Constraining respiration flux and carbon pools in a simple ecosystem carbon model

Olya Skulovich[1], Caroline A. Famiglietti[2], Alexandra G. Konings[2], Pierre Gentine[1]

[1]Earth and Environmental Engineering, Columbia University, New York, NY 10027, USA

[2]Department of Earth System Science, Stanford University, Stanford, USA

Key Points:

- Assimilation of globally available data like solar-induced fluorescence and vegetation optical depth improves model results.
- The assimilated data helps constrain the fluxes and pools (e.g. soil) that are not directly observed.
- The effect of the additional data depends on the site-level conditions, data quality, and representation in the model.

Corresponding author: Olya Skulovich, os2328@columbia.edu

Abstract

Incorporating observational data in carbon-cycle models provides a systematic framework for understanding complex ecosystem carbon dynamics, contributing essential insights for climate change mitigation and land ability to continue acting as a carbon sink. This study addresses the challenge of accurately quantifying carbon fluxes and pools, focusing on the information content of remote sensing observations. The research explores the impact of assimilating multiple observational datasets into the CARbon DAta MOdel fraMework (CARDAMOM). Satellite observations such as solar-induced fluorescence (SIF) and vegetation optical depth (VOD) are used as proxies for photosynthesis and aboveground biomass, respectively. The study aims to answer key questions about the reliability of remote sensing data in constraining the ecosystem respiration flux and sizes and dynamics of carbon pools and the relative usefulness of SIF and VOD across five FLUXNET sites. We conclude that assimilating remote SIF and VOD instead of site-based net ecosystem exchange did not deteriorate and even improved model predictions for all metrics except for interannual variability. Notably, the improved results correspond to a consistent shift in values for crucial model parameters across all five investigated sites.

Plain Language Summary

Carbon-cycle models allow us to study how terrestrial ecosystems absorb carbon from the atmosphere and release it back and what its afterlife is in plants and soil. Scientists use observed data to accurately quantify these processes and incorporate them into models to constrain model parameters. We study how two satellite-based measurements, one used to substitute for photosynthesis, which is not measurable directly, and one used to substitute for aboveground biomass, which is scarcely available, help to improve the model's performance.

1 Introduction

Accurately quantifying terrestrial carbon sinks/sources and carbon pools is critical for reliable projections of carbon emissions and climate change mitigation, yet large uncertainties still exist among the components of the global carbon budget (Arneth et al., 2017; Piao et al., 2018; Gasser et al., 2020; Friedlingstein et al., 2022). Carbon is taken up by the terrestrial biosphere through photosynthesis and released via autotrophic and

heterotrophic respiration as well as disturbances (wildfires, windthrows). Yet, how those fluxes and their partitioning will change under elevated CO_2 remains a question of debate (e.g., Xu et al., 2015; Kirschbaum & McMillan, 2018; L. Liu et al., 2020). To tackle this challenge, it is necessary to understand the potentially competing processes that affect carbon uptake under global change, such as plants' physiological response to elevated CO_2, higher temperatures, increased aridity, increased extreme events frequency, and other climatological shifts associated with climate change on plant-level and ecosystem-level scales (Cox et al., 2000; Tharammal et al., 2019; Song et al., 2019; Denissen et al., 2022). Nevertheless, large amounts of global and site-level observational data have become available and can now be used to constrain many of those processes. Over the last couple of decades, our understanding of photosynthesis and our capacity to constrain it at the global scale has dramatically increased through the use, first, of vegetation indices, (e.g., Wu et al., 2009) and then more recently of Solar Induced Fluorescence or SIF, (e.g., X. Yang et al., 2015; Zhang et al., 2014), a proxy for gross primary productivity (GPP). However, quantifying global respiration fluxes is currently impossible through direct measurements, and they can only be indirectly inferred, such as using statistical upscaling from local measurements or using process-based models, (e.g., Jian et al., 2018; Jung et al., 2019; Nathaniel et al., 2023). In essence, from an observational standpoint, there are no global constraints on the respiration part of the land carbon budget, leading to major uncertainties in our capacity to understand and predict the terrestrial carbon cycle.

