m^{-2}\ \mu m^{-1}\ sr^{-1}$, slope = 0.78-0.82/0.78-0.88, respectively (**Fig. A3-A5**). Once we considered the spatial heterogeneity of PFTs within the satellite grid, the simulation biases can be further reduced to less than 0.18 $mW\ m^{-2}\ \mu m^{-1}\ sr^{-1}$ (RMSE < 0.12 $W\ m^{-2}\ \mu m^{-1}\ sr^{-1}$, **Fig. A7-A8**).

The MLR-SIF model has two advantages: first, it can directly account for the influence of incoming radiation (i.e., PAR), rather than exclusively calculating NPQ with an empirical nonlinear function of the degree of light saturation; second, qL-based SIF parameterization can generate better temporal dynamics during non-peak and shoulder seasons, as well as capture the peak season magnitude with sub-grid model parameters. On the other hand, the NPQ-based SIF parameterization only shows strong agreement during the peak season and offers minimal improvement when considering trait-specific photosynthetic parameters and spatial heterogeneity. Several studies have also investigated the use of the MLR-SIF model to estimate GPP, with SIF as a shortcut that could reduce the degree of model structure complexity and parameter uncertainty, compared to the NPQ-based model. At the leaf level, Han et al. (2022a) confirmed that the MLR-SIF model can accurately reproduce photosynthesis for both C_3 and C_4 species under diverse environmental conditions, with a regression slope of around 0.80 when compared with measurements. At the canopy level, the MLR-SIF can also be applied by replacing q_L and $\phi_{PSIImax}$ with the quantum yield of PSII and NPQ mechanistically. Liu et al. (2022) reported estimations with a half-hourly RMSE of 5.62 $\mu mol\ m^{-2}\ s^{-1}$ (R^2 = 0.85) and a daily RMSE of 3.25 $\mu mol\ m^{-2}\ s^{-1}$ (R^2 > 0.90). While further tests need to be carried out across diverse ecosystems, the simplicity of the MLR-SIF model offers a promising opportunity to apply it at larger regional and global scales, either to infer GPP or SIF.

4.2. Comparison to existing model parameterization

In this study, we examined five foliar traits in the SCOPE, including Cab, Cca, Cdm, Cw, and $Vcmax_{25}$, without any parameter calibration to understand the roles of these

parameters in GPP and SIF simulations. Although we set these parameters as time-invariant in simulations, it has been found that seasonally varying parameters like $Vcmax_{25}$ only had a slight impact on simulation improvement in SCOPE (Prikaziuk et al., 2023). The $Vcmax_{25}$ the control simulations we used were PFT-specific reported by Groenendijk et al. 2011, instead of the default SCOPE values as the other four traits used. Other studies have already compared PFT-specific and default SCOPE parameters. For instance, Prikaziuk et al. 2023 opted for $Vcmax_{25}$ values of 10 to 40 ($\mu mol\ m^{-2}s^{-1}$) for forest PFTs and 50 ($\mu mol\ m^{-2}s^{-1}$) for croplands in GPP simulations and found that the PFT-specific $Vcmax_{25}$ outperformed the default SCOPE values (60 $\mu mol\ m^{-2}s^{-1}$) in GPP simulation.

Previous studies have also examined the impact of Cab, Cca, Cdm, and Cw on SIF simulations in the SCOPE model. For example, Koffi et al. 2015 found that SIF only increases with Cab when Cab is smaller than 20 $\mu g\ cm^{-2}$ while it almost remains constant as Cab further increases. Verrelst et al. 2015 reported that Cab not only accounts for SIF variability, but also has a profound effect on its spectra shape. Cdm and Cw were reported to have a negligible impact on the visible and near-infrared spectral features when spectra are less than 700 nm (Verrelst et al., 2015). These findings indicate that SIF model parameter uncertainties can have different impacts on SIF simulations, and well-constrained parameters can reduce the bias in SCOPE simulations (Prikaziuk et al., 2023). A more nuanced investigation (such as the foliar trait-specific values used in this study) on the model parameters is certainly needed to better characterize the temporal dynamic of SIF.

We found that foliar trait-specific photosynthetic parameters had a significant impact on SIF simulations only for qL-based SIF parameterizations, up to 70 % in the diurnal cycle (**Fig. 10b**). This finding contrasts with earlier studies, e.g., Koffi et al. 2015, which found that SIF only slightly increases with an increase of $Vcmax_{25}$ (especially for $Vcmax_{25}$ > 125 $\mu mol\ m^{-2}s^{-1}$). However, it supports the theoretical arguments by Han et al. 2022b, which reported linear relationships between q_L ×SIF and photosynthetic capacity parameters ($Vcmax_{25}$, and maximum J_a ($Jmax_{25}$)). We argue that the previously reported insensitivity of simulated SIF to $Vcmax_{25}$ (Koffi et al., 2015) is likely a consequence of the NPQ-based SIF parameterization. This finding indicates that the model structures can have a considerable impact on parameter uncertainties (Bonan et al., 2011). The NPQ-based SIF parameterization can significantly mute the sensitivities of parameters and sub-grid heterogeneity treatments.

4.3. Impact of k_{DF} in SIF simulations

In our qL-based SIF parameterization, the estimation of Fe requires parameter k_{DF} as an input, which we set to 10 in our study. In the original SCOPE model, the k_{DF} was assumed to be larger than 17.5, according to **Eq. (8-9)**, and was assumed to be 19 by Gu et al. (2019). However, its value has been set to around 9 in other studies (e.g., Pfundel, 1998; Atherton et al., 2016; Han et al., 2022a). Liu et al. (2022) developed a leaf-level concurrent instrument to estimate the k_{DF} (see their Fig. 3). By measuring upward and downward fluorescence emission and the maximum fluorescence quantum yield (presented in photon flux density) of winter wheat, they obtained the maximum

fluorescence quantum yield of PSII under dark-adapted conditions. Their results showed a small range of k_{DF} variation, and its value should be smaller than 11.0.

To assess the impact of the choice of k_{DF} value on the performance of the baseline qL-based SIF parameterization, we performed a sensitivity analysis of SIF to k_{DF} set as 5, 10, and 20 (**Fig. 13**). The MEF between simulations and observations are above 0, except for the simulations in JERC that are incapable of reproducing the observed peak-season and daytime SIF magnitude (**Sect. 3.2**). As k_{DF} changed, the bias was up to 92 %, RMSE varied from 0.07 to 0.33 $mW\ m^{-2}\ \mu m^{-1}\ sr^{-1}$, and slope values varied from 0.4 to 2.1 (all exclude the metrics in JERC when k_{DF}=5). These results confirmed that k_{DF} had a significant impact on the accuracy of SIF simulations, and further investigation is needed to understand its uncertainty.

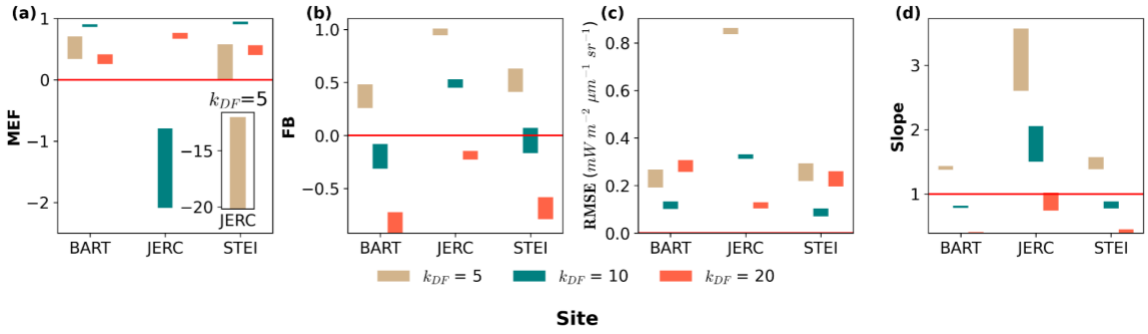

Fig. 13. Performance of baseline qL-para simulations of 16-day mean SIF against satellite observations (TROPOMI ESA, TROPOMI Caltech, and OCO-2) as k_{DF} changed: (a) MEF; (b) FB; (c) RMSE; and (d) Slope values. The upper/lower bounds of the bar represent corresponding maximum/minimum statistical metrics. The red horizontal line represents the best agreement that can be achieved with observations for each metric.

4.4. Other caveats

This study aimed to assess the performance of SIF and GPP simulations in the SCOPE model across three NEON forest sites, which present diverse climate and LC conditions. However, there are some aspects that need to be further investigated.

