Two for one: Partitioning CO_2 fluxes and understanding the relationship between solar-induced chlorophyll fluorescence and gross primary productivity using machine learning

Weiwei Zhan [a], Xi Yang [b], Youngryel Ryu [c], Benjamin Dechant [c], Yu Huang [a], Yves Goulas [d], Minseok Kang [e], Pierre Gentine [a],[*]

[a] Department of Earth and Environmental Engineering, Columbia University, New York, NY, United States
[b] Department of Environmental Sciences, University of Virginia, Charlottesville, VA, United States
[c] Department of Landscape Architecture and Rural Systems Engineering, Seoul National University, Republic of Korea
[d] Dynamic Meteorology Laboratory, Ecole Polytechnique, Palaiseau, France
[e] National Center for AgroMeteorology, Seoul, Republic of Korea

ARTICLE INFO

ABSTRACT

Accurately partitioning net ecosystem exchange (NEE) into ecosystem respiration (ER) and gross primary productivity (GPP) is critical for understanding the terrestrial carbon cycle. The standard partitioning methods rely on simplified empirical models, which have inherent structural errors. These structural errors lead to biased GPP and ER estimation, especially during extreme events (e.g., drought) and human disturbances (e.g., crop harvest). Recently, solar-induced chlorophyll fluorescence (SIF) has been shown to be well correlated to GPP, thus offering a path to improve the NEE partitioning by constraining GPP. However, the ecosystem-scale relationship between GPP and SIF remains limited. Here, we show that neural networks informed by SIF observations (NN_{SIF}) can be successfully used to partition NEE, while simultaneously learning the ecosystem-scale GPP-SIF relationship. NN_{SIF} was compared against standard partitioning methods and NN without SIF constraint (NN_{noSIF}), using field data from different ecosystems and synthetic data generated by a coupled fluorescence-photosynthesis model (SCOPE). NN_{SIF} showed superior performance as: (1) it effectively improves the ER estimation, especially at high temperature, (2) it better captures the moisture limitation on ER, (3) it more accurately estimates LUE variations to stress, and (4) it uniquely captures the rapid GPP drop after land management (harvest). Furthermore, NN_{SIF} can retrieve the GPP-SIF relationship at the ecosystem scale, and elucidate how this relationship responds to environmental conditions. Overall, our algorithm provides the first direct and non-empirical estimate of the ecosystem-scale GPP-SIF relationship, without relying on any prior empirical assumptions on the relationships between CO_2 fluxes, climatic drivers, and SIF. The new knowledge learned by NN_{SIF} can help better estimate global-scale GPP using satellite SIF, especially during extreme events and in the presence of land management.

1. Introduction

Terrestrial ecosystems approximately absorb one-third of anthropogenic CO_2 emissions, as gross primary productivity (GPP) outpaces ecosystem respiration (ER) and other carbon losses (e.g. disturbances) (Keenan et al., 2019; McNicol et al., 2018). Accurate estimation of GPP and ER is thus important for future climate projections (Schimel et al., 2015). However, currently, it is impossible to directly observe daytime GPP or ER at the ecosystem scale (Desai et al., 2008; Keenan et al., 2019). Instead, GPP and ER are typically inferred from measurements of net ecosystem exchange (NEE), the net CO_2 exchange between an ecosystem and the atmosphere, defined as NEE = ER – GPP. Accurately partitioning NEE into ER and GPP is critical for understanding the responses of the terrestrial carbon cycle to a changing climate, especially during extreme events (Zscheischler et al., 2014).

Various NEE partitioning methods have been proposed in recent years (Desai et al., 2008; Keenan et al., 2019; Lasslop et al., 2012, 2010; Reichstein et al., 2005; Spielmann et al., 2019; Stoy et al., 2006; Tramontana et al., 2020). Among all, two methods are widely applied in the EC community. The nighttime (NT) method uses night data to

parameterize a temperature-driven ER model, by fitting its temperature sensitivity (E_0) and reference respiration (R_{ref}). Then it extrapolates the fitted model to daytime temperatures to retrieve GPP as the residual, assuming that the model remains the same during daytime (Reichstein et al., 2005). The daytime (DT) method fits a light- and temperature-driven model, using both day and night data, to estimate GPP and ER (Lasslop et al., 2010). Both NT and DT methods rely on simplified empirical models which, inherently, have structural errors. One major structural error originates from the simplified ER parameterization. For instance, both methods assume that (1) the fitted ER parameters, E_0 and R_{ref}, remain the same through day and night, and (2) ER originates from a single pool driven by either air or soil temperature. However, these simplified assumptions have been found to generate biased ER and GPP estimates (Keenan et al., 2019; Wohlfahrt and Galvagno, 2017). Furthermore, these empirical models could collapse during extreme events (e.g., drought) and human disturbances (e.g., harvest), which are critical for the global carbon cycle (Zscheischler et al., 2014). Such collapse is partly due to the limited drivers used in these empirical models. To account for other unselected drivers (e.g., soil moisture, microbial activity, biomass), standard methods adjust the parameters of empirical models using a moving window approach. Even so, they still fail to capture the responses to fast environmental changes (e.g., land management or fast atmospheric forcing) (Riederer et al., 2014; Tramontana et al., 2020).

To tackle the limitations of standard partitioning methods, some studies have developed machine learning (ML)-based methods, which do not rely on empirical models but infer the relationships purely from data. Oikawa et al. (2017) trained neural networks (NN) to estimate nighttime ER, and then extrapolated the fitted NN to daytime to partition NEE. However, similar to NT, this method assumes that day and night ER are regulated by the same mechanisms. Instead of solely using night data, Tramontana et al. (2020) used day and night data to develop a new framework, in which NN was constrained by the observed NEE to simultaneously estimate GPP and ER. This approach outperforms standard methods for capturing additional ecological patterns, e.g., hysteresis in ER and GPP diurnal cycles. However, the observed NEE is the only constraint in their framework, and thus NEE uncertainties could have large effects on the GPP and ER estimates (Tramontana et al., 2020). Therefore, additional constraints can be imposed to reduce these uncertainties (Spielmann et al., 2019) and improve the partitioning estimates.

Chlorophyll fluorescence (ChlF), an optical signal emitted from the excited chlorophyll, can be a promising constraint on plant photosynthesis (Porcar-Castell et al., 2014). Given that ChlF competes for the same excitation energy as the light reactions of photosynthesis, ChlF can be an informative indicator of the photosynthetic mechanism (Baker, 2008). Pulse-amplitude modulation (PAM) can actively measure ChlF using controlled excitation light, but it is limited to leaf-level measurements (Magney et al., 2017; Mohammed et al., 2019). Alternatively, passive measurements of solar-induced chlorophyll fluorescence (SIF) can monitor plant photosynthesis across canopy, regional, and global scales without using artificial light (Frankenberg et al., 2011; Joiner et al., 2013), but they are impacted by canopy structure, thus deviating from the leaf-level physiological response (Guanter et al., 2014; Zeng et al., 2019). Compared to traditional vegetation indices that are mainly sensitive to vegetation structure, SIF can mechanistically constrain GPP through directly informing the actual electron transport rate which links the light and dark reactions (Gu et al., 2019; Ryu et al., 2019). This unique physiology dependence, in addition to the growing number of SIF measurements across multiple spatiotemporal scales (Frankenberg et al., 2011; Guan et al., 2016; Joiner et al., 2013, 2011; Sun et al., 2017; Yang et al., 2015), demonstrate the great potential of SIF to resolve the NEE partitioning problem (Kira et al., 2021).

However, the ecosystem-level relationship between SIF and GPP still remains unclear, especially at shorter time scales. A growing number of studies have highlighted a near-linear GPP-SIF relationship over different ecosystems across long time scales (weekly to monthly) (Frankenberg et al., 2011; Guan et al., 2016; Magney et al., 2019; Sun et al., 2018; Yang et al., 2015). But at shorter time scales, the GPP-SIF relationship shows stronger responses to physiological factors (e.g., atmospheric dryness). Under water-stressed conditions, SIF is found to decouple with GPP so that their linear relationship can break down (Helm et al., 2020; Kim et al., 2021; Marrs et al., 2020; Wohlfahrt et al., 2018). Yet, our understanding of this non-linear GPP-SIF relationship remains limited, especially at the ecosystem scale where vegetation structure and strong microclimate variations can have a large impact on the observed SIF and GPP (Dechant et al., 2020; Gentine et al., 2019). Furthermore, till now, the ecosystem-scale GPP-SIF relationships are all retrieved using GPP partitioned by standard methods, which highly rely on empirical models. The structural errors of these empirical models could bias the GPP-SIF relationship. Therefore, a non-empirical understanding of the ecosystem-scale GPP-SIF relationship is critically necessary to better capture GPP using satellite SIF, and also to provide constraints for ecosystem models representing SIF (Van der Tol et al., 2014; van der Tol et al., 2009).

In this study, we show that a new NN framework informed by SIF observations (NN_{SIF}) can tackle two problems simultaneously: (1) partitioning NEE into GPP and ER without prescribed empirical models, and (2) unraveling the GPP-SIF relationship at the ecosystem scale. As for the first goal, we compared NN_{SIF}-partitioned CO_2 fluxes against the results of NT, DT, and the method of Tramontana et al. (2020) (i.e., NN without SIF information, NN_{noSIF}). Specifically, we evaluated whether NN_{SIF} can reduce the structural errors of standard methods, and whether it can alleviate the equifinality issue of NN_{noSIF}. As for the second goal, we evaluated the GPP-SIF relationship retrieved by NN_{SIF}, NT, DT, and NN_{noSIF}, and explored whether NN_{SIF} can better elucidate the physiological controls on the GPP–SIF relationship at the ecosystem scale. To the best of our knowledge, this study represents the first direct and non-empirical assessment of the ecosystem-scale GPP-SIF relationship, as it does not rely on any prior empirical assumptions on the relationships between CO_2 fluxes and climatic drivers (e.g., the light response curve), nor does it use any empirical formulations on the GPP-SIF relationship (e.g., linear or hyperbolic equations).

2. Materials and methods

2.1. Machine learning algorithms

We proposed a ML-based framework aiming at 1) more accurately partitioning NEE, 2) retrieving the ecosystem-scale GPP-SIF relationship. Similar to the scheme of Tramontana et al. (2020), we used two subnetworks to predict ER and GPP respectively. Then we calculated the difference between predicted ER and GPP (Eq. (1)) to retrieve the final output NEE. The difference is that Tramontana et al. (2020) used NEE as the only constraint, while we used both NEE and SIF constraints to better partition NEE (our first goal). The SIF constraint was introduced by adding another subnetwork following the GPP subnetwork (Fig. 1). This additional subnetwork predicted SIF using the previous GPP estimates. Then, two final outputs, NEE and SIF, were compared against their observations to minimize the errors, by which the relationships between inputs (climatic drivers), intermediate outputs (ER, GPP), and final outputs (NEE, SIF) were learned by NN_{SIF}. In this way, ER and GPP were simultaneously constrained by both NEE and SIF observations, as shown by the dashed green arrows in Fig. 1. In addition, the proposed framework is also intended to unravel the ecosystem-scale GPP-SIF relationship (our second goal). In our model, SIF is predicted as a function of GPP and environmental factors (Eq. (2)). A well-trained NN_{SIF} can learn this function, without any empirical assumptions, and thus unravel how the GPP-SIF relationship responds to physiological factors at the ecosystem scale. Note that NN_{SIF} does not artificially enforce stronger linear relationships between GPP and SIF, although it includes the observed SIF as a constraint. Indeed, any complicated GPP-SIF
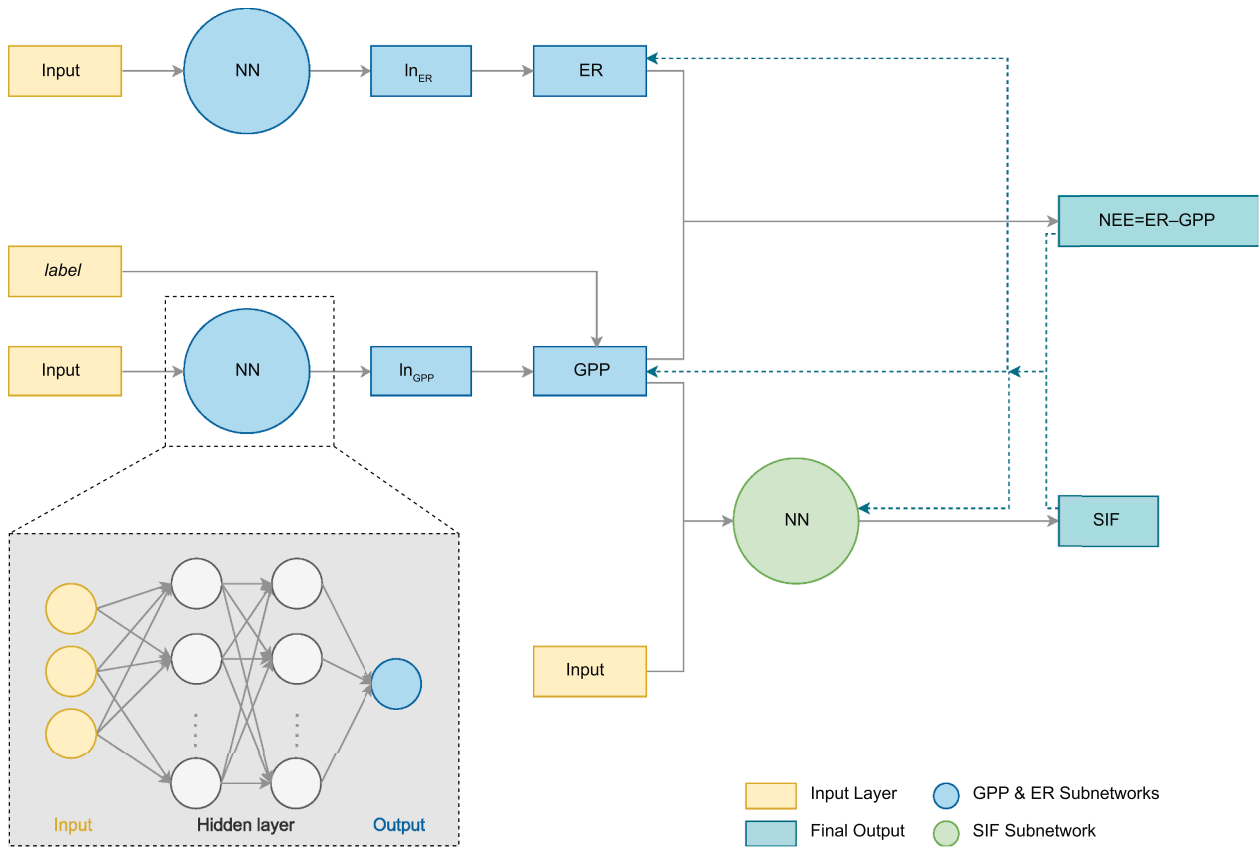
Fig. 1. *Schematic of the NN_SIF architecture.* NN$_{SIF}$ consists of three subnetworks. The shaded rectangle in the lower-left panel shows the detailed structure of each subnetwork. In the feedforward process (solid gray arrows), NEE and SIF were predicted using the current weights of three subnetworks. Then they were compared against observations through the loss function (Eq. (3)). In the backpropagation process (dashed green arrows), weights of each subnetwork were updated together, end-to-end, to minimize the loss function, and to make the predicted NEE and SIF closer to their observations. By this means, GPP and ER were simultaneously constrained by NEE and SIF observations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