Process-based models allow us to combine knowledge of physical, chemical, and biological processes with the collected data to achieve interpretable carbon cycle analysis. Process-based models depict ecological processes with models of varying complexity and different level of abstraction. Yet, while those models are good at capturing some aspects of the system, they rely on several structural assumptions, and the model parameters should be carefully calibrated to improve the model accuracy (Y.-P. Wang et al., 2009). For example, Li et al. (2021) lists model structure and model assumptions uncertainties among the main processes contributing to the overall model uncertainties. These uncertainties are due to our incomplete understanding of some ecological mechanisms (for instance, belowground processes and microbial interactions, (e.g., Hartmann et al., 2020)), an abundance of empirical equations with parameters that are not necessarily applicable globally, and model-specific simplifications. Model parameters can be calibrated via data assimilation (also called "model-data fusion") to best match observational data,

including data uncertainty to quantify model parametric and prediction uncertainties, (e.g., Li et al., 2020). Models can even constrain processes that are not directly observed (Talagrand, 1997) because of the internal physical and biological constraints linking observed to unobserved variables (e.g., mass balance for carbon pools). In other words, since these models quantify internal system interconnections and dynamics, constraints introduced via observational data are propagated through the model and can constrain the rest of unobservable fluxes or pools. As such, unobserved fluxes and pools, such as respiration and soil carbon, can potentially be constrained through the assimilation of observable variables that are indirectly related to those processes.

A caveat of the data assimilation approach lies in the tradeoff between model complexity and the demand for data to be assimilated (e.g., Famiglietti et al., 2021). Indeed, the more complex the models are, the more parameters are needed for their description. Then, the more parameters the model comprises, the more likely it is that the model output will match equally well the observations with *different* combinations of the parameters. This phenomenon is called equifinality (Beven, 1993; Beven & Freer, 2001). Simply put, it means that the model can give the right answer (i.e., is optimized and corresponds well with the observational data) but for the wrong reason (i.e., the resulting combination of the model parameter is inconsistent with the *true* system dynamic). Generally speaking, equifinality is reduced when more data is assimilated to constrain multiple different components of the model. In practice, assimilating multiple datasets can be limited by data availability and the technical complexity of the data assimilation process.

In this work, we study the effects of assimilating multiple observational datasets in an ecosystem carbon model to quantify their impact on constraining unobserved respiration flux and carbon pools. The physical models' equations constrain the relationship between non-observed variables and observed variables, which are assimilated. The model used for this data assimilation framework is the CARbon DAta MOdel fraMework model or CARDAMOM for short (A. Bloom & Williams, 2015; A. A. Bloom et al., 2016). This model is actively used by the scientific community to model the terrestrial carbon cycle for process understanding and has been proven to successfully capture spatial patterns and temporal trends, as well as inter-annual variability of the various variables of interest (A. A. Bloom et al., 2018; López-Blanco et al., 2019; Quetin et al., 2019; Y. Yang et al., 2019; Yin et al., 2020; A. Norton et al., 2021), including in a benchmarking effort

across more than 200 eddy-covariance sites (Y. Yang et al., 2022). The model has been previously applied to study global and local mechanisms for diverse ecosystems, from Arctic (López-Blanco et al., 2019) to tropics (Yin et al., 2020). CARDAMOM was created as a relatively simple model so that the low number of the model parameters could be better constrained by observational data and thus reduce the risk of equifinality. Model complexity and equifinality in the context of CARDAMOM's predictive skill were previously examined by Famiglietti et al. (2021) and concluded that the model skill depends on properly constraining the model parameters. Building on that, we look into the model skill when the parameters and variables of interest are constrained by indirect observations. In particular, we use SIF as a proxy for GPP, and vegetation optical depth (VOD) as a proxy for above-ground biomass and use physical constraints to build additional constraints on belowground carbon pools and ecosystem respiration. VOD is a variable derived from remote sensing observations as an attenuation of the surface reflectance by the wet biomass and can serve as a measure of total biomass and water-related vegetation stress (Konings et al., 2017).