First, SIF simulation during the peak season of 2020 in BART was underestimated (**Fig. 8a and 11a**), due to high vapor pressure deficit (VPD) input (**Fig. A9**), which can disrupt the balance between carboxylation and light harvesting, leading to a decrease in SIF (Paul-Limoges et al., 2018). Moreover, anomalies in the meteorological variables (**Fig. A8 and A9**) could also result in a high-magnitude offset in JERC. Similarly, in the early summers of 2020 and 2021 in JERC (**Fig. 10a and 12a**), positive anomalies may be explained by earlier warming caused by higher temperatures and shortwave radiation (**Fig. A10**). These positive anomalies could cause an overestimation of PAR and temperature-dependent parameters, such as Vcmax, which can ultimately result in a broadband SIF flux overestimation of more than 30 % (Verrelst et al., 2015). In addition, this double peak season pattern is also shown in the LAI time series (**Fig. 3c**), indicating the potential uncertainties from satellite LAI input. Therefore, SIF modeling under external forcing anomalies needs to be further constrained, including meteorological variables and LAI.

Furthermore, the study utilized photosynthetic parameters sourced from NEON measurements and the TRY database, and treated them as time-invariant. However, this approach neglects their seasonal variability. Recent studies have emphasized the

importance of incorporating the seasonality of Vcmax, chlorophyll content, and carotenoid content into simulations of GPP and SIF fluxes (Wang et al., 2019; Wu et al., 2022; Magney et al., 2019). Additionally, we opted to use the median of NEON trait measurements from all species instead of dominant species within certain plots (i.e., around each flux tower or within corresponding LC types, as shown in **Fig. A6**), which may introduce uncertainty at the ecosystem scale. For example, Wang et al. (2022) employed leaf spectroscopy across NEON domains and uncovered substantial inter- and intra-species trait variations that depend on local environmental factors (e.g., elevation, mean annual temperature and precipitation).

Lastly, by selecting PFT-specific coefficient sets for the qL-based SIF parameterization (e.g., a_{q_L} and b_{q_L}), we did not consider differences within PFTs. Moreover, these coefficients from Han et al. (2022a) appeared insufficient for evergreen forests. Since no set of coefficients could represent evergreen trees in the temperate climate class, we chose broadleaf deciduous trees (BDT) for the JERC-EF simulation regime. Therefore, future research is needed to examine the degree to which these qL-related parameters vary in PFTs or across species within the same PFT, and how such variations may impact SIF simulations.

5. CONCLUSIONS

In this study, we looked at how different model structures (qL-based vs NPQ-based parameterization), parameter uncertainties (foliar trait-specific vs PFT-specific/default parameters), and landscape heterogeneity (heterogeneous vs. homogeneous sub-grid) can affect SIF modeling. We used three NEON forest sites as a testbed and conducted simulations within the SCOPE radiative transfer framework. Our findings revealed that qL-based SIF parameterization outperformed NPQ-based in capturing the observed temporal dynamics from satellite SIF (TROPOMI and OCO-2), especially during non-growing seasons. Additionally, using foliar trait-specific parameters significantly improved the performance of both GPP and SIF (qL-based) simulations at both daily and seasonal levels. More importantly, the choice of SIF model structure can impact the parameter sensitivity; for example, as qL-based SIF parameterization is more sensitive to variations in model parameters, e.g., the discrepancy between the qL-para and qL-ctrl simulations are larger than 30 % in all the three sites, but are less than 15 % between NPQ-para and NPQ-ctrl. Finally, we found that only when model parameters were set to be foliar trait-specific and spatial heterogeneity was accounted for, the qL-based SIF parameterization produced the closest agreement between simulated and observed SIF across ecoregions and timescales, from diurnal, to seasonal, and interannual scales. In summary, this study underscores the importance of physiological-meaningful combinations of model structures, foliar trait-specific parameters, and sub-grid spatial heterogeneity in the realism and accuracy of SIF simulations. These insights are vital for advancing our knowledge in evaluating,

understanding, and predicting SIF and GPP, and ultimately the global carbon budget and its response to climate change.

REFERENCE

Atherton, J., Nichol, C.J., & Porcar-Castell, A. (2016). Using spectral chlorophyll fluorescence and the photochemical reflectance index to predict physiological dynamics. Remote Sensing of Environment, 176, 17-30

Baldocchi, D.D. (2003). Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. Global Change Biology, 9, 479-492

Bonan, G.B., & Doney, S.C. (2018). Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models. Science, 359

Bonan, G.B., Lawrence, P.J., Oleson, K.W., Levis, S., Jung, M., Reichstein, M., Lawrence, D.M., & Swenson, S.C. (2011). Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. Journal of Geophysical Research, 116

Brown, L.A., Meier, C., Morris, H., Pastor-Guzman, J., Bai, G., Lerebourg, C., Gobron, N., Lanconelli, C., Clerici, M., & Dash, J. (2020). Evaluation of global leaf area index and fraction of absorbed photosynthetically active radiation products over North America using Copernicus Ground Based Observations for Validation data. Remote Sensing of Environment, 247

Camino, C., Gonzalez-Dugo, V., Hernandez, P., & Zarco-Tejada, P.J. (2019). Radiative transfer Vcmax estimation from hyperspectral imagery and SIF retrievals to assess photosynthetic performance in rainfed and irrigated plant phenotyping trials. Remote Sensing of Environment, 231

Chen, R., Liu, L., Liu, Z., Liu, X., Kim, J., Kim, H.S., Lee, H., Wu, G., Guo, C., & Gu, L. (2024). SIF-based GPP modeling for evergreen forests considering the seasonal variation in maximum photochemical efficiency. Agricultural and Forest Meteorology, 344

Collatz, G. J., Ball, J. T., Grivet, C., & Berry, J. A. (1991). Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer. Agricultural and Forest meteorology, 54(2-4), 107-136.

Cui, T., Sun, R., Xiao, Z., Liang, Z., & Wang, J. (2020). Simulating spatially distributed solar-induced chlorophyll fluorescence using a BEPS-SCOPE coupling framework. Agricultural and Forest Meteorology, 295

De Cannière, S., Herbst, M., Vereecken, H., Defourny, P., & Jonard, F. (2021). Constraining water limitation of photosynthesis in a crop growth model with sun-induced chlorophyll fluorescence. Remote Sensing of Environment, 267

Dewitz, J., and U.S. Geological Survey, 2021, National Land Cover Database (NLCD) 2019 Products (ver. 2.0, June 2021): U.S. Geological Survey data release, https://doi.org/10.5066/P9KZCM54.

Farquhar, G. D., von Caemmerer, S. V., & Berry, J. A. (1980). A biochemical model of photosynthetic CO 2 assimilation in leaves of C 3 species. planta, 149, 78-90.

Fuster, B., Sánchez-Zapero, J., Camacho, F., García-Santos, V., Verger, A., Lacaze, R., Weiss, M., Baret, F., & Smets, B. (2020). Quality Assessment of PROBA-V LAI, fAPAR and fCOVER Collection 300 m Products of Copernicus Global Land Service. Remote Sensing, 12

Galmes, J., Medrano, H., & Flexas, J. (2007). Photosynthetic limitations in response to water stress and recovery in Mediterranean plants with different growth forms. New Phytol, 175, 81-93

Groenendijk, M., Dolman, A.J., van der Molen, M.K., Leuning, R., Arneth, A., Delpierre, N., Gash, J.H.C., Lindroth, A., Richardson, A.D., Verbeeck, H., & Wohlfahrt, G. (2011). Assessing parameter variability in a photosynthesis model within and between plant functional types using global Fluxnet eddy covariance data. Agricultural and Forest Meteorology, 151, 22-38

Gu, L., Han, J., Wood, J.D., Chang, C.Y., & Sun, Y. (2019). Sun-induced Chl fluorescence and its importance for biophysical modeling of photosynthesis based on light reactions. New Phytol, 223, 1179-1191

Guanter, L., Alonso, L., Gómez-Chova, L., Amorós-López, J., Vila, J., & Moreno, J. (2007). Estimation of solar-induced vegetation fluorescence from space measurements. Geophysical Research Letters, 34

Guanter, L., Bacour, C., Schneider, A., Aben, I., van Kempen, T., Maignan, F., Retscher, C., Köhler, P., Frankenberg, C., Joiner, J., & Zhang, Y. (2021).