relationship can be learned by NN$_{SIF}$, not only the linear one, according to the Universal Approximation Theorem (Hornik et al., 1989). Hence, NN$_{SIF}$ can capture complex and non-linear relationships between GPP and SIF.

The schematic of the NN$_{SIF}$ is shown in Fig. 1. The NN$_{SIF}$ framework has three subnetworks: one for GPP prediction, one for ER prediction, and another one for SIF prediction. The NN$_{SIF}$ inputs are meteorological and biotic drivers. To constrain GPP to be zero at night, we multiplied the GPP neuron by another input *label*, which equals zero at night and equals one during daytime. Here, nighttime was defined as times when PAR was smaller than 50 μmol m^{-2} s^{-1}. To accelerate the training process, we performed data normalization for all inputs except *label*: we subtracted the mean and divided by the standard deviation to ensure that input variables have zero mean and unit variance. The optimal subnetwork structure here has three layers: (1) an input layer driven by input variables, (2) two hidden layers that use rectified linear unit (ReLU) as the activation function, (3) an output layer that estimates the logarithmic values of GPP and ER to ensure non-negative results. Finally, all subnetwork outputs were combined at the last layer, using the following Eqs. (1) and (2), to get final NEE and SIF outputs. In Eqs. (1) and 2, RH is relative humidity. T$_{air}$ is air temperature. SWC is soil water content, and when SWC measurements is not available, evaporative fraction (fraction of the latent heat flux to available energy, EF) is used to indicate the soil moisture variations (Gentine et al., 2007, 2011).

$$NEE = ER - GPP \qquad (1)$$

$$SIF = f(GPP, RH, T_{air}, SWC \text{ or } EF) \qquad (2)$$

The input variables for NN$_{SIF}$ are summarized in Table 1. As for ER

Table 1
The input variables of NN$_{SIF}$ to estimate GPP, ER, and SIF.

Subnetwork	Temporal resolution	Input variables	Data source
ER	(Half-)hourly	T$_{air}$	Measured
		SWC	Measured
		W$_s$	Measured
		W$_d$	Measured
	Daily	NEE$_{night}$	Calculated from the measured NEE
		EF[1]	Calculated from the measured heat fluxes
	4-day	LAI	MCD15A3H
GPP	(Half-)hourly	T$_{air}$	Measured
		RH	Measured
		SWC	Measured
		PAR	Measured
		W$_s$	Measured
		W$_d$	Measured
	Daily	EF[1]	Calculated from the measured heat fluxes
	4-day	LAI	MCD15A3H
SIF	(Half-)hourly	GPP	Predicted internally by NN$_{SIF}$
		T$_{air}$	Measured
		RH	Measured
		SWC	Measured
	Daily	EF[1]	Calculated from the measured heat fluxes

[1] : EF is only used when the measured SWC is not available.

and GPP estimation, T_{air} was selected due to its effect on the chemical reactions in plants. SWC and RH were selected to account for the water stress, and the effect of atmospheric dryness on stomatal conductance. We selected RH rather than vapor pressure deficit (VPD), given the high collinearity between T_{air} and VPD (Novick et al., 2016; Zhou et al., 2019a, 2019b). Photosynthetic active radiation (PAR) was selected since it is critical for photosynthesis light reactions. To account for the effect of substrate availability on ER, we calculated the daily average of nighttime NEE (NEE_{night}) for ER prediction (Tramontana et al., 2020). Here, NEE_{night} was estimated only for days when there were at least 8 h nighttime NEE records, and for days with less than 8 h records, we used linear interpolation to fill the missing gaps of NEE_{night}. We also included leaf area index (LAI) to account for the biomass effect (Reichstein et al., 2003), which was found to improve the estimation of ER, GPP, and NEE in this study (Table S1). The LAI data were obtained from Moderate Resolution Imaging Spectroradiometers (MODIS). We identified good-quality LAI (i.e., generated by the main algorithm, no significant cloud) based on the quality flag, and filled the missing gaps using the multi-year average (2002–2020) of the same day. Wind speed (W_s) and wind direction (W_d) were included for ER and GPP estimations due to their effects on flux tower footprint, but they were not included for SIF prediction, since the footprint of SIF measurements remained invariant. SIF was predicted based on the GPP value which was estimated before the NN_{SIF} optimization. We used RH, T_{air}, and SWC to model the physiological controls on the GPP-SIF relationship. We also tested the inclusion of the escape ratio, f_{esc}, to reflect the impact of canopy structure (Guanter et al., 2014). f_{esc} was calculated using an empirical method (Zeng et al., 2019), as NIRv/fPAR, where NIRv is the near-infrared reflectance of vegetation (Badgley et al., 2017) and fPAR is the fraction of canopy-absorbed PAR. However, f_{esc} did not significantly improve the NN_{SIF} performance (Table S1), and was finally excluded for SIF prediction to simplify the NN_{SIF} inputs.

To optimize NN_{SIF} end-to-end, we minimized the loss function that quantifies the differences between observations and predictions of both NEE and SIF. In the loss function (Eq. (3)), NEE_{obs} and SIF_{obs} are NEE and SIF observations respectively. NEE_{NN} and SIF_{NN} are the predictions of NEE and SIF respectively. N is the number of data points. J is the total loss. This loss function is the weighted sum of the mean squared error of NEE and SIF, which is achieved through multi-task learning.

$$J = w_1 \frac{1}{N}(NEE_{obs} - NEE_{NN})^2 + w_2 \frac{1}{N}(SIF_{obs} - SIF_{NN})^2 \quad (3)$$

The weights of NEE task (w_1) and SIF task (w_2) were inferred purely from the data, based on the uncertainties of NEE and SIF (Kendall et al., 2018). The term with higher uncertainty will be automatically assigned with a lower weight. To verify this statement, we injected increasing noises to SIF to amplify its uncertainties, while keeping the NEE data unchanged. We find that the SIF weight gradually decreases with the increasing SIF noise level, however, it does not become zero even when the noise is high and the GPP-SIF relationship is hard to recognize (Fig. S1), emphasizing that even in this case the SIF information is still being used by the algorithm.

To train and validate NN_{SIF}, we randomly divided data into three subsets: 50% for the training set to fit the model, 20% for the validation set to assess the model while tuning hyperparameters, and 30% for the test set to evaluate the final tuned model. To avoid overtraining, early stopping was triggered if the model performance on the validation set was not improved for over 10 epochs. Adam gradient descent algorithm (Kingma and Ba, 2014) was used to optimize NN_{SIF}. The hyperparameters, e.g., the number of hidden layers, were tuned using SHERPA Library (Hertel et al., 2018) to identify the minimum size NN_{SIF} that can achieve good performance. In the final optimum NN_{SIF}, there are two hidden layers with 64 nodes in each subnetwork. Note that NN_{SIF} with the same structure (e.g., the same number of layers and nodes) can generate slightly different results during each training process, as its weight initialization is randomized each time. Hence, we used

the deep ensembles method to quantify the structural uncertainty of NN_{SIF}. Specifically, we retrained the same NN_{SIF} 1000 times with different weight initializations. The standard deviation of predictions from 1000 models was used as an estimate of the NN_{SIF} structural uncertainty (Caldeira and Nord, 2020).

In addition to the NN_{SIF}, we also trained NN_{noSIF}, a model that has the same GPP and ER subnetworks as NN_{SIF} but does not have the SIF subnetwork. Both daytime and nighttime samples were used to train NN_{SIF} and NN_{noSIF}. The difference between NN_{SIF} and NN_{noSIF} is that when processing daytime samples, the GPP retrieval in NN_{SIF} is additionally constrained by daytime SIF observations, and this constraint in turn indirectly affects the ER estimation. Note that although NN_{noSIF} shares some similarities with the approach of Tramontana et al. (2020) (e.g., two subnetworks for GPP and ER estimation), there are distinct differences between these two approaches (as we aimed to have a fair comparison between our model with and without SIF): (1) their input predictors are different; (2) NN_{noSIF} does not use a light use efficiency (LUE) equation for the GPP estimation.

2.2. In-situ datasets

2.2.1. Tower-based SIF measurements

We used field SIF observations from three sites with different vegetation types. One specific NN_{SIF} was built and trained for each site. Overall, NN_{SIF} shows good performance across three sites in predicting NEE and SIF (Table 2). The first site is Harvard Forest in Massachusetts, USA (42.538°N, 72.171°W). SIF was measured by an automated spectroscopy system installed 5 m above the canopy (Yang et al., 2018b, 2015). Canopy reflectance was measured by narrowband photodiode sensors which are 10 m above the canopy. Measurements were carried out during the growing season (May to October) in 2013 and 2014. SIF was retrieved at 760 nm using the spectral fitting method (Meroni et al., 2009). The second site is a wheat cropland at Montfavet near Avignon, France (43.917°N, 4.879°E). SIF and canopy reflectance were measured by TriFLEX, which was deployed at the height of 21 m to observe from nadir. SIF was retrieved at 760 nm using the nFLD method (Daumard et al., 2010). Measurements were conducted during the growing season in 2010. The growing season is from March 3, 2010 (end of tillering) to July 13, 2010 (harvest). The third site is a rice paddy located in Cheorwon, Gangwon province, South Korea (38.201°N, 127.251°E). SIF was observed by an automated spectroscopy system that has a QE Pro-spectrometer with very high spectral resolution (Dechant et al., 2019). SIF was retrieved by the singular vector decomposition (SVD) method (Yang et al., 2018a). Canopy reflectance was measured with two instruments that simultaneously monitored the incoming and outgoing

Table 2

The performance of NN_{SIF} and NN_{noSIF} in estimating NEE and SIF. We train NN_{SIF} 1000 times to quantify the model uncertainty. Statistics in this table are the mean and the standard deviation (in brackets) of R^2 and RMSE from the 1000 models. The unit for NEE RMSE is $\mu mol\ m^{-2}\ s^{-1}$. The unit for SIF RMSE is $mW\ m^{-2}\ nm^{-1}$ for Cheorwon, and is $mW\ m^{-2}\ nm^{-1}\ sr^{-1}$ for other two sites.

		Observed NEE v.s. estimated NEE		Observed SIF v.s. estimated SIF	
		R^2	RMSE	R^2	RMSE
Harvard Forest	NN_{SIF}	0.90 (± 0.003)	2.78 (± 0.04)	0.78 (± 0.01)	0.19 (± 0.01)
	NN_{noSIF}	0.92 (± 0.003)	2.54 (± 0.11)	/	/
Avignon	NN_{SIF}	0.92 (± 0.003)	3.16 (± 0.07)	0.90 (± 0.01)	0.16 (± 0.01)
	NN_{noSIF}	0.93 (± 0.005)	3.07 (± 0.12)	/	/
Cheorwon	NN_{SIF}	0.89 (± 0.004)	2.00 (± 0.05)	0.85 (± 0.01)	0.57 (± 0.03)
	NN_{noSIF}	0.88 (± 0.01)	2.14 (± 0.10)	/	/

radiation (Dechant et al., 2020). The observations were conducted during 2016. The growing season of the Cheorwon rice paddy site is from the end of April (Day of Year (DOY), 120) to harvest in early September (DOY, 248). Note that Cheorwon observations were made with a cosine corrector (as opposed to bare fibers in other two sites), and thus have the different SIF unit (mw m^{-2} nm^{-1} v.s. mw m^{-2} nm^{-1} sr^{-1}). For more details on the three sites, the instruments, and SIF retrieval methods, the reader is referred to Yang et al. (2015), Goulas et al. (2017), and Yang et al. (2018a).