When investigating the effects of assimilating additional datasets in CARDAMOM, we employ a multi-objective approach and examine the impact of different model formulations, varying data quality (including information content and data uncertainty), and implicit and explicit constraints introduced in the model. We investigate the following research questions: 1) Given the error-prone nature of NBE at the global scale (e.g., Deng & Chen, 2011; Chevallier et al., 2019; Cui et al., 2022), how well can remote sensing data such as SIF and VOD help constrain the carbon cycle instead? 2) Which of the SIF and VOD data is more useful in this task, and how does it depend on local conditions? 3) Can respiration flux and carbon pools be reliably inferred from data assimilation, given indirect observational constraints? 4) What is the role of data availability and assumed uncertainty on the assimilation results? The analysis is conducted for five FLUXNET sites across different biomes.

2 Methods

The CARDAMOM framework consists of two major parts: the carbon cycle model and the data assimilation infrastructure. Carbon cycle models have different "flavors" and are called DALEC (Data Assimilation Linked Ecosystem Carbon model) versions that vary depending on the physical processes and parameters included (e.g., with/without

fires, with/without water cycle, etc. A. Bloom and Williams (2015); Quetin et al. (2020)). The general model structure is relatively simple and, in most of the DALECs, includes six carbon pools: foliar, labile, wood, fine roots, litter, and soil organic matter (SOM). An additional advantage of the CARDAMOM framework is the inclusion of a series of ecological and dynamic constraints (EDC) on model parameters and initial conditions (A. Bloom & Williams, 2015). These "common sense" constraints drive ecosystem variables towards more consistent and realistic solutions, thus further helping to reduce equifinality. The optimized model parameters and initial conditions are time-invariant and site-specific. They are inferred using a Metropolis–Hastings Markov chain Monte Carlo (MCMC) approach (Haario et al., 2001). The model is run at monthly resolution.

To test whether we can better constrain respiration fluxes and carbon pools by assimilating SIF and VOD, we first take a medium-complexity version of the DALEC model (for a detailed description of the model, refer to A. Bloom and Williams (2015) and Famiglietti et al. (2021)). In this configuration, 33 model parameters and initial conditions are being optimized via data assimilation. Site-specific meteorological data (air temperature, shortwave radiation, atmospheric CO_2 concentration, vapor pressure deficit, precipitation, and wind speed) drive the model dynamics, while observational data further constrain model parameters. In particular, eddy covariance net ecosystem exchange (NEE) measurements from FLUXNET (Pastorello et al., 2020), leaf area index (LAI) estimates from the Copernicus Global Land Service (Fuster et al., 2020), and *in situ* biomass surveys are used for assimilation into the model. The same or similar combination of data for data assimilation has previously been used in CARDAMOM (López-Blanco et al., 2019; Quetin et al., 2020; Famiglietti et al., 2021). In the previously reported configurations, NEE had the highest impact on the model performance (Famiglietti et al., 2021). In CARDAMOM studies that have a regional or global focus, beyond FLUXNET sites, atmospheric inversion of net biosphere exchange (NBE) is used in the assimilation with the caveat that it can have large uncertainties (e.g., H. Wang et al., 2019; Cui et al., 2022). Hence, we investigate how excluding it and including SIF and VOD datasets for data assimilation, as well as other modeling choices and assumptions, affect the model performance.

The analysis is run over several eddy-covariance sites, including Harvard Forest EMS Tower, USA (US-Ha1), Puechabon, France (Fr-Pue), Le Bray, France (Fr-LBr), Hyytiala, Finland (Fi-Hyy), and Howard Springs, Australia (AU-How), see Table 1.