Han, J., Chang, C.Y., Gu, L., Zhang, Y., Meeker, E.W., Magney, T.S., Walker, A.P., Wen, J., Kira, O., McNaull, S., & Sun, Y. (2022a). The physiological basis for estimating photosynthesis from Chla fluorescence. New Phytol, 234, 1206-1219

Han, J., Gu, L., Wen, J., & Sun, Y. (2022b). Inference of photosynthetic capacity parameters from chlorophyll a fluorescence is affected by redox state of PSII reaction centers. Plant Cell Environ, 45, 1298-1314

He, L., Wood, J.D., Sun, Y., Magney, T., Dutta, D., Köhler, P., Zhang, Y., Yin, Y., & Frankenberg, C. (2020). Tracking Seasonal and Interannual Variability in Photosynthetic Downregulation in Response to Water Stress at a Temperate Deciduous Forest. Journal of Geophysical Research: Biogeosciences, 125

Joiner, J., Yoshida, Y., Köehler, P., Campbell, P., Frankenberg, C., van der Tol, C., Yang, P., Parazoo, N., Guanter, L., & Sun, Y. (2020). Systematic Orbital Geometry-Dependent Variations in Satellite Solar-Induced Fluorescence (SIF) Retrievals. Remote Sensing, 12

Kattge, J., DÍAz, S., Lavorel, S., Prentice, I.C., Leadley, P., BÖNisch, G., Garnier, E., Westoby, M., Reich, P.B., Wright, I.J., Cornelissen, J.H.C., Violle, C., Harrison, S.P., Van Bodegom, P.M., Reichstein, M., Enquist, B.J., Soudzilovskaia, N.A., Ackerly, D.D., Anand, M., Atkin, O., Bahn, M., Baker, T.R., Baldocchi, D., Bekker, R., Blanco, C.C., Blonder, B., Bond, W.J., Bradstock, R., Bunker, D.E., Casanoves, F., Cavender-Bares, J., Chambers, J.Q., Chapin Iii, F.S., Chave, J., Coomes, D., Cornwell, W.K., Craine, J.M., Dobrin, B.H., Duarte, L., Durka, W., Elser, J., Esser, G., Estiarte, M., Fagan, W.F., Fang, J., FernÁNdez-MÉNdez, F., Fidelis, A., Finegan, B., Flores, O., Ford, H., Frank, D., Freschet, G.T., Fyllas, N.M., Gallagher, R.V., Green, W.A., Gutierrez, A.G., Hickler, T., Higgins, S.I., Hodgson, J.G., Jalili, A., Jansen, S., Joly, C.A., Kerkhoff, A.J., Kirkup, D., Kitajima, K., Kleyer, M., Klotz, S., Knops, J.M.H., Kramer, K., KÜHn, I., Kurokawa, H., Laughlin, D., Lee, T.D., Leishman, M., Lens, F., Lenz, T., Lewis, S.L., Lloyd, J., LlusiÀ, J., Louault, F., Ma, S., Mahecha, M.D., Manning, P., Massad, T., Medlyn, B.E., Messier, J., Moles, A.T., MÜLler, S.C., Nadrowski, K., Naeem, S., Niinemets, Ü., NÖLlert, S., NÜSke, A., Ogaya, R., Oleksyn, J., Onipchenko, V.G., Onoda, Y., OrdoÑEz, J., Overbeck, G., Ozinga, W.A., PatiÑO, S., Paula, S., Pausas, J.G., PeÑUelas, J., Phillips, O.L., Pillar, V., Poorter, H., Poorter, L., Poschlod, P., Prinzing, A., Proulx, R., Rammig, A., Reinsch, S., Reu, B., Sack, L., Salgado-Negret, B., Sardans, J., Shiodera, S., Shipley, B., Siefert, A., Sosinski, E., Soussana, J.F., Swaine, E., Swenson, N., Thompson, K., Thornton, P., Waldram, M., Weiher, E., White, M., White, S., Wright, S.J., Yguel, B., Zaehle, S., Zanne, A.E., & Wirth, C. (2011). TRY – a global database of plant traits. Global Change Biology, 17, 2905-2935

Koehler, P., Frankenberg, C., Magney, T.S., Guanter, L., Joiner, J., & Landgraf, J. (2018). Global retrievals of solar induced chlorophyll fluorescence with TROPOMI: first results and inter-sensor comparison to OCO-2. Geophys Res Lett, 45, 10456-10463

Koffi, E.N., Rayner, P.J., Norton, A.J., Frankenberg, C., & Scholze, M. (2015). Investigating the usefulness of satellite-derived fluorescence data in inferring gross primary productivity within the carbon cycle data assimilation system. Biogeosciences, 12, 4067-4084

Kramer, D. M., Johnson, G., Kiirats, O., & Edwards, G. E. (2004). New fluorescence parameters for the determination of QA redox state and excitation energy fluxes. Photosynthesis research, 79, 209-218.

Lasslop, G., Reichstein, M., Papale, D., Richardson, A.D., Arneth, A., Barr, A., Stoy, P., & Wohlfahrt, G. (2009). Separation of net ecosystem exchange into assimilation

and respiration using a light response curve approach: critical issues and global evaluation. Global Change Biology, 16, 187-208

Leblanc, S.G., & Fournier, R.A. (2014). Hemispherical photography simulations with an architectural model to assess retrieval of leaf area index. Agricultural and Forest Meteorology, 194, 64-76

Lee, J.E., Frankenberg, C., van der Tol, C., Berry, J.A., Guanter, L., Boyce, C.K., Fisher, J.B., Morrow, E., Worden, J.R., Asefi, S., Badgley, G., & Saatchi, S. (2013). Forest productivity and water stress in Amazonia: observations from GOSAT chlorophyll fluorescence. Proc Biol Sci, 280, 20130171

Li, R., Lombardozzi, D., Shi, M., Frankenberg, C., Parazoo, N.C., Kohler, P., Yi, K., Guan, K., & Yang, X. (2022a). Representation of Leaf-to-Canopy Radiative Transfer Processes Improves Simulation of Far-Red Solar-Induced Chlorophyll Fluorescence in the Community Land Model Version 5. J Adv Model Earth Syst, 14, e2021MS002747

Li, X., Shabanov, N.V., Chen, L., Zhang, Y., & Huang, H. (2022b). Modeling solar-induced fluorescence of forest with heterogeneous distribution of damaged foliage by extending the stochastic radiative transfer theory. Remote Sensing of Environment, 271

Liu, J., Bowman, K.W., Schimel, D.S., Parazoo, N.C., Jiang, Z., Lee, M., Bloom, A.A., Wunch, D., Frankenberg, C., Sun, Y., O'Dell, C.W., Gurney, K.R., Menemenlis, D., Gierach, M., Crisp, D., & Eldering, A. (2017). Contrasting carbon cycle responses of the tropical continents to the 2015-2016 El Nino. Science, 358

Liu, Z., Zhao, F., Liu, X., Yu, Q., Wang, Y., Peng, X., Cai, H., & Lu, X. (2022). Direct estimation of photosynthetic CO2 assimilation from solar-induced chlorophyll fluorescence (SIF). Remote Sensing of Environment, 271

Lombardozzi, D.L., Wieder, W.R., Sobhani, N., Bonan, G.B., Durden, D., Lenz, D., SanClements, M., Weintraub-Leff, S., Ayres, E., Florian, C.R., Dahlin, K., Kumar, S., Swann, A.L.S., Zarakas, C., Vardeman, C., & Pascucci, V.

Magney, T.S., Bowling, D.R., Logan, B.A., Grossmann, K., Stutz, J., Blanken, P.D., Burns, S.P., Cheng, R., Garcia, M.A., Köhler, P., Lopez, S., Parazoo, N.C., Raczka, B., Schimel, D., & Frankenberg, C. (2019). Mechanistic evidence for tracking the seasonality of photosynthesis with solar-induced fluorescence. Proc Natl Acad Sci U S A, 116, 11640-11645

Miller, J. B. (1967). A formula for average foliage density. Australian Journal of Botany, 15(1), 141-144.

Mohammed, G.H., Colombo, R., Middleton, E.M., Rascher, U., van der Tol, C., Nedbal, L., Goulas, Y., Perez-Priego, O., Damm, A., Meroni, M., Joiner, J., Cogliati, S., Verhoef, W., Malenovsky, Z., Gastellu-Etchegorry, J.P., Miller, J.R., Guanter, L.,

Moreno, J., Moya, I., Berry, J.A., Frankenberg, C., & Zarco-Tejada, P.J. (2019). Remote sensing of solar-induced chlorophyll fluorescence (SIF) in vegetation: 50 years of progress. Remote Sens Environ, 231

Muñoz Sabater, J. (2019): ERA5-Land monthly averaged data from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.68d2bb30 (Accessed on DD-MMM-YYYY)

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. Journal of hydrology, 10(3), 282-290.

NEON (National Ecological Observatory Network)1. Plant foliar traits (DP1.10026.001), RELEASE-2023. https://doi.org/10.48443/tmrs-fb32. Dataset accessed from https://data.neonscience.org/data-products/DP1.10026.001/RELEASE-2023 on December 25, 2023.

NEON (National Ecological Observatory Network)2. Plant presence and percent cover (DP1.10058.001), RELEASE-2023. https://doi.org/10.48443/9579-a253. Dataset accessed from https://data.neonscience.org/data-products/DP1.10058.001/RELEASE-2023 on December 25, 2023.