2.2.2. Eddy-covariance flux and meteorological measurements

The in-situ datasets also include EC flux and meteorological measurements. The Harvard Forest dataset is available from the Harvard Forest Data Archive (Munger and Wofsy, 2021). Avignon and Cheorwon have half-hourly measurements, while Harvard Forest was recorded hourly. The meteorological measurements include T_{air}, RH, soil temperature (T_{soil}), SWC, W_s, W_d, PAR, and diffuse PAR. For Avignon and Cheorwon, we calculated the daily-averaged evaporative fraction (fraction of latent heat to available energy), to approximate SWC changes (Gentine et al., 2007, 2011), because SWC measurements of these two sites were not available, or had long-term gaps (more than 2 months) due to malfunctioning issues. The fraction of the canopy-absorbed PAR (fPAR) was quantified using multiple sensors under and above the canopy (Goulas et al., 2017; Kim et al., 2019; Yang et al., 2015). LAI was acquired from the MCD15A3H dataset through Global Subset Tool (https://modis.ornl.gov/globalsubset/), with the resolution of 500 m and 4-day. We extracted the LAI data from the pixel where the site is located in. NEE was measured by the EC technique and went through quality control procedures (Papale et al., 2006). Here, we did not fill NEE gaps, but only used good-quality NEE observations to avoid any circularities. NEE is gap-filled using its relationship with environmental inputs (Aubinet et al., 2012), while here, NN$_{SIF}$ was trained to learn this relationship. Thus, using gap-filled NEE to train NN$_{SIF}$ causes circularities. We also partitioned NEE using standard NT and DT methods by REddyProc (Wutzler et al., 2018). T_{air} was selected as the driving temperature for NT and DT. For details about the instrumentation, the reader is referred to Munger and Wofsy (2021), Goulas et al. (2017), and Hwang et al. (2020).

2.3. Synthetic dataset

In the real world, the true GPP and ER during daytime are unknown. Hence, an analysis based on synthetic data, which provides a reference for GPP and ER, is used to evaluate NN$_{SIF}$, NT, DT, and NN$_{noSIF}$. We generated synthetic data using the Soil-Canopy-Observation of Photosynthesis and Energy fluxes (SCOPE). SCOPE is a radiative transfer model that simulates ChlF of leaf and canopy, and CO_2 fluxes between soil, vegetation, and atmosphere (Van der Tol et al., 2014; van der Tol et al., 2009). We ran SCOPE (v1.73) at Harvard Forest to simulate GPP, ER, NEE, and SIF at the hourly scale. We think that the simulation of one site is sufficient to test whether NN$_{SIF}$ can learn the functional relationships implemented in SCOPE, as SCOPE is not structurally changing from one site to another (only the climatic inputs and parameters are changing). The simulation period was the same as the observation (growing season in 2013 and 2014). The canopy-scale SCOPE outputs were used in this study. We drove SCOPE using meteorological and LAI time series. Details about the inputs and parameters of SCOPE simulation are summarized in Table S2.

The procedure to extract canopy-scale GPP and ER from SCOPE outputs is as follows. For the GPP simulation, we extracted the outputs of canopy apparent photosynthesis, i.e., carboxylation minus photorespiration (Wohlfahrt and Gu, 2015). For the ER simulation, it was the sum of two parts: leaf respiration (R_{leaf}) and soil respiration (R_s). SCOPE simulates R_{leaf} using the Lloyd and Taylor equation driven by canopy temperature (T_{canopy}) (Lloyd and Taylor, 1994). The R_{ref} parameter of R_{leaf} was downregulated by the maximum carboxylation rate (V_{cmax})

(Collatz et al., 1991). With regard to the R_s part, the original SCOPE model forces R_s to zero. Hence, we used Eq. (4) driven by T_{soil} to model R_s (Migliavacca et al., 2011). Note that precipitation is used in the original model of Migliavacca et al. (2011), but in SCOPE, SWC is directly used to downscale V_{cmax} (Bayat et al., 2019). Thus, we used SWC instead to describe the regulation of moisture on R_s, following Zhang et al. (2018). In Eq. (4), E_0 is temperature sensitivity. T_{ref} is the reference temperature (15 °C). T_0 is constant of −46.02 °C. SWC_0 is the soil field capacity. Here, the stem and root respiration were not included, as SCOPE does not explicitly simulate the biochemical processes of plant stem and root.

$$R_s = R_{ref} \times \exp\left(E_0 \left(\frac{1}{T_{ref} - T_0} - \frac{1}{T_{soil} - T_0} \right) \right) \times \left(1 - \left(\frac{SWC}{SWC_0} - 1 \right)^2 \right) \quad (4)$$

The SCOPE-simulated ER and GPP have similar magnitudes to the in-situ data partitioned by NT (Fig. S2), with an R^2 of 0.83 for GPP and an R^2 of 0.54 for ER. We believe that this agreement is acceptable, as even for the in-situ ER and GPP, their estimates still show noticeable differences among different partitioning methods (Fig. 5). For example, the R^2 between NT and DT is 0.92 for GPP and 0.58 for ER at Harvard Forest (Fig. S3 (b, f)), which is similar to the above-mentioned R^2 between the SCOPE simulations and NT estimates. To imitate in-situ data, we added artificial noises to the simulated NEE and SIF. For SIF, we added a random Gaussian noise with zero mean and 0.1 mw m^{-2} nm^{-1} sr^{-1} standard deviation, based on typical SIF retrieval errors. For NEE, the noise have two components: (1) a white Gaussian noise with zero mean and 0.1 μmol m^{-2} s^{-1} standard deviation; (2) a heteroscedastic random noise that increases with NEE magnitude (~ 8% of the NEE magnitude) (Lasslop et al., 2008; Tramontana et al., 2020).

We applied NN$_{SIF}$, NT, DT, and NN$_{noSIF}$ methods to partition the SCOPE-simulated NEE, and then the partitioned ER and GPP were compared against the SCOPE simulations. The inputs for these four methods were the same as those used to partition in-situ NEE of the three sites, except that W_d was not included for NN$_{SIF}$ and NN$_{noSIF}$, as W_d was not available from SCOPE simulations. We found that NN$_{SIF}$ can successfully emulate SCOPE in predicting GPP (R^2=0.991), ER (R^2=0.858), NEE (R^2=995), and SIF (R^2=0.988) (Fig. S4).

In addition to normal climate conditions, we also tested the NN$_{SIF}$ performance under severe water and heat stress where the GPP-SIF relationship can be highly non-linear. We used the SCOPE model to run a 3-month extremely drought simulation. The input climatic data are visualized and summarized in Fig. S5: T_{air} was gradually increased from 20 °C to over 40 °C, while both SWC and RH gradually decreased over time. The air VPD can be extremely high and can even approach 10 kPa. As a result, the SCOPE-simulated GPP-SIF relationship is highly non-linear, and this relationship is even slightly inverted at the highest VPD (Fig. S6 (a)). We found that NN$_{SIF}$ still shows good performance in predicting GPP (R^2=0.967), ER (R^2=0.857), NEE (R^2=0.987), and SIF (R^2=0.977), although the R^2 value is slightly lower than the original experiment under normal climate simulations (Fig. S7).

2.4. Statistical analysis

With regard to the first goal of NN$_{SIF}$, i.e., partitioning NEE, we compared GPP and ER partitioned by NN$_{SIF}$ with those derived by NT, DT, and NN$_{noSIF}$. The comparison was performed on both synthetic and in-situ data. We also compared the ecosystem LUE estimated by all methods. LUE was calculated as normalizing GPP by canopy-absorbed PAR (APAR). To avoid outliers, we selected LUE within the range of 1–99th percentiles. Besides, we specifically verified whether these four methods can correctly retrieve the functional relationship between climatic drivers and CO_2 fluxes (e.g., the GPP-PAR relationship). This verification was performed on SCOPE simulations which can provide the known (as imposed) reference relationship. To obtain the functional relationship, we firstly sorted data into 20 bins according to quantiles of

the climatic driver, then we averaged the climatic driver and the CO_2 flux in each bin to get the response curve.

As for the second goal of NN_{SIF}, i.e., unraveling the ecosystem-scale GPP-SIF relationship, we performed linear regression to evaluate the GPP-SIF relationship retrieved by NN_{SIF}, NT, DT, and NN_{noSIF}. Furthermore, we applied Pearson correlation to analyze the relationship between the GPP/SIF ratio and physiological factors. To avoid outliers, GPP/SIF ratio within the range of 1–99th percentile was selected. For both linear regression and Pearson correlation, a significant threshold of 0.05 was used to decide whether to reject a null hypothesis.

Furthermore, we analyzed how the NN_{SIF} performance was affected by the uncertainties stemming from the input data (i.e., SIF, NEE, and LAI). We imposed noises to both NEE and SIF SCOPE simulations, and the noise intensity gradually increased from *level1* to *level10*. Gaussian noises were added to SIF. The standard deviation of noises ranged 0.05–0.5 and increased by 0.05 per level, which approximately corresponded to 10–50% of the SCOPE SIF magnitude. Heteroscedastic noises

were added to NEE. The standard deviation of noises ranged from 4% to 40% of NEE magnitude, and increased by 4% per level. We then explored how the increasing uncertainties from SIF and NEE affected the accuracy of NN_{SIF} (Fig. S16). In addition to SIF and NEE, we also analyzed the effect of LAI uncertainties (Fig. S17). Gaussian noises were added to the input LAI data, and the noise standard deviation gradually increased from 0.2 to 2 across 10 levels, which approximately corresponds to 10–50% of the LAI magnitude.

3. Results

3.1. Partitioning of CO_2 fluxes using machine learning

3.1.1. Comparisons of ecosystem photosynthesis among different methods

As for GPP estimation, we find high R^2 among NN_{SIF}, NT, DT, and NN_{noSIF} methods (Fig. S8). However, high R^2 is not sufficient to indicate good performances of these four methods, as high R^2 can be mainly due
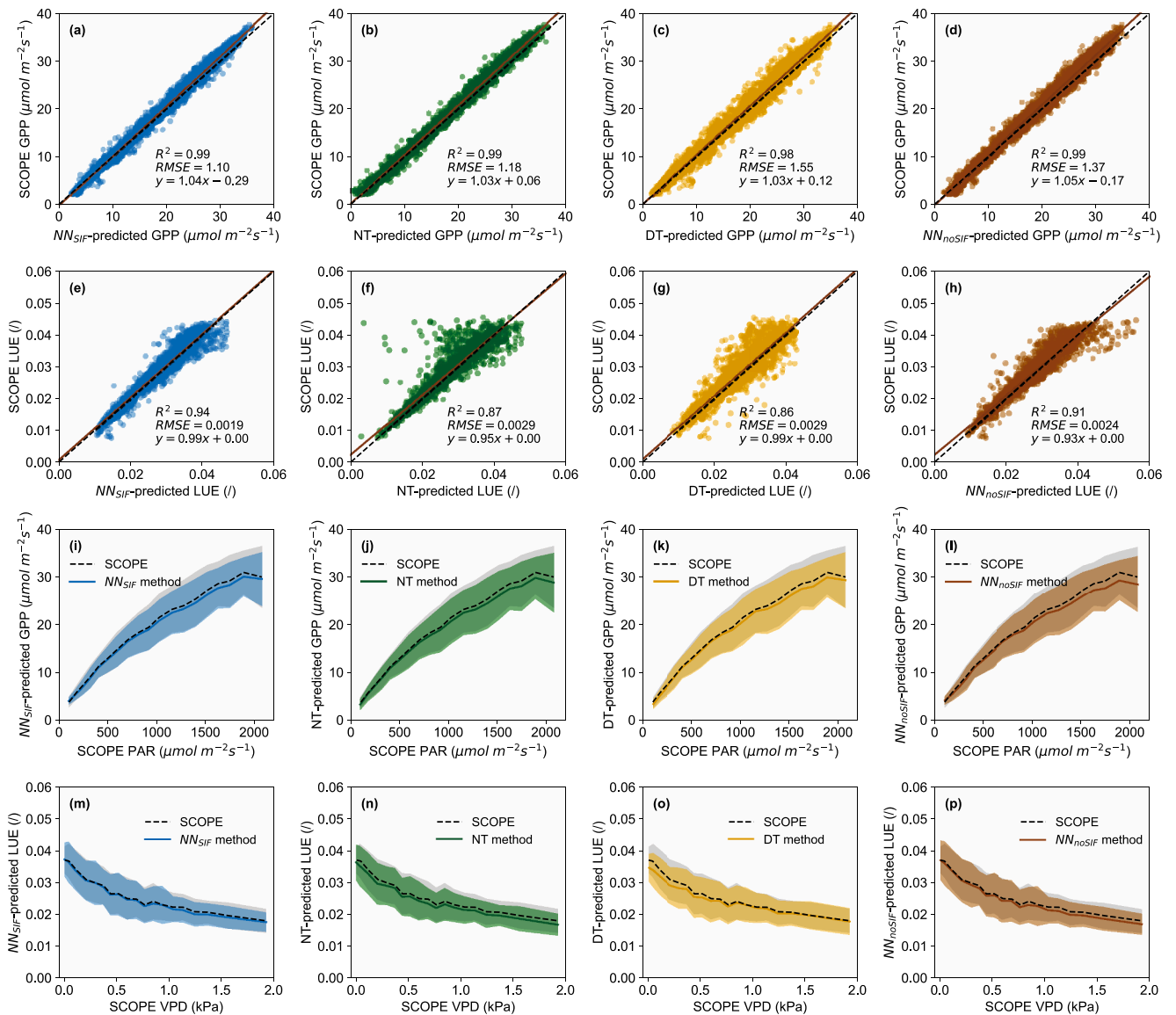

Fig. 2. *The performance of NN_{SIF}, NT, DT, and NN_{noSIF} in estimating SCOPE GPP and LUE.* Here, the NEE partitioning is based on SCOPE simulations which can provide the "truth" LUE. The SCOPE simulations are based on data from the Harvard Forest site. *The first and second rows*: comparisons of GPP (a-d) and LUE (e-h) between SCOPE references and the predictions of four methods. The brown solid lines denote the fitted regression models, and black dashed lines denote the diagonal lines. *The third and fourth row*: the GPP-PAR (i–l) relationship and the LUE-VPD (m–p) relationship retrieved by the four methods. Shaded areas indicate the standard deviation of the relationship. Dark dashed lines denote the SCOPE relationship. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to the strong diurnal and seasonal cycles driven by APAR. The LUE comparison, which reduces the effects of the strong diurnal and seasonal cycles, can more stringently evaluate the method performance in estimating ecosystem photosynthesis. Among four methods, NN_{SIF} shows the highest agreement with the reference SCOPE LUE, with the highest R^2 and the lowest RMSE (Fig. 2(e)). In addition, we also take only midday GPP (11:00–13:00 local time) for comparison, which can reduce the diurnal cycle effect. The midday GPP comparison shows that NT and NN_{SIF} achieve slightly better agreement with the reference SCOPE GPP based on the R^2 and RMSE values (Fig. S9).