Table 1. Summary of eddy-covariance sites, showing their location, FLUXNET code, observational time period, mean climate information, and ecosystem type. Ecosystem type is denoted using the International Geosphere-Biosphere Programme (IGBP) classification. DBF: deciduous broadleaf forest; EBF: evergreen broadleaf forest; ENF: evergreen needleleaf forest; WSA: woody savanna. Simple aridity index is calculated as De Martonne aridity index (De Martonne, 1923) $AI = \frac{P}{T_a+10}$.

Site code	Lat	Lon	Elevation, m	IGBP	Data record	Mean annual temp, C	Mean annual precip-itation, mm/year	Simple aridity index	Reference
AU-How	-12.49	131.15	42	WSA	2001-2014	27	1449	Moderate humid	(Beringer et al., 2007)
FI-Hyy	61.85	24.29	181	ENF	1999-2014	3.8	709	Very humid	(Suni et al., 2003)
FR-LBr	44.72	-0.77	61	ENF	1998-2008	13.6	900	Moderate humid	(Berbigier et al., 2001)
FR-Pue	43.74	3.59	270	EBF	2000-2014	13.5	883	Moderate humid	(Rambal et al., 2004)
US-Ha1	42.54	-72.17	340	DBF	1998-2012	6.2	1071	Excessive humid	(Munger & Wofsy, 2014)

2.1 Including new observational datasets

To include a new observational dataset in CARDAMOM, the following steps need to be performed: 1) define a functional form that represents the data as a function of existing and new model variables and parameters; 2) add data likelihood to the full model likelihood; 3) define a prior for any new model parameters. We discuss the assimilation of SIF and VOD in the following sections.

2.1.1 SIF model

SIF is assumed to be a linear function of GPP (Wood et al., 2017); hence, in the model, it is included as follows.

$$SIF = p_{sif} F_{GPP} \tag{1}$$

where F_{GPP} is the GPP flux and p_{sif} is a proportionality coefficient $[m^2 s^{-2} \mu m^{-1} sr^{-1}]$.

Synthetic and observational SIF data are considered, with the synthetic SIF defined directly from FLUXNET GPP data for the corresponding FLUXNET site. This

allows assessing the effect of data quality (signal vs. noise) on the model performance. Indeed, by construction, the synthetic SIF data contains the signal from the site-measured GPP with no additional noise related to observational uncertainties, sensor uncertainties, SIF retrieval model uncertainties, as well as the scale mismatch between the site level data and the grid level remote sensing data. The observational SIF, on the other hand, contains the GPP-related signal along with the noise, with the unknown ratio of the two. To which degree that affects the model performance defines the model sensitivity to the data uncertainty.

An average between daytime- GPP_{DT}^{FLUX} and nighttime- GPP_{NT}^{FLUX} derived GPP is considered the site GPP.

$$SIF_{synth} = p_{sif}\left(\frac{GPP_{DT}^{FLUX} + GPP_{NT}^{FLUX}}{2}\right) \tag{2}$$

The slope coefficient p_{sif} is individually fit to the FLUX site GPP data such that the final synthetic SIF data is in the same range as the observational SIF data. GOME-2 SIF data (Joiner et al., 2023) is used as SIF observational data. From GOME-2, both SIF and normalized SIF (normalized by photosynthetically active radiation – PAR) data are used. A similar amplitude of the synthetic and observational data is necessary to assess the effect of the data quality (noisiness) on the model performance and isolate it from other effects. An example data time series is shown in Fig. S5.

The SIF likelihood function is constructed similarly to the likelihood function for other assimilated data in CARDAMOM with the observation ($\boldsymbol{O}$) probability given a set of model parameters $\boldsymbol{x}$ is

$$P(\boldsymbol{O}|\boldsymbol{x}) = e^{-0.5\sum_{i=1}^{n}\frac{\left(M_i^{SIF}-O_i^{SIF}\right)^2}{\sigma_i^2}} \tag{3}$$

where O_i^{SIF} is the ith SIF observation, M_i^{SIF} is the corresponding model SIF, and σ_i is the ith error variance for each observation with no error covariance between observation errors assumed (A. Bloom & Williams, 2015).