Pacheco-Labrador, J., Perez-Priego, O., El-Madany, T.S., Julitta, T., Rossini, M., Guan, J., Moreno, G., Carvalhais, N., Martín, M.P., Gonzalez-Cascon, R., Kolle, O., Reischtein, M., van der Tol, C., Carrara, A., Martini, D., Hammer, T.W., Moossen, H., & Migliavacca, M. (2019). Multiple-constraint inversion of SCOPE. Evaluating the potential of GPP and SIF for the retrieval of plant functional traits. Remote Sensing of Environment, 234

Parazoo, N.C., Magney, T., Norton, A., Raczka, B., Bacour, C., Maignan, F., Baker, I., Zhang, Y., Qiu, B., Shi, M., MacBean, N., Bowling, D.R., Burns, S.P., Blanken, P.D., Stutz, J., Grossmann, K., & Frankenberg, C. (2020). Wide discrepancies in the magnitude and direction of modeled solar-induced chlorophyll fluorescence in response to light conditions. Biogeosciences, 17, 3733-3755

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., Isaac, P., Polidori, D., Reichstein, M., Ribeca, A., van Ingen, C., Vuichard, N., Zhang, L., Amiro, B., Ammann, C., Arain, M.A., Ardo, J., Arkebauer, T., Arndt, S.K., Arriga, N., Aubinet, M., Aurela, M., Baldocchi, D., Barr, A., Beamesderfer, E., Marchesini, L.B., Bergeron, O., Beringer, J., Bernhofer, C., Berveiller, D., Billesbach, D., Black, T.A., Blanken, P.D., Bohrer, G., Boike, J., Bolstad, P.V., Bonal, D., Bonnefond, J.M., Bowling, D.R., Bracho, R., Brodeur, J., Brummer, C., Buchmann, N., Burban, B., Burns, S.P., Buysse, P., Cale, P., Cavagna, M., Cellier, P., Chen, S., Chini, I., Christensen, T.R., Cleverly, J., Collalti, A., Consalvo, C., Cook, B.D., Cook, D., Coursolle, C., Cremonese, E., Curtis, P.S., D'Andrea, E., da Rocha, H., Dai, X., Davis, K.J., Cinti, B., Grandcourt, A., Ligne, A., De Oliveira, R.C., Delpierre, N., Desai,

A.R., Di Bella, C.M., Tommasi, P.D., Dolman, H., Domingo, F., Dong, G., Dore, S., Duce, P., Dufrene, E., Dunn, A., Dusek, J., Eamus, D., Eichelmann, U., ElKhidir, H.A.M., Eugster, W., Ewenz, C.M., Ewers, B., Famulari, D., Fares, S., Feigenwinter, I., Feitz, A., Fensholt, R., Filippa, G., Fischer, M., Frank, J., Galvagno, M., Gharun, M., Gianelle, D., Gielen, B., Gioli, B., Gitelson, A., Goded, I., Goeckede, M., Goldstein, A.H., Gough, C.M., Goulden, M.L., Graf, A., Griebel, A., Gruening, C., Grunwald, T., Hammerle, A., Han, S., Han, X., Hansen, B.U., Hanson, C., Hatakka, J., He, Y., Hehn, M., Heinesch, B., Hinko-Najera, N., Hortnagl, L., Hutley, L., Ibrom, A., Ikawa, H., Jackowicz-Korczynski, M., Janous, D., Jans, W., Jassal, R., Jiang, S., Kato, T., Khomik, M., Klatt, J., Knohl, A., Knox, S., Kobayashi, H., Koerber, G., Kolle, O., Kosugi, Y., Kotani, A., Kowalski, A., Kruijt, B., Kurbatova, J., Kutsch, W.L., Kwon, H., Launiainen, S., Laurila, T., Law, B., Leuning, R., Li, Y., Liddell, M., Limousin, J.M., Lion, M., Liska, A.J., Lohila, A., Lopez-Ballesteros, A., Lopez-Blanco, E., Loubet, B., Loustau, D., Lucas-Moffat, A., Luers, J., Ma, S., Macfarlane, C., Magliulo, V., Maier, R., Mammarella, I., Manca, G., Marcolla, B., Margolis, H.A., Marras, S., Massman, W., Mastepanov, M., Matamala, R., Matthes, J.H., Mazzenga, F., McCaughey, H., McHugh, I., McMillan, A.M.S., Merbold, L., Meyer, W., Meyers, T., Miller, S.D., Minerbi, S., Moderow, U., Monson, R.K., Montagnani, L., Moore, C.E., Moors, E., Moreaux, V., Moureaux, C., Munger, J.W., Nakai, T., Neirynck, J., Nesic, Z., Nicolini, G., Noormets, A., Northwood, M., Nosetto, M., Nouvellon, Y., Novick, K., Oechel, W., Olesen, J.E., Ourcival, J.M., Papuga, S.A., Parmentier, F.J., Paul-Limoges, E., Pavelka, M., Peichl, M., Pendall, E., Phillips, R.P., Pilegaard, K., Pirk, N., Posse, G., Powell, T., Prasse, H., Prober, S.M., Rambal, S., Rannik, U., Raz-Yaseef, N., Rebmann, C., Reed, D., Dios, V.R., Restrepo-Coupe, N., Reverter, B.R., Roland, M., Sabbatini, S., Sachs, T., Saleska, S.R., Sanchez-Canete, E.P., Sanchez-Mejia, Z.M., Schmid, H.P., Schmidt, M., Schneider, K., Schrader, F., Schroder, I., Scott, R.L., Sedlak, P., Serrano-Ortiz, P., Shao, C., Shi, P., Shironya, I., Siebicke, L., Sigut, L., Silberstein, R., Sirca, C., Spano, D., Steinbrecher, R., Stevens, R.M., Sturtevant, C., Suyker, A., Tagesson, T., Takanashi, S., Tang, Y., Tapper, N., Thom, J., Tomassucci, M., Tuovinen, J.P., Urbanski, S., Valentini, R., van der Molen, M., van Gorsel, E., van Huissteden, K., Varlagin, A., Verfaillie, J., Vesala, T., Vincke, C., Vitale, D., Vygodskaya, N., Walker, J.P., Walter-Shea, E., Wang, H., Weber, R., Westermann, S., Wille, C., Wofsy, S., Wohlfahrt, G., Wolf, S., Woodgate, W., Li, Y., Zampedri, R., Zhang, J., Zhou, G., Zona, D., Agarwal, D., Biraud, S., Torn, M., & Papale, D. (2020). The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. Sci Data, 7, 225

Paul-Limoges, E., Damm, A., Hueni, A., Liebisch, F., Eugster, W., Schaepman, M.E., & Buchmann, N. (2018). Effect of environmental conditions on sun-induced fluorescence in a mixed forest and a cropland. Remote Sensing of Environment, 219, 310-323

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and earth system sciences*, *11*(5), 1633-1644.

Pfündel, E. (1998). Estimating the contribution of photosystem I to total leaf chlorophyll fluorescence. *Photosynthesis research*, *56*, 185-195.

Porcar-Castell, A., Tyystjarvi, E., Atherton, J., van der Tol, C., Flexas, J., Pfundel, E.E., Moreno, J., Frankenberg, C., & Berry, J.A. (2014). Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges. J Exp Bot, 65, 4065-4095

Prikaziuk, E., Migliavacca, M., Su, Z., & van der Tol, C. (2023). Simulation of ecosystem fluxes with the SCOPE model: Sensitivity to parametrization and evaluation with flux tower observations. Remote Sensing of Environment, 284

Raczka, B., Porcar-Castell, A., Magney, T., Lee, J.E., Köhler, P., Frankenberg, C., Grossmann, K., Logan, B.A., Stutz, J., Blanken, P.D., Burns, S.P., Duarte, H., Yang, X., Lin, J.C., & Bowling, D.R. (2019). Sustained Nonphotochemical Quenching Shapes the Seasonal Pattern of Solar-Induced Fluorescence at a High-Elevation Evergreen Forest. Journal of Geophysical Research: Biogeosciences, 124, 2005-2020

Reichstein, M., Tenhunen, J., Roupsard, O., Ourcival, J.M., Rambal, S., Miglietta, F., Peressotti, A., Pecchiari, M., Tirone, G., & Valentini, R. (2003). Inverse modeling of seasonal drought effects on canopy CO2/H2O exchange in three Mediterranean ecosystems. Journal of Geophysical Research: Atmospheres, 108

Rogers, A. (2014). The use and misuse of V(c,max) in Earth System Models. Photosynth Res, 119, 15-29

Ryu, Y., Berry, J.A., & Baldocchi, D.D. (2019). What is global photosynthesis? History, uncertainties and opportunities. Remote Sensing of Environment, 223, 95-114

Shi, S., Cong, W., Lu, S., Zhao, T., Wang, F., & Lu, Q. (2022). Can SIF and NPQ be used in the photosynthesis rate simulation of plants subjected to drought? Environmental and Experimental Botany, 203

Song, Y., Wang, L., & Wang, J. (2021). Improved understanding of the spatially-heterogeneous relationship between satellite solar-induced chlorophyll fluorescence and ecosystem productivity. Ecological Indicators, 129

Sun, Y., Gu, L., Wen, J., van der Tol, C., Porcar-Castell, A., Joiner, J., Chang, C.Y., Magney, T., Wang, L., Hu, L., Rascher, U., Zarco-Tejada, P., Barrett, C.B., Lai, J., Han, J., & Luo, Z. (2023a). From remotely sensed solar-induced chlorophyll fluorescence to ecosystem structure, function, and service: Part I-Harnessing theory. Glob Chang Biol, 29, 2926-2952