Moreover, we evaluate whether different partitioning methods correctly learn the functional relationships implemented in SCOPE. All four methods can effectively reproduce the GPP-PAR relationship (Fig. 2 (i, l)) as well as the LUE-VPD relationship (Fig. 2 (m–p)). However, the GPP response curve retrieved by NN_{noSIF} shows relatively large deviations at higher PAR conditions (Fig. 2(l)), and the DT method slightly underestimates LUE at humid conditions (low VPD, Fig. 2(o)). Overall, NN_{SIF} performs well in retrieving the GPP-PAR and LUE-VPD functional relationships, across the whole range of PAR and VPD (Fig. 2 (i, m)).

When compared to standard methods, NN_{SIF} has better agreement with DT than with NT for LUE estimation at the (half-)hourly time scale, and this trend is consistent for both synthetic (Fig. 3 (a, b)) and *in-situ* (Fig. 3 (d–e, g–h, j–k)) data. One reason for this higher agreement with DT is that the DT GPP is estimated directly by the fitted model and thus has less noise, while the NT GPP is estimated indirectly by subtracting the estimated ER from NEE observations, which is more susceptible to the measurement noises especially at the half-hourly time scale. After aggregating to daily and weekly scales, NN_{SIF} shows higher correlations with NT than with DT (except the Cheorwon site), as most of the high-frequency noises are removed by the temporal aggregation. Besides, the NN_{SIF}-estimated LUE shows good consistency with NN_{noSIF}, with an R^2 no less than 0.64 (Fig. 3 (c, f, i, l)).

Compared to other methods, NN_{SIF} can better capture the instantaneous effect of the harvest on GPP variations. The rice paddy in Cheorwon was harvested in early September 2016 (DOY, 248) (Yang et al., 2018a). The observed SIF dramatically decreases after the harvest (Fig. 4(a)). As for NN_{SIF}-predicted GPP, both its half-hourly value and midday average show a sudden drop following the harvest, and the
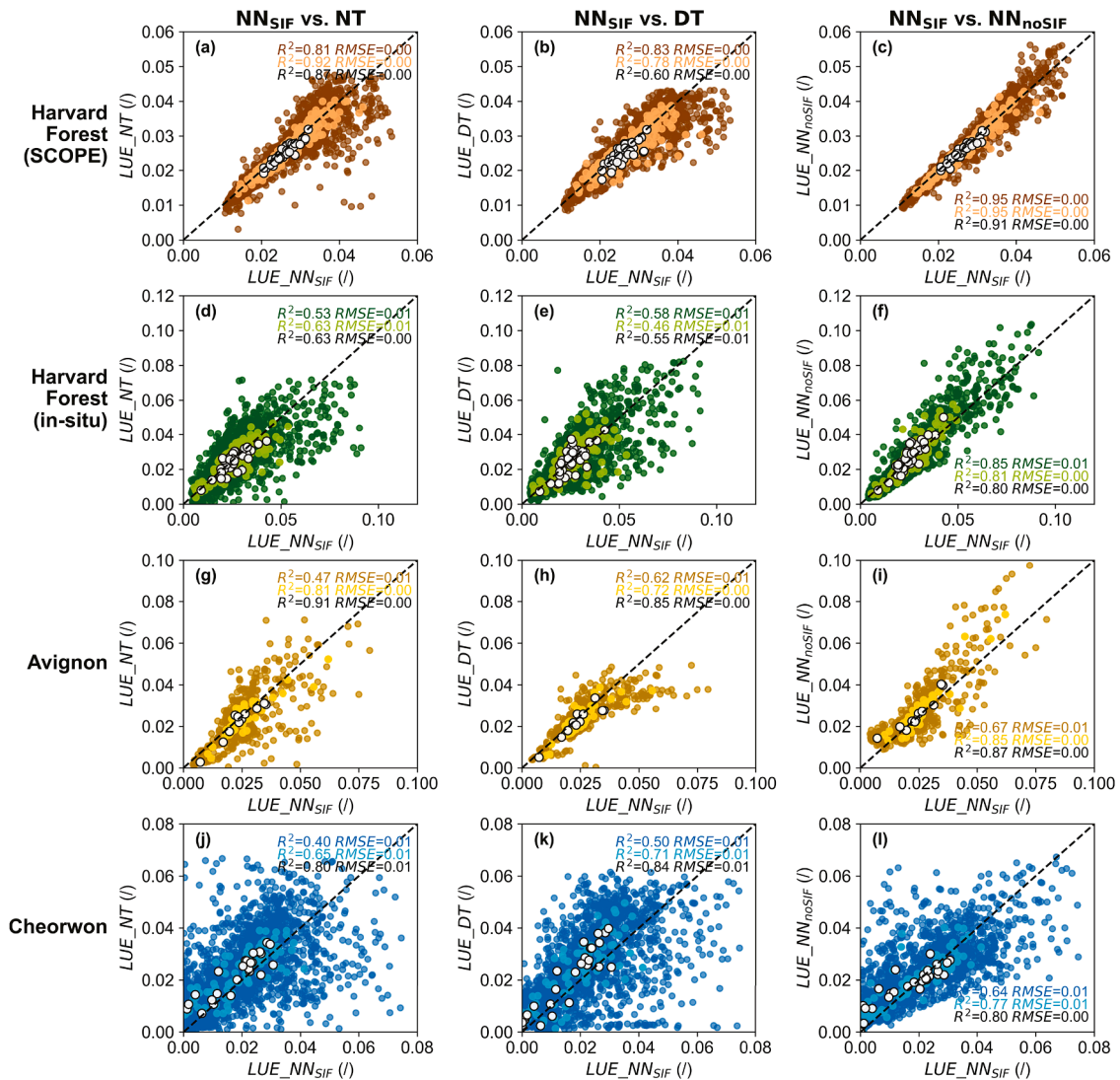

Fig. 3. *Comparisons of NN_{SIF}-predicted LUE against LUE estimated by NT, DT, and NN_{noSIF}.* The first row (a–c) shows comparisons of SCOPE simulations obtained from Harvard Forest. The other rows show comparisons of *in-situ* data at Harvard Forest (d–f), Avignon (g–i), and Cheorwon (j–l). Different colors denote different time scales. White points represent the weekly scale. Light-colored points (light brown, light green, light yellow, and light blue) represent the daily scale. Dark-colored points (dark brown, dark green, dark yellow, and dark blue) represent the (half-)hourly scale. Here, daily average was estimated only for days when there were at least 8 h LUE records, and weekly average was estimated only for weeks when there were at least 5-day records. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
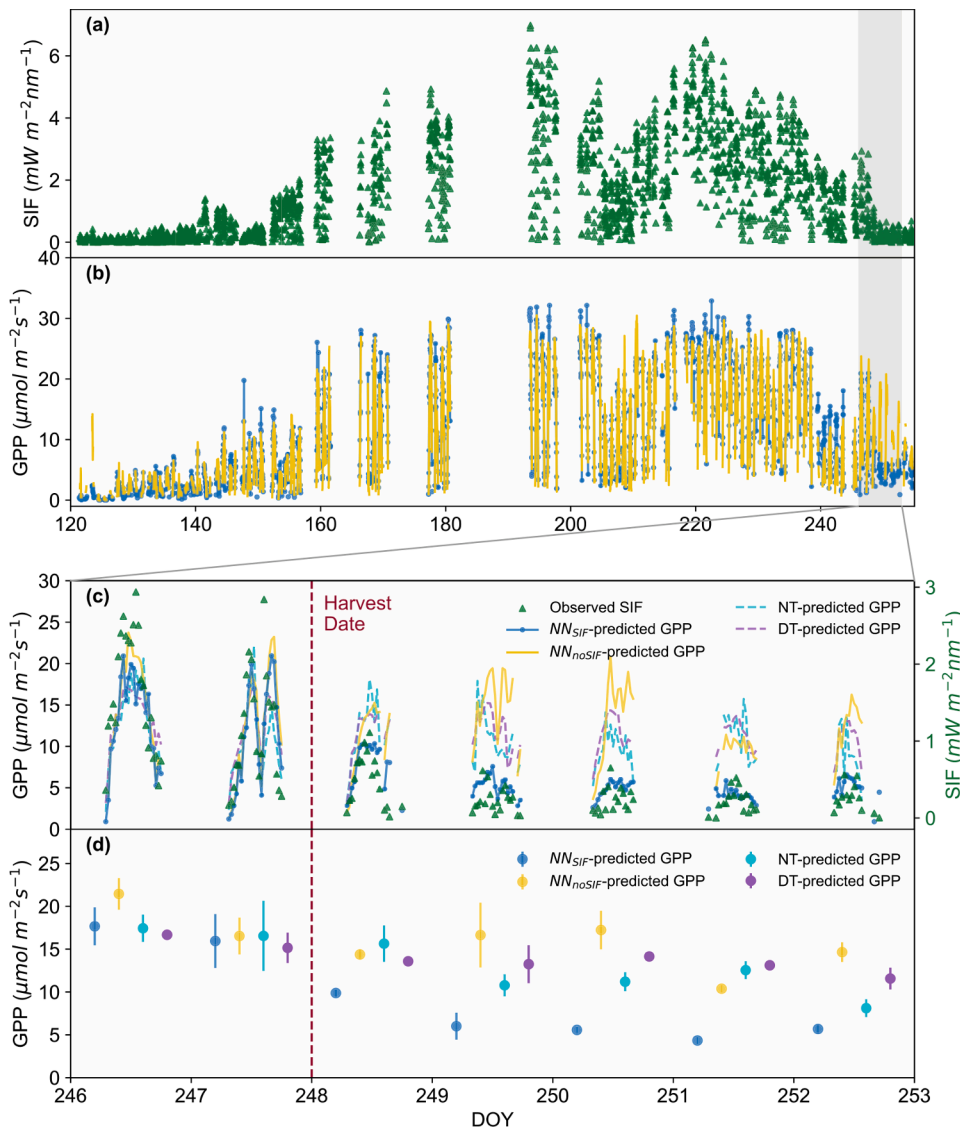
Fig. 4. *Time series of half-hourly (a) SIF and (b) GPP at the Cheorwon site. (c) and (d) are the zoom-in figure of the gray shadow zone in (a) and (b). (c) shows the diurnal variations of SIF and GPP partitioned by four methods during the 7-day window before and after the harvest date. (d) shows midday (11:00–13:00 local time) average of GPP partitioned by four methods. The error bar indicates the standard deviation during the midday period. The average was estimated only for days without missing datapoints during the midday period. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)*

half-hourly value is no larger than 10 μmol m^{-2} s^{-1} (Fig. 4 (c, d)). As for the other three methods, their midday average does not show a very sharp drop and remains at a similar magnitude to the pre-harvest values (Fig. 4(d)), which suggests that NN$_{SIF}$ is more sensitive to the rapid GPP variations due to harvest.

3.1.2. Comparisons of ecosystem respiration among different methods

As for ER estimation, NN$_{SIF}$ shows different consistency with the other three methods. NN$_{SIF}$ shows relatively strong consistency with NN$_{noSIF}$ with an R^2 no less than 0.57 (Fig. 5 (c, f, l)). When compared to standard methods, NN$_{SIF}$ has better agreement with NT than with DT for both synthetic (Fig. 5 (a, b)) and *in-situ* (Fig. 5 (d, e, g-h, j, k)) data. Overall, the consistency of ER estimation is evidently weaker than that of GPP estimation. However, despite these noticeable ER discrepancies, their magnitude is similar to the inherent differences between NT and DT methods according to the RMSE values (Fig. S3).

When tested on SCOPE simulations, NN$_{SIF}$ shows good performance in ER estimation, especially at high temperatures (Fig. 6). NN$_{SIF}$-predicted ER shows high consistency with SCOPE ER with the highest R^2 and the lowest RMSE (Fig. 6(a)). However, NT and NN$_{noSIF}$ underestimate ER at high T$_{air}$ (Fig. 6 (d, j)), while DT underestimates ER at low T$_{air}$ (Fig. 6(g)). The NN$_{SIF}$ outperformance is further corroborated from the retrieved ER-T$_{air}$ relationship. The SCOPE ER keeps increasing with

T$_{air}$, with higher slope ($\frac{\Delta ER}{\Delta T_{air}}$) at high T$_{air}$ (Fig. 6 (b, e, h, k)). Both NT and NN$_{noSIF}$ fail to reproduce the SCOPE relationship when T$_{air}$ is higher than the 95th percentile of nighttime T$_{air}$ (Fig. 6 (e, k)). The results of NN$_{SIF}$ and DT are generally consistent with the SCOPE relationship. However, the DT relationship (the green solid line) shows relatively larger deviations from the SCOPE relationship (the black dashed line) across the whole T$_{air}$ range (Fig. 6(h)). Overall, NN$_{SIF}$ shows the lowest deviations and can most correctly estimate the ER-T$_{air}$ relationship (Fig. 6 (a, b)).