2.1.2 VOD model

Vegetation optical depth (VOD), based on microwave remote sensing, provides a constraint on aboveground biomass. Unlike SIF, VOD representation in the model is not as straightforward. VOD is measured as the attenuation of electromagnetic waves, which is proportional to the mass of water in the vegetation (Jackson & Schmugge, 1991; Wigneron et al., 2017; Konings et al., 2017). Hence, VOD reflects vegetation water content that can be representative of both variability in total biomass and plant relative water content that depends on meteorological conditions. There is no clear separation of these two components even at larger-than-daily time scales (Konings, Holtzman, et al., 2021). Considering the overall simplification of the given DALEC and aiming at keeping the number of new parameters minimal to reduce equifinality, we assume VOD at the monthly resolution to be a function of above-ground biomass only. That is, we assume that variations in relative water content are small, which is a reasonable assumption. Relative water content representation is missing in the carbon model, which is not uncommon for a model of such simplicity but more complex models are starting to include plant hydraulics (e.g., Kennedy et al., 2019) that could resolve water content. For simplicity, we additionally assume that VOD is a linear combination of leaf and wood biomass, with different learnable regression parameters of each biomass type, at each site:

$$VOD = p_{fol} C_{fol} + p_{woo} C_{woo} \tag{4}$$

where C_{fol} is the foliar carbon pool, C_{woo} is the wood carbon pool $[gCm^{-2}]$ and p_{fol} and p_{woo} are the corresponding coefficients $[gC^{-1}m^2]$, since VOD is dimensionless.

In addition to simplifying the relationship between VOD and biomass, this representation poses a numerical issue. Indeed, for forests, in units of carbon mass, C_{woo} can be several orders of magnitude larger than C_{fol}. For example, wood biomass averages 13,000 gCm^{-2} while foliar biomass reaches the maximum of 300 gCm^{-2} based on a CARDAMOM prediction for Harvard Forest. At the same time, the VOD dynamic is primarily due to C_{fol} variability because C_{fol} is much more variable than C_{woo} and because VOD is more sensitive to upper canopy layers than to lower canopy layers (Konings, Saatchi, et al., 2021). In CARDAMOM, both the coefficients p_{woo} and p_{fol} and carbon pool sizes are inferred, so the contradicting forces of pools' dynamic and pool' sizes may result in equifinality and a wide range of predicted pools sizes. This, in turn, would lead to no

or minimal additional information introduced by VOD in comparison to LAI (since LAI is defined in CARDAMOM as $LAI = p_{lai}C_{fol}$). There are several ways to tackle this issue, from changing VOD representation in the model to introducing new ecological and dynamic constraints (EDC). Here, we define narrow, not overlapping priors for p_{woo} and p_{fol} ranges that ensure the relationship between the sizes of the pools is preserved.

Similar to the SIF module, synthetic and observational VOD data are considered to assess the effect of data quality (signal vs. noise) on the model performance. Since the true foliar and wood carbon pool dynamic is unknown, we used the following procedure to construct the synthetic VOD data. First, we run CARDAMOM assimilating NBE, LAI, and ABGB and take model output median foliar and wood pools as true pools for that site. With these pools, we fit p_{fol} and p_{woo} coefficients such that the constructed VOD matches the observational data. Two observational datasets are considered – VODCA (Moesinger et al., 2020) C-band VOD and GLAB-VOD (Skulovich et al., 2024) L-band VOD (extracted for the grid cell closest to each site). VOD likelihood function is constructed in the same way as Eq.3.

2.2 Experimental setup

We compare the base case with no data assimilated ('none' case) to the previously reported combination that includes NBE, LAI, and ABGB ('NBE_LAI_ABGB') and then remove NBE and instead add either SIF, VOD, or both SIF and VOD synthetic and observational data (see Table 2). We examine to what extent remote sensing SIF and VOD can replace NBE and lead to similar or better results.