Sun, Y., Wen, J., Gu, L., Joiner, J., Chang, C.Y., van der Tol, C., Porcar-Castell, A., Magney, T., Wang, L., Hu, L., Rascher, U., Zarco-Tejada, P., Barrett, C.B., Lai, J., Han, J., & Luo, Z. (2023b). From remotely-sensed solar-induced chlorophyll

fluorescence to ecosystem structure, function, and service: Part II-Harnessing data. Glob Chang Biol, 29, 2893-2925

Thum, T., Zaehle, S., Köhler, P., Aalto, T., Aurela, M., Guanter, L., Kolari, P., Laurila, T., Lohila, A., Magnani, F., Van Der Tol, C., & Markkanen, T. (2017). Modelling sun-induced fluorescence and photosynthesis with a land surface model at local and regional scales in northern Europe. Biogeosciences, 14, 1969-1987

Van der Tol, C., Verhoef, W., Timmermans, J., Verhoef, A., & Su, Z. (2009). An integrated model of soil-canopy spectral radiances, photosynthesis, fluorescence, temperature and energy balance. *Biogeosciences*, *6*(12), 3109-3129.

van der Tol, C., Berry, J.A., Campbell, P.K., & Rascher, U. (2014). Models of fluorescence and photosynthesis for interpreting measurements of solar-induced chlorophyll fluorescence. J Geophys Res Biogeosci, 119, 2312-2327

Veefkind, J.P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H.J., de Haan, J.F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., & Levelt, P.F. (2012). TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. Remote Sensing of Environment, 120, 70-83

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

Vilfan, N., van der Tol, C., Muller, O., Rascher, U., & Verhoef, W. (2016). Fluspect-B: A model for leaf fluorescence, reflectance and transmittance spectra. Remote Sensing of Environment, 186, 596-615

Vuichard, N., & Papale, D. (2015). Filling the gaps in meteorological continuous data measured at FLUXNET sites with ERA-Interim reanalysis. Earth System Science Data, 7, 157-171

Wang, R., Chen, J.M., Luo, X., Black, A., & Arain, A. (2019). Seasonality of leaf area index and photosynthetic capacity for better estimation of carbon and water fluxes in evergreen conifer forests. Agricultural and Forest Meteorology, 279

Wang, Y., Zhou, L., Zhuang, J., Sun, L., & Chi, Y. (2021). The spatial heterogeneity of the relationship between gross primary production and sun-induced chlorophyll fluorescence regulated by climate conditions during 2007–2018. Global Ecology and Conservation, 29

Wang, Z., Townsend, P.A., & Kruger, E.L. (2022). Leaf spectroscopy reveals divergent inter- and intra-species foliar trait covariation and trait-environment relationships across NEON domains. New Phytol, 235, 923-938

Wen, J., Köhler, P., Duveiller, G., Parazoo, N.C., Magney, T.S., Hooker, G., Yu, L., Chang, C.Y., & Sun, Y. (2020). A framework for harmonizing multiple satellite instruments to generate a long-term global high spatial-resolution solar-induced chlorophyll fluorescence (SIF). Remote Sensing of Environment, 239

Wilson, J. W. (1963). Estimation of foliage denseness and foliage angle by inclined point quadrats. *Australian journal of botany*, *11*(1), 95-105.

Wu, G., Jiang, C., Kimm, H., Wang, S., Bernacchi, C., Moore, C.E., Suyker, A., Yang, X., Magney, T., Frankenberg, C., Ryu, Y., Dechant, B., & Guan, K. (2022). Difference in seasonal peak timing of soybean far-red SIF and GPP explained by canopy structure and chlorophyll content. Remote Sensing of Environment, 279

Yang, P., Prikaziuk, E., Verhoef, W., & van der Tol, C. (2021a). SCOPE 2.0: a model to simulate vegetated land surface fluxes and satellite signals. Geoscientific Model Development, 14, 4697-4712

Yang, P., van der Tol, C., Campbell, P.K.E., & Middleton, E.M. (2021b). Unraveling the physical and physiological basis for the solar- induced chlorophyll fluorescence and photosynthesis relationship using continuous leaf and canopy measurements of a corn crop. Biogeosciences, 18, 441-465

Yu, L., Wen, J., Chang, C.Y., Frankenberg, C., & Sun, Y. (2019). High-Resolution Global Contiguous SIF of OCO-2. Geophysical Research Letters, 46, 1449-1458

Zeng, Y., Badgley, G., Chen, M., Li, J., Anderegg, L.D.L., Kornfeld, A., Liu, Q., Xu, B., Yang, B., Yan, K., & Berry, J.A. (2020). A radiative transfer model for solar induced fluorescence using spectral invariants theory. Remote Sensing of Environment, 240

Zhang, Y., Fang, J., Smith, W.K., Wang, X., Gentine, P., Scott, R.L., Migliavacca, M., Jeong, S., Litvak, M., & Zhou, S. (2023). Satellite solar-induced chlorophyll fluorescence tracks physiological drought stress development during 2020 southwest US drought. Glob Chang Biol, 29, 3395-3408

Zhao, F., Dai, X., Verhoef, W., Guo, Y., van der Tol, C., Li, Y., & Huang, Y. (2016). FluorWPS: A Monte Carlo ray-tracing model to compute sun-induced chlorophyll fluorescence of three-dimensional canopy. Remote Sensing of Environment, 187, 385-399

APPENDIX

TEXT A1 SCOPE gross photosynthesis

SCOPE output An represents the net leaf assimilated carbon, with leaf day respiration rate (Rd) determined as a fraction of $Vcmax_{25}$: $Rd = Rd_{param} \cdot Vcmax_{25}$. To make SCOPE output An represent gross photosynthesis (Ag), we set Rd_{param} to 0 throughout all simulations to disable leaf respiration.

$$An = Ag - Rd \quad (A1)$$

$$Rd = Rd_{param} \cdot Vcmax_{25} \quad (A2)$$

TEXT A2 Performance evaluation

We conducted an evaluation of the accuracy of simulated GPP and SIF data by comparing them temporally with tower and satellite observations, respectively. The comparison of GPP was conducted on a daily basis, whereas to match the OCO-2 16-day SIF products, all SIF simulations were aggregated to a 16-day scale.

To assess the model performance, we employed a suite of metrics, including Nash-Sutcliffe modeling efficiency (MEF; Nash & Sutcliffe, 1970), root-mean-square error (RMSE), fractional bias (FB), and slope of the linear regression trend. The MEF metric ranges from -Inf to 1 and evaluates the performance of the model in comparison to predicting observation mean. The RMSE metric measures the average magnitude of the prediction residuals. The FB metric measures the relative difference between observation mean and prediction mean. Slope measures how well the simulations are correlated with the observations. The closer the values of MEF and slope are to 1 and the closer the values of RMSE and Bias are to 0, the greater the consistency between the model simulations and the observations.

Table A1 Percentage of each land cover (LC) class within a 0.05° × 0.05° grid of each site. 30 m LC data from the 2019 National Land Cover Database (NLCD), regridded to 300 m based on the dominant LC type in each grid. The LC used in LC-specific simulation (Spatial heterogeneity effect for BART and JERC, see **Sect.3.3**) are shown in bold. For the site BART, the mixed forest (45 %) LC was further equally allocated to deciduous forest and evergreen forest.

Site	Deciduous forest (%)	Evergreen forest (%)	Mixed forest (%)	Cultivated crops (%)	Wetlands (%)	Urban (%)	Others (%)
30 m:							
BART	27	20	38	1	3	8	3
STEI	33	4	21	4	22	4	12
JERC	2	44	4	34	4	4	8
300 m:							
BART	**34**	**13**	**45**	-	3	4	1
STEI	41	3	18	3	24	1	10
JERC	-	**56**	1	**38**	1	-	1

Table A2 The parameter regimes used in the qL-based SIF parameterization (see **Table A3** reported in Han et al., 2022a for exact values), following the Köppen climate classes (Peel et al., 2007). For brevity, BDT refers to Broadleaf deciduous

tree; NET refers to Needleleaf evergreen tree; C3 Crop refers to crop species with C3 metabolism.

Simulation regime	Köppen climate class	Parameter regime
STEI (baseline)	Boreal	BDT Boreal
BART:	Boreal	
Baseline		BDT Boreal
BART-deciduous		BDT Boreal
BART-evergreen		NET Boreal
JERC:	Temperate	
Baseline		BDT Temperate
JERC-evergreen		BDT Temperate
JERC-crop		C3 Crop

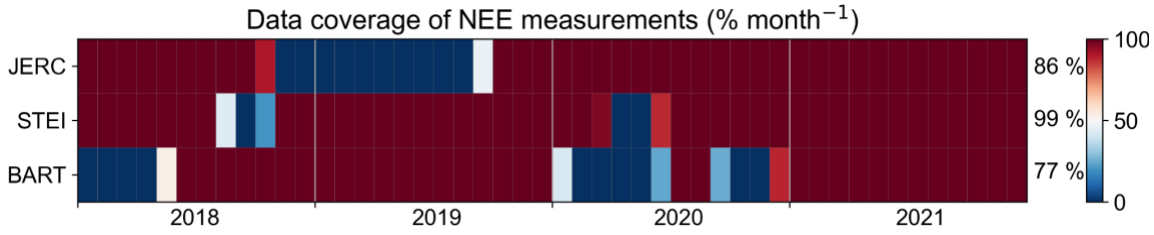

Fig. A1. Data coverage of net ecosystem exchange (NEE) by site and by month during 2018-2021. Average data coverages for all the study period are inset.