The better performance of NN$_{SIF}$ demonstrates the role of the SIF constraint in improving daytime ER estimation. There are distinct differences between day and night SCOPE ER distributions (Fig. 6 (i, l)). The driving temperatures of ER, i.e., T$_{soil}$ and T$_{canopy}$, show different distributions between day and night (Fig. 6 (c, f)). Consequently, daytime ER has higher overall values and higher $\frac{\Delta ER}{\Delta T_{air}}$ slope than night ER (Fig. 6 (i, l)). NN$_{noSIF}$ shows noticeable biases at high temperatures outside the night T$_{air}$ range (Fig. 6 (d, e, j, k)), while NN$_{SIF}$, with the additional SIF constraint, can better estimate daytime ER especially at high T$_{air}$ (Fig. 6 (a, b)).

Besides, NN$_{SIF}$ shows superior performance in capturing the moisture limitation on ER (Fig. 7). The Avignon site is Mediterranean and seasonally water-limited. From late June to the end of observation period (early July), the daily evaporative fraction (EF) abruptly declined with half-hourly VPD higher than 3 kPa (Fig. S10). The four methods
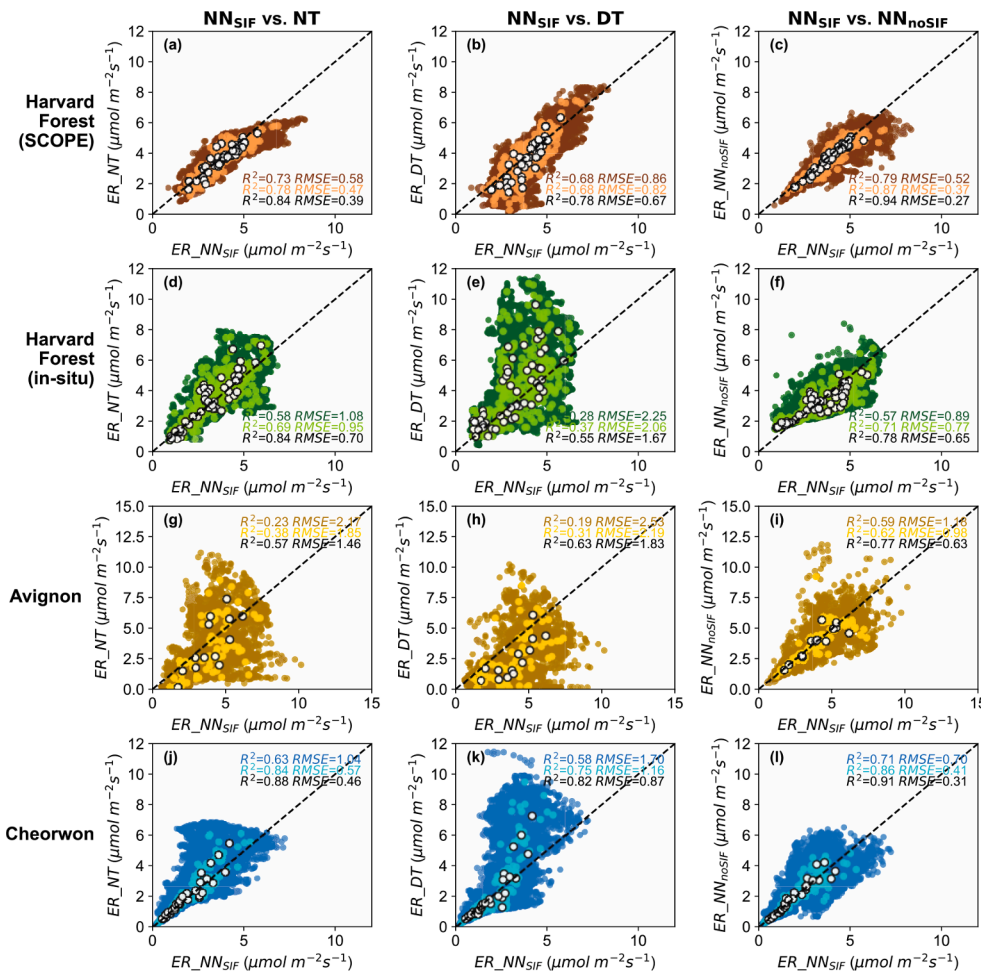
Fig. 5. *Comparisons of NN$_{SIF}$-predicted ER against ER partitioned by NT, DT, and NN$_{noSIF}$. The first row (a–c) shows comparisons of SCOPE simulations obtained from Harvard Forest. The other rows show comparisons of in-situ data at Harvard Forest (d–f), Avignon (g–i), and Cheorwon (j–l). Different colors denote different time scales. White points represent the weekly scale. Light-colored points (light brown, light green, light yellow, and light blue) represent the daily scale. Dark-colored points (dark brown, dark green, dark yellow, and dark blue) represent the (half-)hourly scale. Here, daily average was estimated only for days when there were at least 12 h ER records, and weekly average was estimated only for weeks when there were at least 5-day records. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)*

behave differently during this short-term dry period (~20 days), as indicated by the dark brown points in Fig. 7. NT-predicted ER remains relatively high values even at low EF (Fig. 7(b)). DT-predicted ER is suppressed under the water stress, but it still continues the increasing trend with T$_{air}$ under low EF conditions (Fig. 7(c)). In contrast, NN$_{SIF}$-predicted ER slows the increasing trend when T$_{air}$ approaches 20 °C and, instead, declines with T$_{air}$ especially at low EF (Fig. 7(a)), indicating moisture limitation on ER. NN$_{noSIF}$ similarly captures this temperature optimum effect, but it generates some abnormally high ER (higher than 11 μmol m^{-2} s^{-1}) at medium T$_{air}$ and EF (Fig. 7(d)).

3.2. The GPP-SIF relationship retrieved by NN$_{sif}$

3.2.1. The GPP-SIF relationship at different ecosystems

NN$_{SIF}$ can learn any complicated relationship between GPP and SIF, not limited to the linear one. Here, the GPP-SIF relationships derived by NN$_{SIF}$ show different characteristics across three sites. At the (half-)hourly time scale, the GPP-SIF relationship of Harvard Forest is relatively non-linear (Fig. 8(a)). The GPP/SIF slope slightly flattens with the increasing SIF, indicating a saturation effect of GPP under high irradiance conditions. Besides, the R^2 is improved from 0.86 to 0.87 when we switch from linear to hyperbolic model to fit the hourly relationship. As for Avignon and Cheorwon, the (half-)hourly GPP-SIF relationships are more linear with no obvious saturation effect (Fig. 8 (b, c)). After aggregating from (half-)hourly to daily and weekly scales, the NN$_{SIF}$-derived GPP-SIF relationships demonstrate an enhanced linearity with increasing R^2 across three sites.

In addition, we compared the NN$_{SIF}$ results against previous *in-situ* studies conducted at the three sites (Goulas et al., 2017; Yang et al.,

2018a, 2015). As for Avignon and Cheorwon sites, the half-hourly GPP/SIF slopes predicted by NN$_{SIF}$ (Fig. 8 (a, b)) are close to or even the same as half-hourly slopes obtained from previous studies (16.75 μmol nm sr s^{-1} mW^{-1} for Avignon, 6.16 μmol nm s^{-1} mW^{-1} for Cheorwon). The NN$_{SIF}$ slope of Harvard Forest (Fig. 8(c)) is between the *in-situ* slope at 1:30PM (13.71 μmol nm sr s^{-1} mW^{-1}) and 9:30AM (28.44 μmol nm sr s^{-1} mW^{-1}). These *in-situ* GPP/SIF slopes were obtained based on (half-)hourly SIF observations and GPP partitioned by the NT method.

3.2.2. The responses of GPP-SIF relationship to physiological factors

Using NN$_{SIF}$, we can infer how the ecosystem-scale GPP-SIF relationship responds to physiological factors at the (half-)hourly time scale, e.g., illumination conditions. At Harvard Forest and Avignon, the (half-)hourly GPP-SIF relationship derived by NN$_{SIF}$ shows higher slopes and R^2 on cloudy days than sunny days (Fig. 9 (a, b)). At Cheorwon, both sunny and cloudy days have a very high R^2 with similar slope (Fig. 9(c)). We also evaluated whether this radiation effect can be observed using the observed SIF rather than NN$_{SIF}$-predicted SIF, and found that Harvard Forest and Avignon sites still show higher slopes on cloudy days (Fig. S11 (a, b)). In addition, we conducted correlation analysis to quantitatively evaluate the relationship between the fraction of diffuse PAR (%PAR$_{dif}$) and the GPP/SIF ratio (i.e., the ratio of NN$_{SIF}$ GPP to the observed SIF). There are significant positive correlations between% PAR$_{dif}$ and the (half-)hourly GPP/SIF ratio ($p < 0.05$), especially for Avignon with the highest correlations among three sites (Fig. S12 (d)).

NN$_{SIF}$ can reveal the responses of the (half-)hourly GPP–SIF relationship to atmospheric dryness (Fig. 10). At higher VPD (denoted by
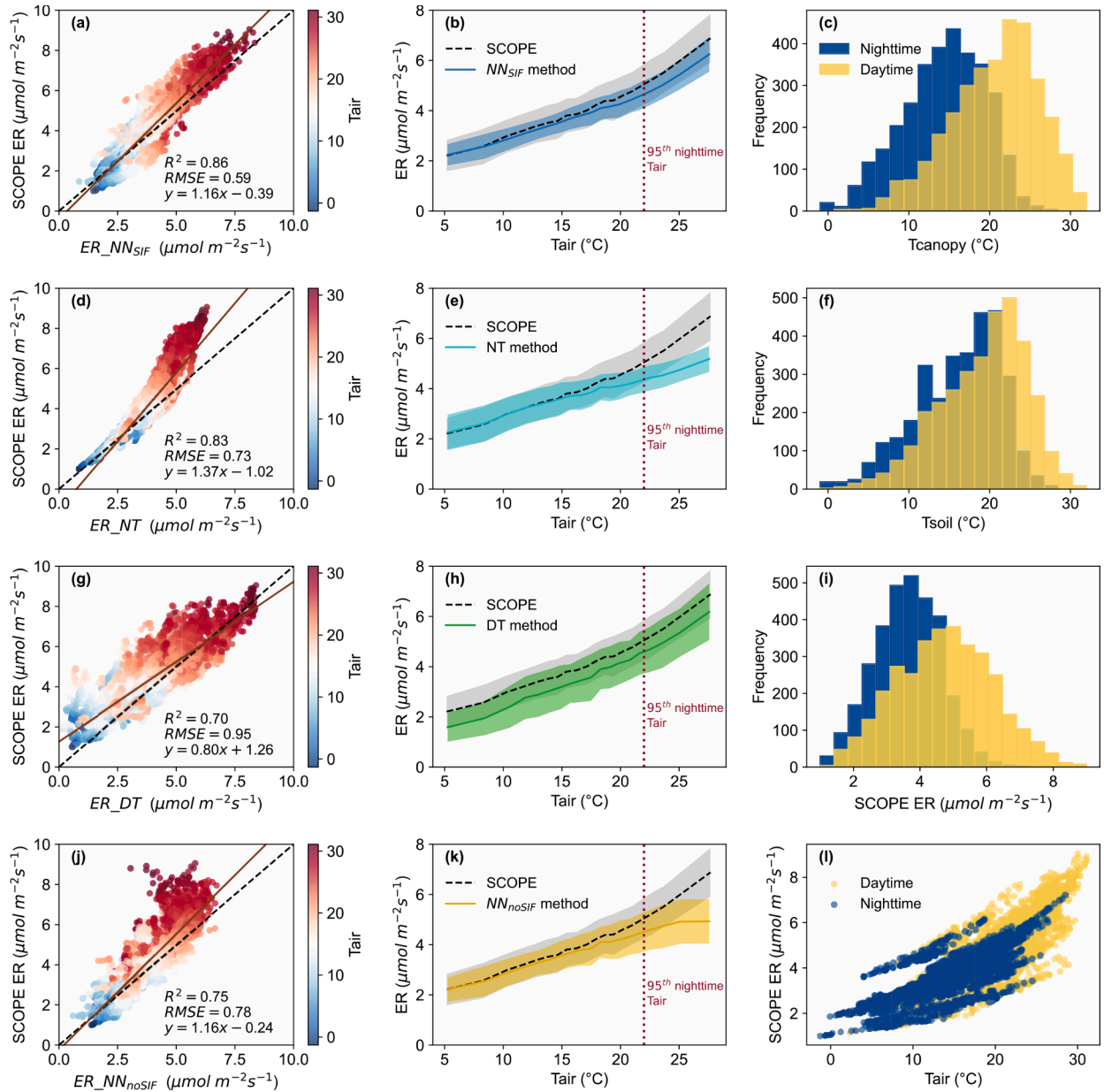
Fig. 6. *Performances of NN$_{SIF}$, NT, DT, and NN$_{noSIF}$ in estimating SCOPE ER.* The SCOPE simulations are obtained from Harvard Forest. *The leftmost column (a, d, g, j):* comparisons between SCOPE ER and ER estimated by four methods. The colormap indicates the changing T$_{air}$. Brown solid lines denote fitted regression models. Black dashed lines are diagonal lines. *The middle column (b, e, h, k):* the ER-T$_{air}$ relationship retrieved by four methods. Shaded areas indicate the standard deviation of the relationship. Dark dashed lines denote the SCOPE relationship. Red vertical dashed lines denote the 95th percentile of night T$_{air}$. **The rightmost column:** (c, f, i) show day and night distributions of T$_{canopy}$, T$_{soil}$, and SCOPE ER. (l) shows the SCOPE ER-T$_{air}$ relationship during day and night. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

blue points), the (half-)hourly GPP corresponding to the same SIF is depressed, i.e., the GPP/SIF ratio exhibits a decreasing trend with the increasing VPD (Fig. 10 (a–d)). This trend is consistent for both *in-situ* (Fig. 10 (a–c)) and synthetic data (Fig. 10(d)). Even for the synthetic data generated under extremely drought conditions, NN$_{SIF}$ can effectively reproduce the highly non-linear GPP-SIF relationship, and can retrieve how this relationship responds to the changing VPD (Fig. 10(e)). Similar VPD response is also visible when we replace the NN$_{SIF}$-predicted SIF with the observed SIF (Fig. S13), but becomes less prominent at Harvard Forest and Cheorwon sites (Fig. S13 (a, c)), which is partly due to the measurement noises in SIF observations. Furthermore, there are significant negative correlations between VPD and the (half-)hourly GPP/SIF ratio ($p < 0.05$) across the three sites. The correlation strength is moderate at Avignon and Harvard Forest, and is relatively lower at

Cheorwon (Fig. S14 (e)).