2.3 Information content of observations

When assimilating observational data, the question of the information content of the available data is often reduced to the question of data uncertainty (e.g., Raupach et al., 2005) or data autocorrelation, as an indicator for data redundancy (e.g. Moore et al., 2011; Williams et al., 2009). Due to the overall data scarcity, it is often assumed that assimilating more data will improve model performance. In this series of experiments, aligned with the research question "What is the role of data availability and assumed uncertainty on the assimilation results?" we test this hypothesis by assimilating only a

Table 2. Data assimilation cases

Experiment	**Case name**	**Assimilated data**
1	'none'	-
2	'ABGB_LAI_NBE'	ABGB, LAI, NBE
3	'LAI_ABGB'	ABGB, LAI
4	'+SIF+synth'	ABGB, LAI, synthetic SIF
5	'+SIF+data'	ABGB, LAI, GOME2 SIF
6	'+SIF+data_n'	ABGB, LAI, GOME2 SIF normalized
7	'+VOD_synth'	ABGB, LAI, synthetic VOD
8	'+VOD_data_C'	ABGB, LAI, VODCA VOD
9	'+VOD_data_L'	ABGB, LAI, GLAB-VOD VOD
10	'+SIF+VOD'	ABGB, LAI, synthetic SIF, GLAB-VOD

portion of the available data and compare the results with assimilating data with reduced or increased uncertainty.

In particular, we consider the following cases:

- Comparing assimilating all available LAI, ABGB, SIF, and VOD observations with assimilating just a part of all available observations, namely 20, 40, 60, and 80% of the data points. In each run, the corresponding number of indices are selected at random, each run is repeated 25 times so that each time it is a different 20% of the data that is used in assimilation (for the 20% case, the principle is the same for all percentage values).
- Assimilating 100% of the available LAI, ABGB, SIF, and VOD, and increasing or reducing the data uncertainty for SIF and VOD. We consider doubling, quadrupling, and reducing the uncertainty in half.

Note that 100% of the available data do not necessarily cover 100% of the time steps defined by the forcing meteorological data. For the sites and time period of this study, NBE and LAI data have more coverage than SIF and VOD data (see Table 3).

Table 3. Number of data points and percentage of the available forcing data for each of the assimilated variables per site

	US-Ha1	**FR-Pue**	**FR-LBr**	**AU-How**	**FI-Hyy**
Forcing time series	180 (100%)	180 (100%)	132 (100%)	168(100%)	192(100%)
LAI	151 (84%)	180 (100%)	120 (91%)	150 (89%)	150 (78%)
NBE	156 (87%)	170 (94%)	108 (82%)	134 (80%)	182 (95%)
ABGB	15 (8%)	13 (7%)	8 (6%)	3 (2%)	6 (3%)
VOD (data)	113 (63%)	44 (24%)	41 (31%)	150 (89%)	105 (55%)
SIF (data)	68 (38%)	93 (52%)	22 (17%)	95 (57%)	58 (30%)

When considering the information content coming from the observations, we can hypothesize the following relationships:

- More data is better than less data. However, to what degree additional data improves the model results depends on the local conditions. For instance, some sites can exhibit minimal stress and can be almost fully described by the drivers only, so additional observational data might not improve the results as much as for sites with more complex vegetation feedback.
- Sometimes, the true uncertainty of the observations is unknown. In that case, there should exist an optimal uncertainty that can be used as an empirical parameter in the model. Indeed, if the uncertainty is too high, the observations do not add much information, and if the uncertainty is too low, the model tries to match noise along with the actual signal from observations.

We examine how these hypotheses hold for assimilating satellite observations in CARDAMOM.

2.4 Performance metrics

For every assessed combination of CARDAMOM inputs, model formulations, and parameters, the model runs at least ten times. Within these runs, the convergence is assessed using Gelman-Rubin diagnostics. Next, the model output is compared to FLUXNET eddy-covariance site data, and different model outputs are compared to each other.