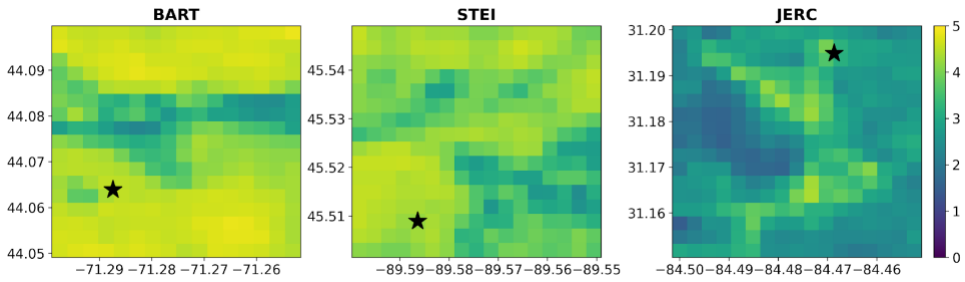

Fig. A2. The growing season mean LAI of our study domains. The flux towers are marked with black stars. LAI data comes from two final estimates (RT5 and RT6, consolidated 50 and 60 days of the Real Time (RT0) products) of CGLS 300 m V1, (Fuster et al., 2021). The growing season was defined as the green days (from averaged green increase day to the averaged green minimum day in a year) from NEON metadata.

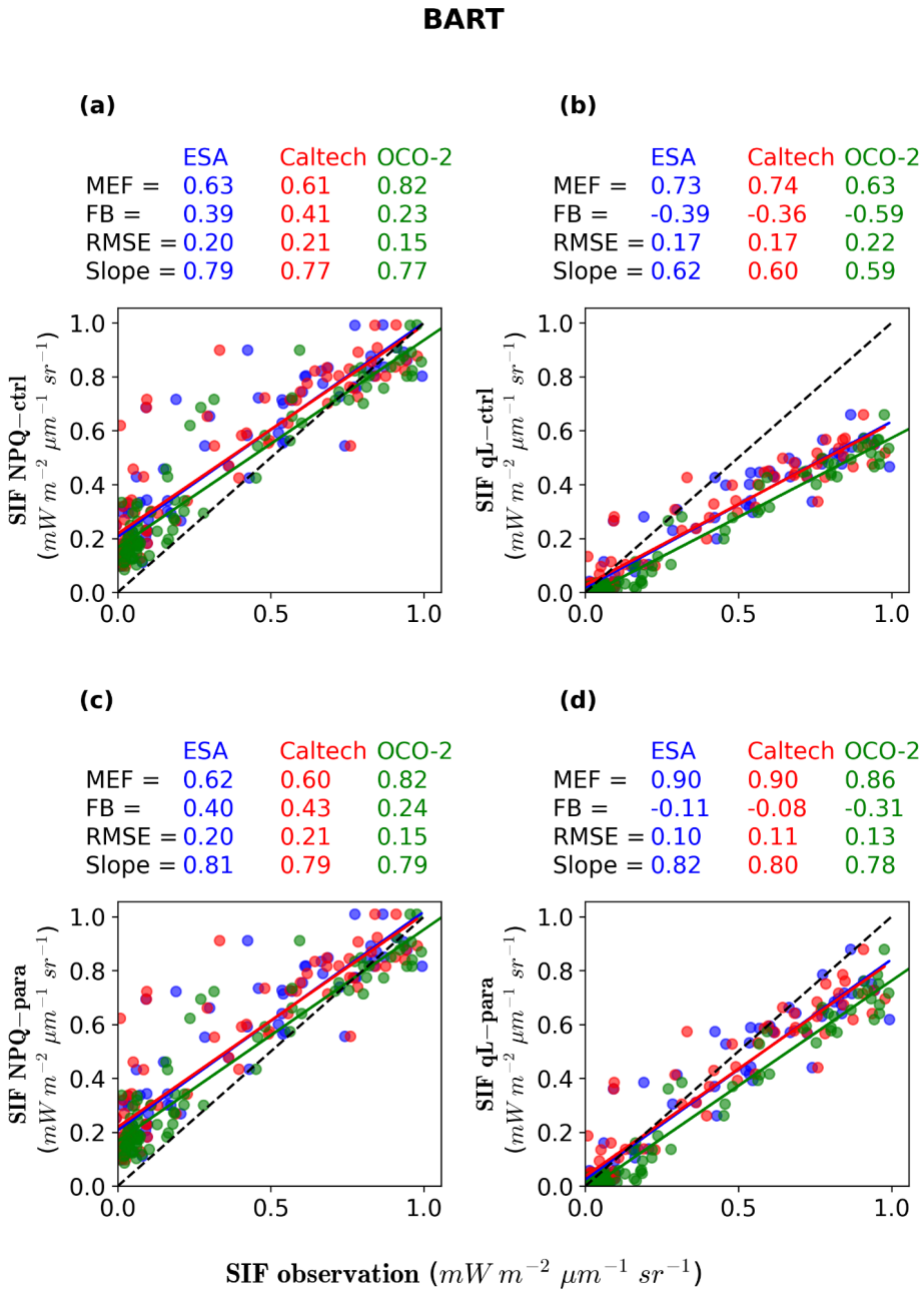

Fig. A3. Comparison of (a) NPQ-ctrl, (b) qL-ctrl, (c) NPQ-para, and (d) qL-para simulations of 16-day mean SIF against satellite observations (TROPOMI ESA, TROPOMI Caltech, and OCO-2) in BART during 2018–2021. The black dashed line represents a 1:1 relationship, and the blue, red, and green lines represent the regression

relationship of simulations with TROPOMI ESA, TROPOMI Caltech, and OCO-2 observations, respectively. The values of MEF, FB, RMSE ($W\ m^{-2}\ \mu m^{-1}\ sr^{-1}$), and the slope values are inset.

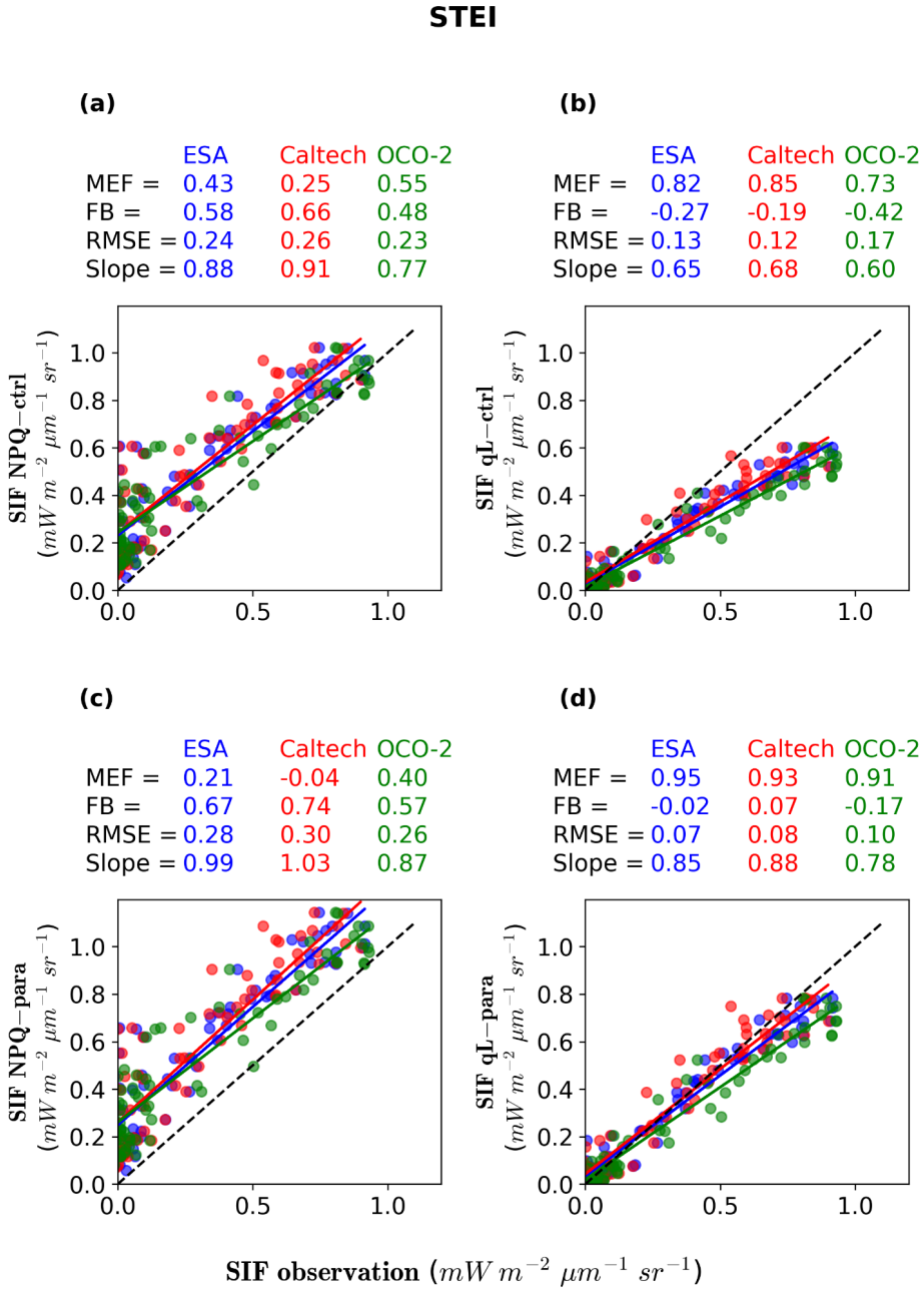

Fig. A4. Same as **Fig. A3**, but for STEI.