4. Discussion

In this study, we developed a new framework informed by SIF that can partition NEE, while simultaneously retrieving the ecosystem-scale GPP-SIF relationship. With regard to GPP estimation, NN$_{SIF}$ can correctly learn the functional relationships (GPP-PAR and LUE-VPD relationships), and can better capture the rapid GPP drop following the harvest. For ER estimation, NN$_{SIF}$ shows better performance in retrieving the ER-T$_{air}$ relationship and capturing the moisture limitation on ER, especially at high T$_{air}$. Furthermore, NN$_{SIF}$ can retrieve the ecosystem-scale GPP-SIF relationship, and elucidate how the GPP-SIF relationship responds to radiation and humidity. Overall, we believe
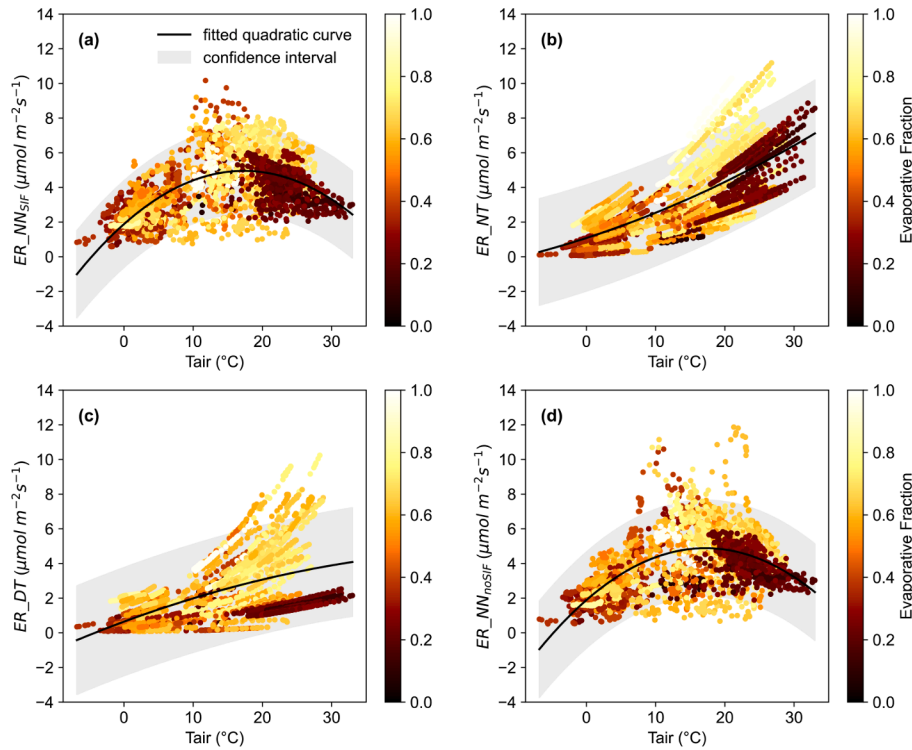
Fig. 7. *The half-hourly relationship between ER and T_{air} at Avignon retrieved by (a) NN_{SIF}, (b) NT, (c) DT, (d) NN_{noSIF} methods.* The colormap indicates the changing evaporative fraction, a proxy for soil moisture. A quadratic polynomial model is used to fit the ER-T_{air} relationship. The shaded area indicates the 95% confidence interval. Here, T_{air} is used to indicate the overall temperature, as soil temperature measurements are not available at the Avignon site.
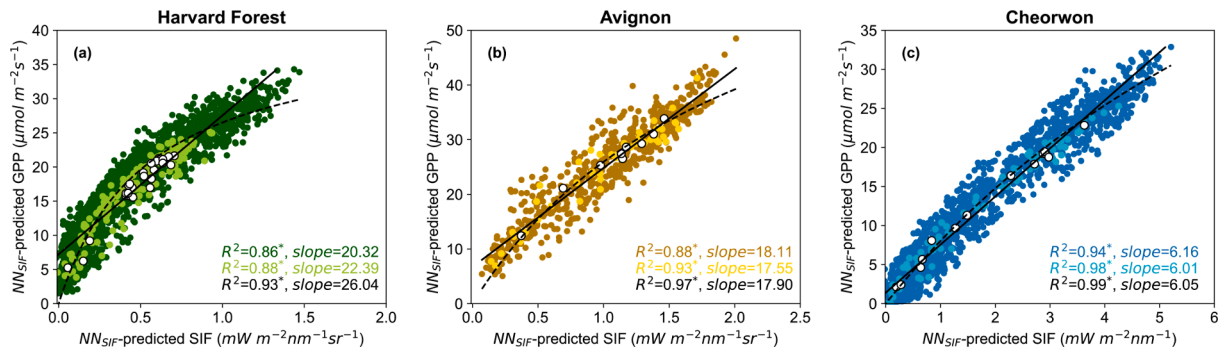

Fig. 8. *GPP-SIF relationships retrieved by NN_{SIF} at (a) Harvard Forest, (b) Avignon, and (c) Cheorwon during the growing season.* Hyperbolic (dashed lines) and linear models (solid lines) are used to fit the (half-)hourly relationship. The superscript '*' indicates statistical significance ($p < 0.05$). The R^2 of hyperbolic models is 0.87, 0.87, and 0.89 for Harvard Forest, Avignon, and Cheorwon respectively. Data points, R^2, and slope values are annotated in different colors for different time scales. White points and black fonts represent the weekly scale. Light-colored points and fonts (light green, light yellow, and light blue) represent the daily scale. Dark-colored points and fonts (dark green, dark yellow, and dark blue) represent the (half-)hourly scale. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

this is the first study that exploits the SIF potential to partition NEE, without imposing any empirical assumptions on the GPP-SIF relationship.

4.1. The role of SIF in CO_2 flux partitioning

One direct effect of the SIF inclusion is to improve the GPP estimation. Compared to other methods, NN_{SIF} can better capture the rapid GPP variations following harvest (Fig. 4) due to two reasons. First, NN_{SIF} incorporates the observed SIF which is directly informative on the rapid variations of vegetation. The second reason is related to the model structure by which SIF is incorporated in NN_{SIF}. Here, SIF is not included as an input for GPP estimation, but as an extra sub-module and also as one part of the loss function, which can more directly affect the NN

optimization by joining the backpropagation process, i.e., the errors between SIF predictions and observations are used to update the NN weights. Indeed, we also improve NN_{noSIF} by including SIF as an additional input for GPP estimation. The new NN_{noSIF} shows improved skills but is still less sensitive than NN_{SIF} in detecting the sharp drop, especially on the harvest day (DOY 248) (Fig. S15 (d)), suggesting that the model structure of NN_{SIF} can better exploit the SIF information to constrain the GPP estimation. As for standard methods, the relatively coarse resolutions of the typical moving windows (NT uses 15-day windows, DT uses a combination of 4-day and 12-day windows) limit their ability to capture the instantaneous effect of harvest. However, this issue can be potentially alleviated in the future, by fitting parameters before and after the event separately.

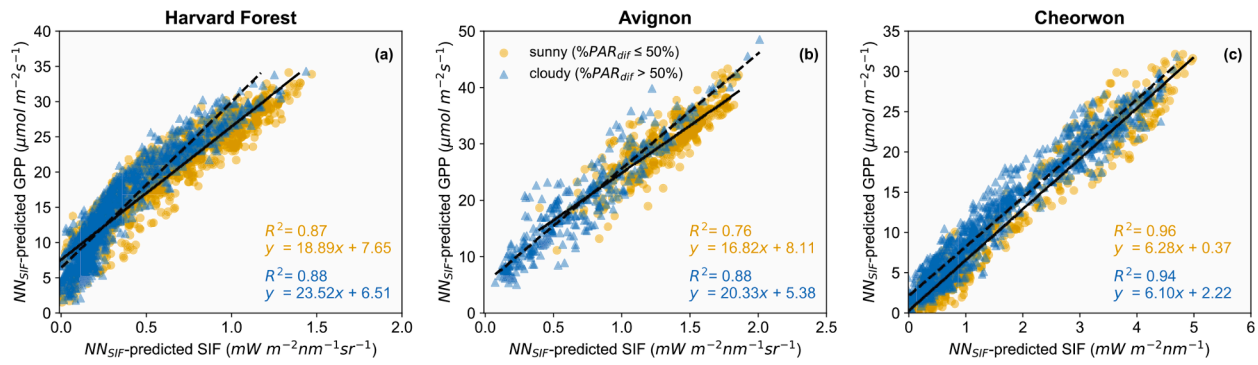Another indirect but important effect of the SIF inclusion is to

Fig. 9. *The (half-)hourly GPP-SIF relationship derived by NN_{SIF} on sunny and cloudy days during the growing season at (a) Harvard Forest, (b) Avignon, and (c) Cheorwon.* Cloudy days are defined as daily mean%PAR_{dif} > 50%. Linear models are fitted for sunny (solid lines) and cloudy (dashed lines) days separately.
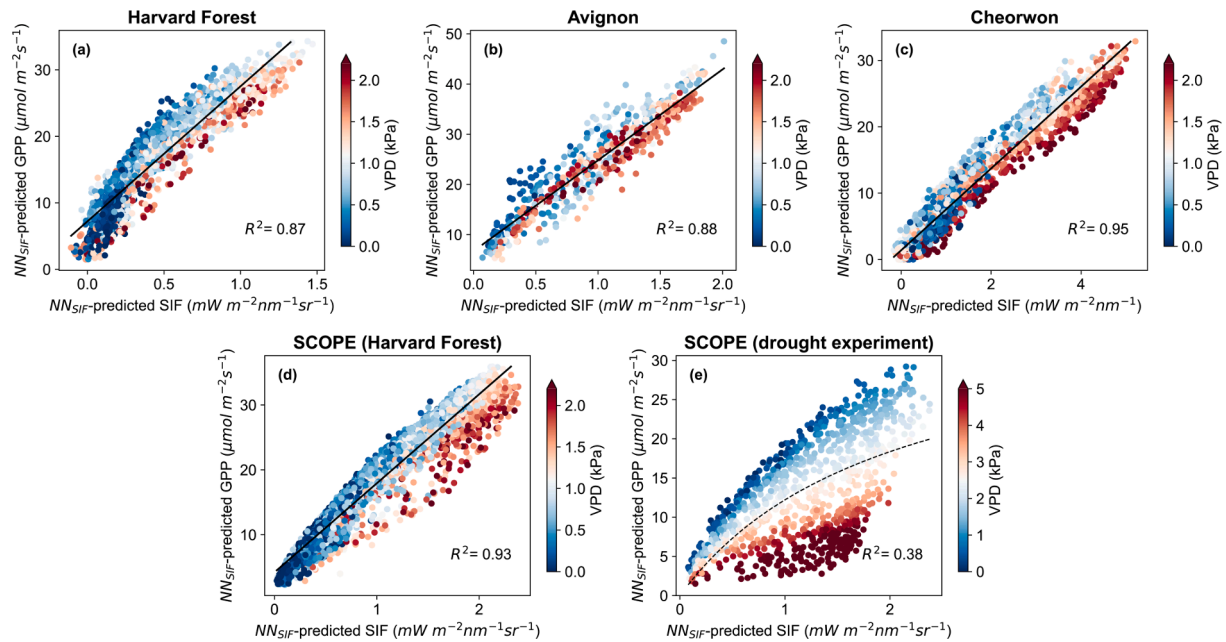

Fig. 10. *The (half-)hourly GPP-SIF relationships derived by NN_{SIF} under different VPD conditions during the growing season.* The colormap in (a-e) indicates the changing VPD. **(a–c)** show NN_{SIF} results obtained from *in-situ* data at **(a)** Harvard Forest, **(b)** Avignon, and **(c)** Cheorwon. **(d)** The NN_{SIF} results obtained from SCOPE simulations at Harvard Forest. **(e)** The NN_{SIF} results obtained from SCOPE simulations under extremely drought conditions. Solid lines denote the fitted linear models, and the dashed line denotes the fitted hyperbolic model.

improve ER estimation, especially during daytime. NN_{SIF} can more accurately retrieve the ER-T_{air} relationship, in particular for high T_{air} (Fig. 6). Both NT and DT assume that the fitted parameters (E_0 and R_{ref}) and the ER-T_{air} relationship remain invariant through day and night, which is the main reason for their deviations in the ER estimation (Fig. 6 (l)). NT fits E_0 and R_{ref} using night data and then extrapolates to daytime, leading to ER underestimation during daytime and at high T_{air} (Fig. 6 (d, e)). DT fits R_{ref} using daytime data, which alleviates daytime ER biases. However, it applies the daytime R_{ref} to nighttime period, leading to larger ER biases at night and at low T_{air} (Fig. 6 (g, h)). Both NN_{SIF} and NN_{noSIF} do not rely on the prescribed empirical relationships of ER, and their difference is that when processing daytime samples, the GPP retrieval in NN_{SIF} is additionally constrained by daytime SIF observations, and this constraint in turn indirectly affects the ER estimation, especially at high temperatures. NN_{noSIF} underestimates ER at high T_{air} which is outside the night T_{air} range (Fig. 6 (j, k)). This suggests that when there are relatively large differences between daytime and nighttime T_{air} distributions, the only NEE constraint shows weakened skills in accurately estimating GPP and ER. As for NN_{SIF}, SIF indirectly constrains the daytime ER through its constraint on GPP. More

specifically, SIF can better constrain GPP, and then the estimated ER covaries to ensure that ER – GPP = NEE still holds true. The SIF inclusion enables NN_{SIF} to substantially reduce the structural errors of standard methods, and to improve the accuracy of NN_{noSIF}, especially for high T_{air} during daytime. This advantage makes NN_{SIF} promising to investigate the responses of CO_2 fluxes to extreme events.