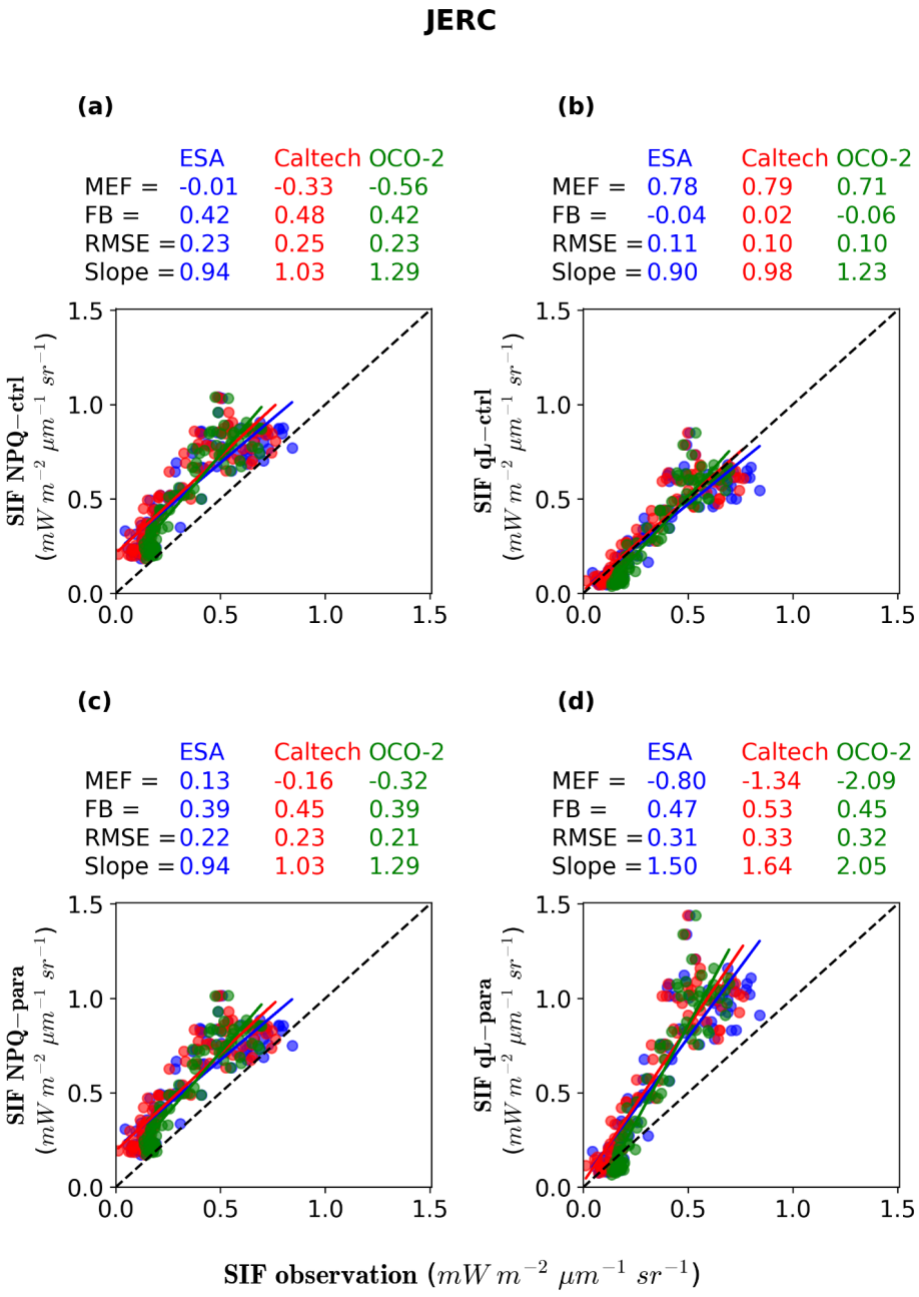

Fig. A5. Same as **Fig. A3**, but for JERC.

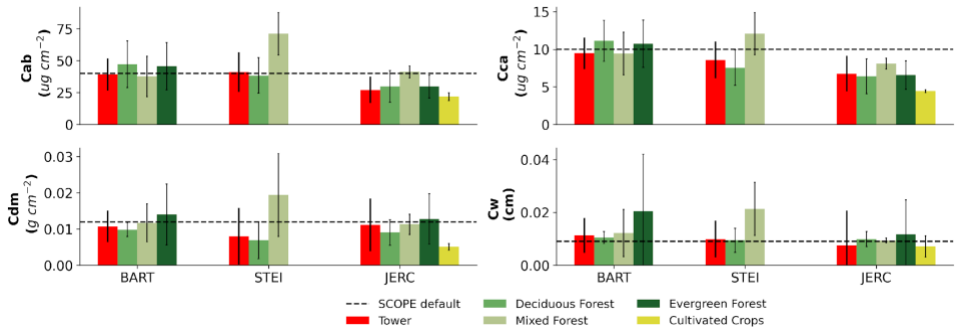

Fig. A6. Comparison of median photosynthetic parameters from NEON and SCOPE default (black dashed line). The error bar represents the standard deviations of all sample values. The "tower" represents the location of sampling that are placed in the 90 % flux area of the primary and secondary airsheds of each NEON tower (https://www.neonscience.org/sites/default/files/All_NEON_TOS_Plots_V9_0.zip).

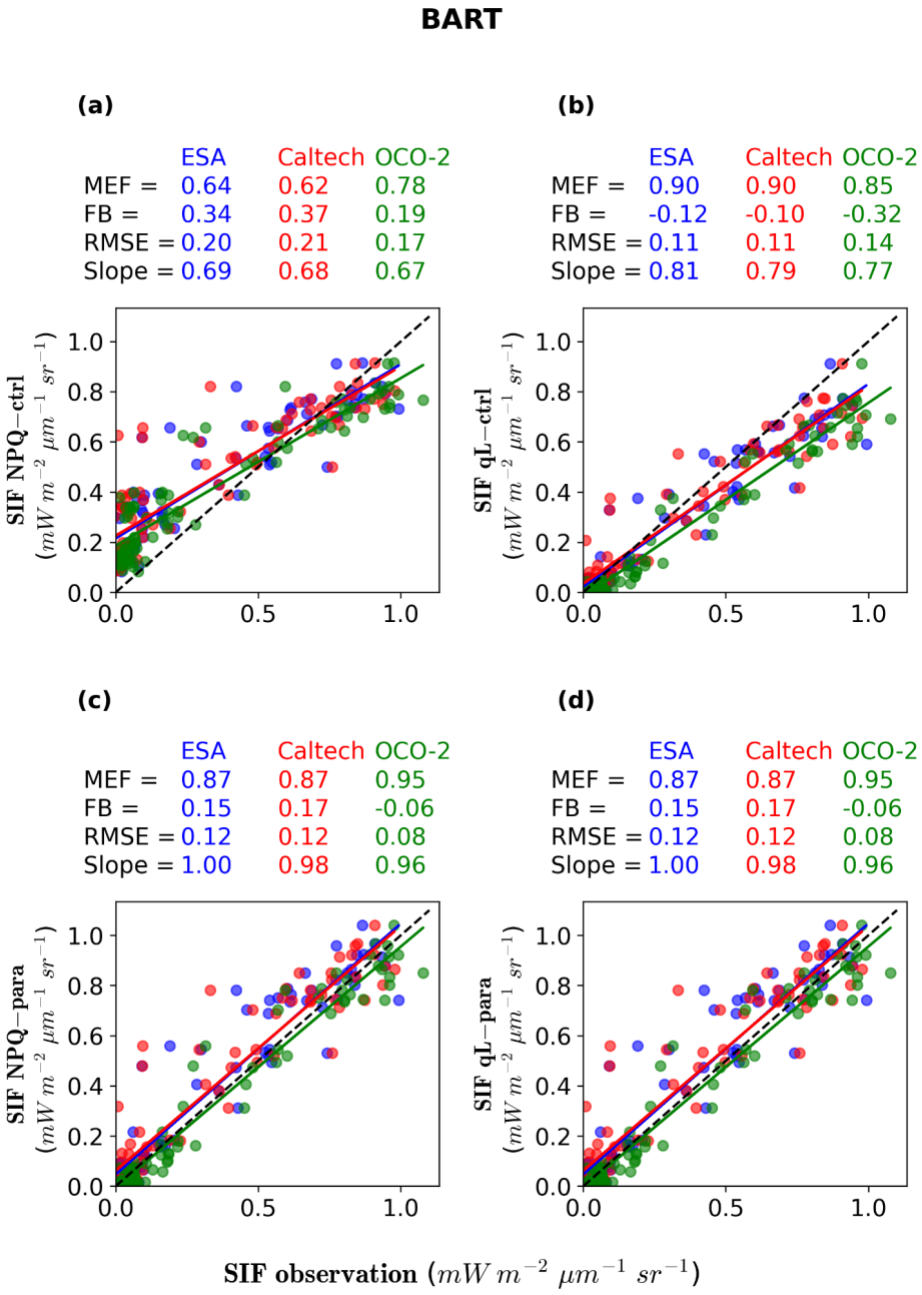

Fig. A7. Same as **Fig. A3**, but for using the LC-specific LAI and foliar trait values in weighted average simulations (weighting values see **Table A1**).