Another advantage of NN_{SIF} is that it can better capture the effect of water stress on ER. Standard methods indirectly consider the soil moisture effect by adjusting parameters (R_{ref}, E_0) using moving windows, and they can capture the moisture limitation on ER at long time scales (Tramontana et al., 2020). However, for the Avignon site, the observed water-stressed period is relatively abrupt and short (~20 days, Fig. S10). The moving window-based parameterization shows limited skills in capturing the instantaneous effect of such short-term water stress. In particular, within each time window where the fitted parameters cannot be adjusted greatly, NT- and DT-estimated ER would simply increase with T_{air} (Fig. 7 (b, c)), as assumed by the Lloyd and Taylor equation used in NT and DT methods. While for NN_{SIF} and NN_{noSIF}, they directly include the soil moisture as an input, and thus can more effectively capture the decline of ER with T_{air} under water stress (Fig. 7 (a,

d)). Besides, NN_{SIF} improves the accuracy of NN_{noSIF} and avoids the abnormally high ER estimates (Fig. 7 (a, d)). The better performance of NN_{SIF} suggests that the ML approach, combined with SIF information, can better estimate moisture dependences of the NEE partitioning.

4.2. Physiological controls on the ecosystem-scale GPP-SIF relationship

The objectives of NN_{SIF} are not limited to better partitioning NEE, but also to derive the ecosystem-scale GPP-SIF relationship. We believe that NN_{SIF} does not artificially enforce stronger linear relationships between GPP and SIF. This statement can be supported by two examples. First, when tested on synthetic data generated under extremely drought conditions, NN_{SIF} can correctly reproduce the highly non-linear GPP-SIF relationship (Figs. S6 and 10(e)), suggesting that it goes beyond the simple linear GPP-SIF scaling. Second, when tested on *in-situ* data, NN_{SIF} can derive either linear or non-linear GPP-SIF relationships across different ecosystems (Fig. 8). Hence, we believe that the NN_{SIF}-retrieved GPP-SIF relationship is unbiased and informative.

NN_{SIF} can not only retrieve the GPP-SIF relationship, but also unravel how this relationship responds to physiological factors. The GPP/SIF ratio predicted by NN_{SIF} increases at higher%PAR_{dif} conditions (Fig. 9). This illumination effect has also been reported across different ecosystems (Chen et al., 2020; Kim et al., 2021; Li et al., 2020; Magney et al., 2019), which is associated with different responses of LUE and effective fluorescence yield (SIF_{yield}, i.e., $\frac{SIF}{APAR}$). Higher%PAR_{dif} allows for deeper light penetration into the canopy, which enhances the overall canopy LUE and thus increases the GPP/SIF ratio. Additionally, LUE can be enhanced due to the NPQ suppression under lower irradiance of cloudy days (Maguire et al., 2020; Porcar-Castell et al., 2014). Oppositely, canopy NIR reflectance is enhanced by higher%PAR_{dif} around the solar noon (Ryu et al., 2010), which likely increases f_{esc} according to the empirical f_{esc} equation (Zeng et al., 2019). This can further increase SIF_{yield} and decrease the GPP/SIF ratio. Here, NN_{SIF} shows that LUE responses dominate over SIF_{yield} responses, and the GPP/SIF ratio increases on cloudy days.

The NN_{SIF}-retrieved GPP/SIF ratio exhibits a significant decreasing trend with the increasing VPD (Fig. 10). Similarly, the Cheorwon *in-situ* study reported such responses to RH (Yang et al., 2018a). This moisture effect is associated with the stomatal responses to atmospheric dryness. Under water stress, plants decrease the stomatal conductance to alleviate water loss, thereby decreasing GPP and LUE (Ball et al., 1987; Collatz et al., 1991; Katul et al., 2010; Medlyn et al., 2011). In parallel, SIF_{yield} is relatively insensitive to stomatal variations and shows only moderate changes (Gu et al., 2019; He et al., 2020). As a result, the GPP/SIF ratio decreases under water stress. This stomatal effect has been reported by *in-situ* studies at the leaf, canopy, and ecosystem scales (Flexas et al., 2002; Helm et al., 2020; Marrs et al., 2020; Wohlfahrt et al., 2018), suggesting that the linear GPP-SIF scaling, as typically imposed in satellite studies, may collapse under water stress. Here, NN_{SIF} can effectively retrieve the moisture dependence of the GPP-SIF relationship, whether tested under normal climates or extremely drought conditions (Fig. 10 (a–e)).

4.3. Uncertainties and advantages of this study

The uncertainties from the input data (e.g., SIF) potentially affect the accuracy of the NN_{SIF} model. Hence, we impose noises with increasing levels to the SIF and NEE in the SCOPE simulations, and examine their effects on the NN_{SIF} performance. At the highest noise level, the GPP, NEE, and SIF predictions still show good accuracy, with R^2 higher than 0.9 (Fig. S16). The ER prediction is more susceptible to noises, but still shows acceptable accuracy (R^2=0.6, RMSE=1.1 $\mu mol\ m^{-2}\ s^{-1}$) at the highest noise level (Fig. S16 (a, c)). In particular, we find that NN_{SIF}-predicted SIF and NEE are actually much closer to the "true" values, compared to the original noisy SIF and NEE that are used to train NN_{SIF}

(Fig. S16 (b, d)). This suggests that NN_{SIF} can filter out most of the noises, and extract the meaningful deterministic signal from the highly noisy data. This advantage implies the potential of NN_{SIF} to be implemented at larger spatial scales using satellite SIF, which typically has higher noise levels than field observations used here. Similarly, we also impose increasing Gaussian noises to the input LAI data, and find that NN_{SIF} shows good robustness to the LAI uncertainty (Fig. S17). In addition to the uncertainties stemming from input data, the NN_{SIF} model itself still has structural uncertainty due to the randomness in its weight initializations and in the training process. We quantified the structural uncertainty by training an ensemble of models. The standard deviation in most cases is smaller than 0.6 $\mu mol\ m^{-2}\ s^{-1}$ for GPP (Fig. S18 (b)), and smaller than 0.25 $\mu mol\ m^{-2}\ s^{-1}$ for ER (Fig. S18 (d)), which is lower than the NN_{noSIF} structural uncertainty (Fig. S19), and can be considered negligible compared to the errors due to input noises.

There are still some limitations in NN_{SIF} despite its promising results. First, we think our current model cannot explicitly simulate the light inhibition on ER (Kok effect) (Keenan et al., 2019; Wehr et al., 2016; Wohlfahrt et al., 2005). In the current version of NN_{SIF}, light intensity is not included as a predictor for ER estimation, as the direct inclusion would make the estimated ER replicate the strong diurnal cycle of PAR, which is unrealistic, as there is some equifinality in the entire model when PAR is used for both GPP and ER estimation. Future studies could include the radiation effect via a more subtle way, so as to further exploit the SIF potential in detecting the Kok effect. Second, although NN_{SIF} is robust to the random Gaussian noises of LAI data, the model performance could still be potentially compromised if the LAI bias is varying seasonally, e.g., the maximum of the estimated GPP and ER could be flawed if the LAI bias only occurs during the peak growing season. Also, some ecosystems have different phenological changes between overstory and understory species. In this case, the satellite LAI potentially biases the estimated GPP and ER of the whole ecosystem, as it mainly reflects the phenological changes of the overstory. Hence, the quality of LAI data should be carefully examined when applying NN_{SIF} in future studies, especially for sites with heterogeneous species. Finally, although NN_{SIF} outperforms other methods when testing on SCOPE simulations, the quality of these results is dependent on how realistic our synthetic simulations are in representing the functional relationships of the real-world ecosystem (e.g., GPP-SIF and ER-T_{air} relationships). Also, among the tested methods, NN_{SIF} has the structure most similar to SCOPE as they both simultaneously simulate GPP and SIF. It requires further investigation to evaluate whether NN_{SIF} holds similar outperformance when tested on other ecosystem models.

A recent study also used SIF observations to partition NEE (Kira et al., 2021). However, there are distinct differences between that method and NN_{SIF}. Firstly, similar to standard partitioning methods, their method also relies on empirical models which have structural errors, while NN_{SIF} learns purely from data without empirical models. This can be a major advantage during extreme events and human disturbances where these empirical models could collapse. Secondly, contrary to Kira et al. (2021), we imposed no empirical formulations on the GPP-SIF relationship, but let NN_{SIF} learn this relationship and its physiological responses. These learned functions can be then leveraged with satellite SIF to achieve a better global-scale GPP estimation.

5. Conclusions

In this study, we developed a novel framework, NN_{SIF}, to tackle two problems simultaneously: (1) partitioning NEE, while (2) learning the GPP-SIF relationship at the ecosystem scale. NN_{SIF} was compared against standard partitioning methods and NN_{noSIF}, using synthetic and field data. Due to the SIF inclusion, NN_{SIF} shows better performance in four aspects: (1) it more accurately retrieves the ER-T_{air} relationship, especially at high T_{air}, (2) it more effectively captures the moisture limitation on ER, (3) it better estimates the LUE variations to stress and

(4) it uniquely captures the sudden GPP drop after harvest. Furthermore, NN$_{SIF}$ can retrieve the ecosystem-scale GPP-SIF relationship, and elucidate how this relationship responds to radiation and humidity.

As far as we know, this is the first effort to exploit the potential of SIF for NEE partitioning, without an empirical formulation of the GPP-SIF relationship. The learned insights are expected to be useful for achieving a better global-scale GPP estimation with satellite SIF. The superior performance of NN$_{SIF}$ suggests that the ML approach, combined with SIF information, holds promising ability to solve the longstanding NEE partitioning problem. Future studies could further exploit the power of ML and SIF by incorporating physical knowledge, e.g., using a hybrid physics-machine learning approach (Reichstein et al., 2019), in order to achieve more accurate NEE partitioning.

Data availability statement

The eddy-covariance and SIF observations of the Harvard Forest site are available from the Harvard Forest Data Archive (https://harvardforest.fas.harvard.edu/harvard-forest-data-archive). The Cheorwon site observations are public on the FLUXNET website (https://fluxnet.org/doi/FLUXNET-CH4/KR-CRK). The Avignon site observations are available through European Fluxes Database Cluster (http://www.europe-fluxdata.eu/). The SCOPE simulations and the codes for NEE partitioning are available at GitHub (https://github.com/Weiwei047/SIF_partitioning) or at zenodo (https://zenodo.org/record/6450447#.YlTl9NPMLtN).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.agrformet.2022.108980.

References

Aubinet, M., Vesala, T., Papale, D., 2012. Eddy covariance: a Practical Guide to Measurement and Data Analysis. Springer Science & Business Media.

Badgley, G., Field, C.B., Berry, J.A., 2017. Canopy near-infrared reflectance and terrestrial photosynthesis. Sci. Adv. 3 (3), e1602244.

Baker, N.R., 2008. Chlorophyll fluorescence: a probe of photosynthesis *in vivo*. Annu. Rev. Plant Biol. 59, 89–113.

Ball, J.T., Woodrow, I.E., Berry, J.A., 1987. A Model Predicting Stomatal Conductance and Its Contribution to the Control of Photosynthesis Under Different Environmental conditions, Progress in Photosynthesis Research. Springer, pp. 221–224.

Bayat, B., van der Tol, C., Yang, P., Verhoef, W., 2019. Extending the SCOPE model to combine optical reflectance and soil moisture observations for remote sensing of ecosystem functioning under water stress conditions. Remote Sens. Environ. 221, 286–301.

Caldeira, J., Nord, B., 2020. Deeply uncertain: comparing methods of uncertainty quantification in deep learning algorithms. Mach. Learn. Sci. Technol. 2 (1), 015002.

Chen, J., Liu, X., Du, S., Ma, Y., Liu, L., 2020. Integrating SIF and clearness index to improve maize GPP estimation using continuous tower-based observations. Sensors 20 (9), 2493.

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. Agric. For. Meteorol. 54 (2–4), 107–136.

Daumard, F., et al., 2010. A field platform for continuous measurement of canopy fluorescence. IEEE Trans. Geosci. Remote Sens. 48 (9), 3358–3368.

Dechant, B., et al., 2020. Canopy structure explains the relationship between photosynthesis and sun-induced chlorophyll fluorescence in crops. Remote Sens. Environ. 241, 111733.

Dechant, B., Ryu, Y., Kang, M., 2019. Making full use of hyperspectral data for gross primary productivity estimation with multivariate regression: mechanistic insights from observations and process-based simulations. Remote Sens. Environ. 234, 111435.

Desai, A.R., et al., 2008. Cross-site evaluation of eddy covariance GPP and RE decomposition techniques. Agric. For. Meteorol. 148 (6–7), 821–838.

Flexas, J., et al., 2002. Steady-state chlorophyll fluorescence (Fs) measurements as a tool to follow variations of net CO2 assimilation and stomatal conductance during water-stress in C$_3$ plants. Physiol. Plant. 114 (2), 231–240.

Frankenberg, C., et al., 2011. New global observations of the terrestrial carbon cycle from GOSAT: patterns of plant fluorescence with gross primary productivity. Geophys. Res. Lett. 38 (17), L17706.

Gentine, P., Entekhabi, D., Chehbouni, A., Boulet, G., Duchemin, B., 2007. Analysis of evaporative fraction diurnal behaviour. Agric. For. Meteorol. 143 (1–2), 13–29.

Gentine, P., Entekhabi, D., Polcher, J., 2011. The diurnal behavior of evaporative fraction in the soil–vegetation–atmospheric boundary layer continuum. J. Hydrometeorol. 12 (6), 1530–1546.

Gentine, P., et al., 2019. Land–atmosphere interactions in the tropics–a review. Hydrol. Earth Syst. Sci. 23 (10), 4171–4197.

Goulas, Y., et al., 2017. Gross primary production of a wheat canopy relates stronger to far red than to red solar-induced chlorophyll fluorescence. Remote Sens. 9 (1), 97. Basel.

Gu, L., Han, J., Wood, J.D., Chang, C.Y.Y., Sun, Y., 2019. Sun-induced Chl fluorescence and its importance for biophysical modeling of photosynthesis based on light reactions. New Phytol. 223 (3), 1179–1191.

Guan, K., et al., 2016. Improving the monitoring of crop productivity using spaceborne solar-induced fluorescence. Glob. Chang. Biol. 22 (2), 716–726.

Guanter, L., et al., 2014. Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. Proc. Natl. Acad. Sci. 111 (14), E1327–E1333.

He, L., et al., 2020. Tracking seasonal and interannual variability in photosynthetic downregulation in response to water stress at a temperate deciduous forest. J. Geophys. Res. Biogeosci. 125 (8), e2018JG005002.

Helm, L.T., Shi, H., Lerdau, M.T., Yang, X., 2020. Solar-induced chlorophyll fluorescence and short-term photosynthetic response to drought. Ecol. Appl.

L. Hertel, J. Collado, P. Sadowski and P. Baldi, 2018. Sherpa: hyperparameter optimization for machine learning models.

Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Netw. 2 (5), 359–366.

Hwang, Y., et al., 2020. Comprehensive assessments of carbon dynamics in an intermittently-irrigated rice paddy. Agric. For. Meteorol. 285, 107933.

Joiner, J., et al., 2013. Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: methodology, simulations, and application to GOME-2. Atmos. Meas. Tech. 6 (10), 2803–2823.

Joiner, J., Yoshida, Y., Vasilkov, A., Middleton, E., 2011. First observations of global and seasonal terrestrial chlorophyll fluorescence from space. Biogeosciences 8 (3), 637.

Katul, G., Manzoni, S., Palmroth, S., Oren, R., 2010. A stomatal optimization theory to describe the effects of atmospheric CO2 on leaf photosynthesis and transpiration. Ann. Bot. 105 (3), 431–442.

Keenan, T.F., et al., 2019. Widespread inhibition of daytime ecosystem respiration. Nat. Ecol. Evol. 3 (3), 407–415.

Kendall, A., Gal, Y., Cipolla, R., 2018. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7482–7491.

Kim, J., et al., 2021. Solar-induced chlorophyll fluorescence is non-linearly related to canopy photosynthesis in a temperate evergreen needleleaf forest during the fall transition. Remote Sens. Environ. 258, 112362.

Kim, J., Ryu, Y., Jiang, C., Hwang, Y., 2019. Continuous observation of vegetation canopy dynamics using an integrated low-cost, near-surface remote sensing system. Agric. For. Meteorol. 264, 164–177.

D.P. Kingma and J. Ba, 2014. Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 https://doi.org/10.48550/arXiv.1412.6980.

Kira, O., et al., 2021. Partitioning net ecosystem exchange (NEE) of CO2 using Solar-Induced chlorophyll Fluorescence (SIF). Geophys. Res. Lett. 48, e2020GL091247 e2020GL091247.

Lasslop, G., et al., 2012. On the choice of the driving temperature for eddy-covariance carbon dioxide flux partitioning. Biogeosciences 9 (12), 5243–5259.

Lasslop, G., Reichstein, M., Kattge, J., Papale, D., 2008. Influences of observation errors in eddy flux data on inverse model parameter estimation. Biogeosciences 5 (5), 1311–1324.

Lasslop, G., et al., 2010. Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation. Glob. Chang. Biol. 16 (1), 187–208.

Li, Z., et al., 2020. Solar-induced chlorophyll fluorescence and its link to canopy photosynthesis in maize from continuous ground measurements. Remote Sens. Environ. 236, 111420.

Lloyd, J., Taylor, J., 1994. On the temperature dependence of soil respiration. Funct. Ecol. 8, 315–323.

Magney, T.S., et al., 2019. Mechanistic evidence for tracking the seasonality of photosynthesis with solar-induced fluorescence. Proc. Natl. Acad. Sci. 116 (24), 11640–11645.

Magney, T.S., et al., 2017. Connecting active to passive fluorescence with photosynthesis: a method for evaluating remote sensing measurements of Chl fluorescence. New Phytol. 215 (4), 1594–1608.

Maguire, A.J., et al., 2020. On the functional relationship between fluorescence and photochemical yields in complex evergreen needleleaf canopies. Geophys. Res. Lett. 47 (9), e2020GL087858 e2020GL087858.

Marrs, J., et al., 2020. Solar-induced fluorescence does not track photosynthetic carbon assimilation following induced stomatal closure. Geophys. Res. Lett. 47 (15), e2020GL087956 e2020GL087956.

McNicol, I.M., Ryan, C.M., Mitchard, E.T., 2018. Carbon losses from deforestation and widespread degradation offset by extensive growth in African woodlands. Nat. Commun. 9 (1), 1–11.

Medlyn, B.E., et al., 2011. Reconciling the optimal and empirical approaches to modelling stomatal conductance. Glob. Chang. Biol. 17 (6), 2134–2144.

Meroni, M., et al., 2009. Remote sensing of solar-induced chlorophyll fluorescence: review of methods and applications. Remote Sens. Environ. 113 (10), 2037–2051.

Migliavacca, M., et al., 2011. Semiempirical modeling of abiotic and biotic factors controlling ecosystem respiration across eddy covariance sites. Glob. Chang. Biol. 17 (1), 390–409.

Mohammed, G.H., et al., 2019. Remote sensing of solar-induced chlorophyll fluorescence (SIF) in vegetation: 50 years of progress. Remote Sens. Environ. 231, 111177.

Munger, W., Wofsy, S., 2021. Canopy-atmosphere exchange of carbon, water and energy at Harvard forest EMS tower since 1991. Environ. Data Initiat. https://doi.org/10.6073/pasta/6e1e3d902387781c1d9822cce8444ede (Accessed 2021-09-23).

Novick, K.A., et al., 2016. The increasing importance of atmospheric demand for ecosystem water and carbon fluxes. Nat. Clim. Chang. 6 (11), 1023–1027.

Oikawa, P., et al., 2017. Revisiting the partitioning of net ecosystem exchange of CO_2 into photosynthesis and respiration with simultaneous flux measurements of $13CO_2$ and CO_2, soil respiration and a biophysical model, CANVEG. Agric. For. Meteorol. 234, 149–163.

Papale, D., et al., 2006. Towards a standardized processing of net ecosystem exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences 3 (4), 571–583.

Porcar-Castell, A., et al., 2014. Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges. J. Exp. Bot. 65 (15), 4065–4095.

Reichstein, M., et al., 2019. Deep learning and process understanding for data-driven Earth system science. Nature 566 (7743), 195–204.

Reichstein, M., et al., 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Glob. Chang. Biol. 11 (9), 1424–1439.

Reichstein, M., et al., 2003. Modeling temporal and large-scale spatial variability of soil respiration from soil water availability, temperature and vegetation productivity indices. Glob. Biogeochem. Cycles 17 (4), 1104–1118.

Riederer, M., Hübner, J., Ruppert, J., Brand, W.A., Foken, T., 2014. Prerequisites for application of hyperbolic relaxed eddy accumulation on managed grasslands and alternative net ecosystem exchange flux partitioning. Atmos. Meas. Tech. 7 (12), 4237–4250.

Ryu, Y., et al., 2010. Testing the performance of a novel spectral reflectance sensor, built with light emitting diodes (LEDs), to monitor ecosystem metabolism, structure and function. Agric. For. Meteorol. 150 (12), 1597–1606.

Ryu, Y., Berry, J.A., Baldocchi, D.D., 2019. What is global photosynthesis? History, uncertainties and opportunities. Remote Sens. Environ. 223, 95–114.

Schimel, D., et al., 2015. Observing terrestrial ecosystems and the carbon cycle from space. Glob. Chang. Biol. 21 (5), 1762–1776.

Spielmann, F., et al., 2019. Gross primary productivity of four European ecosystems constrained by joint CO_2 and COS flux measurements. Geophys. Res. Lett. 46 (10), 5284–5293.

Stoy, P.C., et al., 2006. An evaluation of models for partitioning eddy covariance-measured net ecosystem exchange into photosynthesis and respiration. Agric. For. Meteorol. 141 (1), 2–18.

Sun, Y., et al., 2018. Overview of solar-induced chlorophyll fluorescence (SIF) from the orbiting carbon observatory-2: retrieval, cross-mission comparison, and global monitoring for GPP. Remote Sens. Environ. 209, 808–823.

Sun, Y., et al., 2017. OCO-2 advances photosynthesis observation from space via solar-induced chlorophyll fluorescence. Science 358 (6360), eaam5747.

Tramontana, G., et al., 2020. Partitioning net carbon dioxide fluxes into photosynthesis and respiration using neural networks. Glob. Chang. Biol. 26 (9), 5235–5253.

Van der Tol, C., Berry, J., Campbell, P., Rascher, U., 2014. Models of fluorescence and photosynthesis for interpreting measurements of solar-induced chlorophyll fluorescence. J. Geophys. Res. Biogeosci. 119 (12), 2312–2327.

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. https://doi.org/10.5194/bg-6-3109-2009.

Wehr, R., et al., 2016. Seasonality of temperate forest photosynthesis and daytime respiration. Nature 534 (7609), 680–683.

Wohlfahrt, G., Bahn, M., Haslwanter, A., Newesely, C., Cernusca, A., 2005. Estimation of daytime ecosystem respiration to determine gross primary production of a mountain meadow. Agric. For. Meteorol. 130 (1–2), 13–25.

Wohlfahrt, G., Galvagno, M., 2017. Revisiting the choice of the driving temperature for eddy covariance CO_2 flux partitioning. Agric. For. Meteorol. 237, 135–142.

Wohlfahrt, G., et al., 2018. Sun-induced fluorescence and gross primary productivity during a heat wave. Sci. Rep. 8 (1), 1–9.

Wohlfahrt, G., Gu, L., 2015. The many meanings of gross photosynthesis and their implication for photosynthesis research from leaf to globe. Plant Cell Environ. 38 (12), 2500.

Wutzler, T., et al., 2018. Basic and extensible post-processing of eddy covariance flux data with REddyProc. Biogeosciences 15 (16), 5015–5030.

Yang, K., et al., 2018a. Sun-induced chlorophyll fluorescence is more strongly related to absorbed light than to photosynthesis at half-hourly resolution in a rice paddy. Remote Sens. Environ. 216, 658–673.

Yang, X., et al., 2018b. FluoSpec 2-an automated field spectroscopy system to monitor canopy solar-induced fluorescence. Sensors 18 (7), 2063.

Yang, X., et al., 2015. Solar-induced chlorophyll fluorescence that correlates with canopy photosynthesis on diurnal and seasonal scales in a temperate deciduous forest. Geophys. Res. Lett. 42 (8), 2977–2987.

Zeng, Y., et al., 2019. A practical approach for estimating the escape ratio of near-infrared solar-induced chlorophyll fluorescence. Remote Sens. Environ. 232, 111209.

Zhang, Q., et al., 2018. Changes in photosynthesis and soil moisture drive the seasonal soil respiration-temperature hysteresis relationship. Agric. For. Meteorol. 259, 184–195.

Zhou, S., et al., 2019a. Land–atmosphere feedbacks exacerbate concurrent soil drought and atmospheric aridity. Proc. Natl. Acad. Sci. 116 (38), 18848–18853.

Zhou, S., Zhang, Y., Williams, A.P., Gentine, P., 2019b. Projected increases in intensity, frequency, and terrestrial carbon costs of compound drought and aridity events. Sci. Adv. 5 (1), eaau5740.

Zscheischler, J., et al., 2014. A few extreme events dominate global interannual variability in gross primary production. Environ. Res. Lett. 9 (3), 035001.