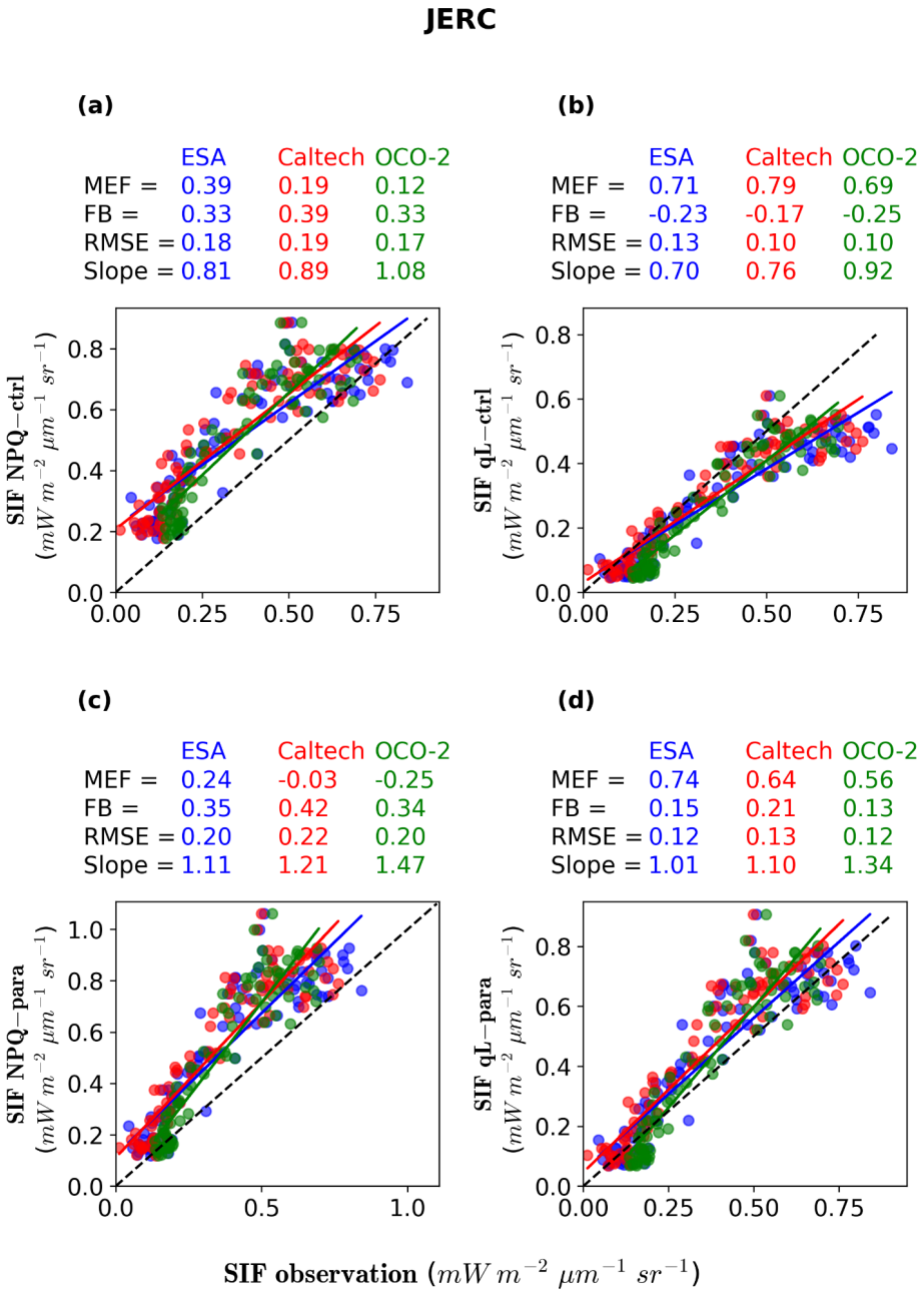

Fig. A8. Same as **Fig. A7**, but for JERC.

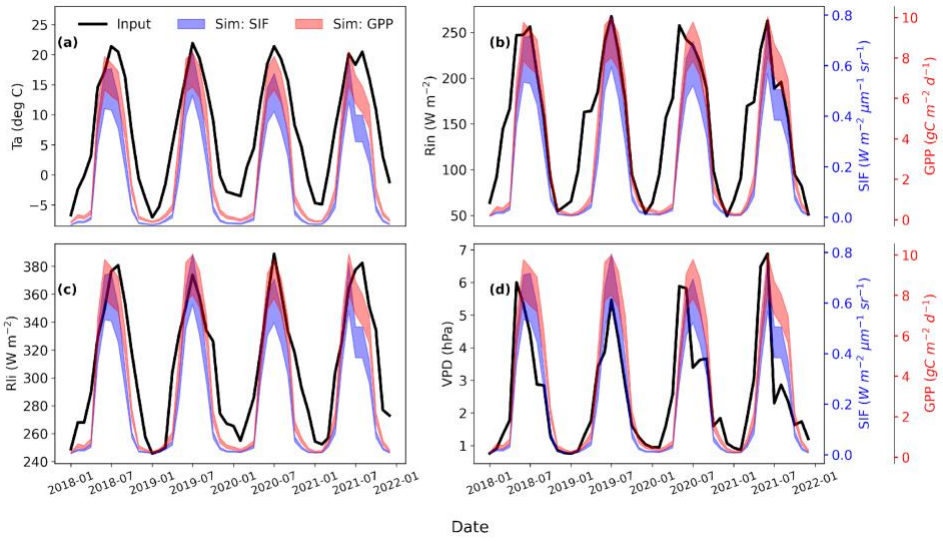

Fig. A9. Monthly time series of GPP and SIF from model simulations and model input meteorological variables in BART from 2018 to 2021: (a) temperature (Ta); (b) shortwave radiation (Rin); (c) longwave radiation (Rli); and (d) vapor pressure deficit (VPD). The shaded regions represent the upper and lower bounds of all simulations.

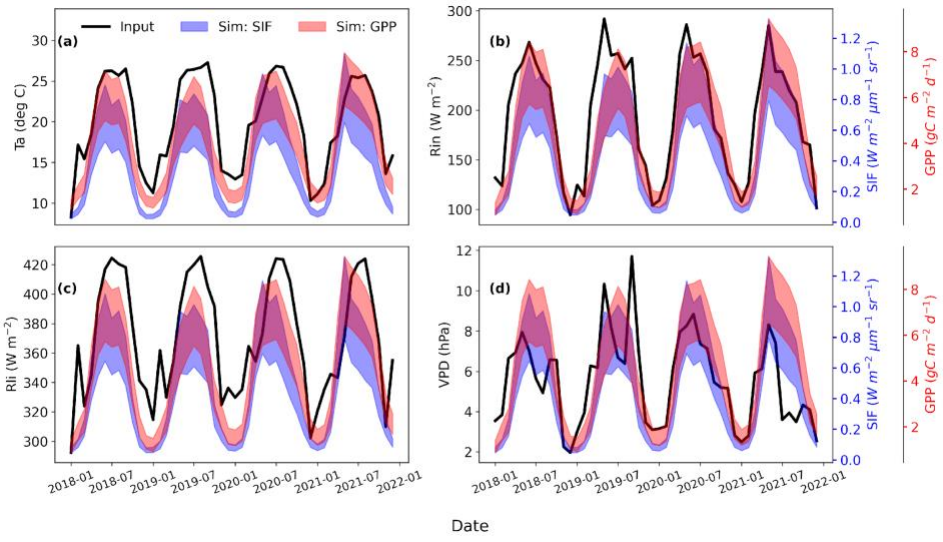

Fig. A10. Same as **Fig. A9**, but for JERC