- **Flux comparison: R^2, RMSE, HI**. The flux comparison is conducted for GPP and respiration based on eddy-covariance partitioning. Median model output is compared with FLUXNET GPP and RECO derived using friction velocity threshold (VUT, Barr et al. (2013)) method and daytime and nighttime partitioning separately. To estimate uncertainties, the FLUXNET data spread is based on the 25th and 75th percentile of the corresponding parameter. The model and site data are compared using the coefficient of determination R^2 and root mean square error (RMSE).

  These metrics allow us to assess the model accuracy; however, in the Bayesian framework, distribution comparison is more meaningful. To achieve this, a histogram intersection (HI, Famiglietti et al. (2021)) is used as an additional skill metric. HI measures the similarity of two (discretized) distributions, with larger HI corresponding to more similar distributions. HI is calculated for GPP and RECO.
- **Interannual variability**. Through the data assimilation framework, CARDAMOM can capture interannual variability in the fluxes. However, the tightness of the fit to the interannual peaks might not necessarily be clear from metrics like R^2, RMSE, or HI. To assess how different combinations of model parameters affect the interannual variability, we calculate the following metric (IAM - interannual anomaly metric):

$$\mathrm{IAM} = \sum_{i=1}^{n} (V_i - V_i^{seas})^2 \tag{5}$$

  where V_i is the ith observation or model output of a parameter V (for example, GPP), V_i^{seas} is the ith value of the seasonal cycle calculated for this parameter, and i can vary from 1 to n, where n is the total number of observations or model outputs. The seasonal cycle in this approach is identical from year to year. It is calculated by matching a periodic signal, a sine wave with a period of 365.25 days, to the observations, with other parameters of the sine wave fitted using *curve_fit()*

function in Python. Comparison to a sine-wave for a signal with a strong seasonal component is defined to emphasize year-to-year anomalies. For a signal with a weaker seasonality, the fitted sine wave will have a small amplitude and be close to the long-term average. We calculate this metric for the observational data and then compare the value to the metric calculated for different model configurations. With similar R^2 and RMSE, IAM helps to assess another aspect of the model performance. For consistent comparison, we take the ratio $\text{IAM}^{model\_output}/\text{IAM}^{data}$. If this value is close to 1, the model and the data have very similar interannual variability, and if the value is close to 0, the model significantly underestimates interannual variability present in the data.

- **Pool constraints: relative change and biomass-VOD relationship**. Biomass data is mainly constrained through indirect proxies (LAI, VOD, NDVI, etc.) apart from temporarily and spatially sparse surveys. For example, for the sites discussed in this study, ABGB measurements are available only for 2-8% of the total length of the meteorological observations (See Table 3). In addition, ABGB comprises the sum of labile, foliar, root, and wood carbon pools in CARDAMOM without any partitioning information. Hence, we assume the comparison of the model pool dynamic to the ABGB observational data is insufficient. Instead, we compare the relative model performance for different combinations of the assimilated data, answering the question, "Can carbon pools be reliably inferred from data assimilation, given indirect observational constraints?" comparing the output carbon pools' mean and distribution.

  Additionally, the modeled biomass is compared to empirical above-ground biomass for a given VOD using the relationship from Y. Y. Liu et al. (2015) (Supplementary Eq. 2 and Supplementary Figure 4 of the original article). The authors used reliable VOD observations and benchmark biomass maps to obtain the VOD-biomass relationship. While this relationship does not necessarily represent the ground truth biomass for a given VOD and depends on the choice of a reference for VOD in the original paper (since VOD derived from different frequency bands can be mutually biased), the closeness of the model output to the paper-derived limits shows us the ability of CARDAMOM model in a particular configuration to constrain carbon pools.

3 Results

Overall, more than 300 different scenarios were considered, and more than 5 000 CARDAMOM runs were performed (with 10 to 25 repetitions per scenario). Gelman-Rubin metric indicates convergence of most of the runs and most of the scenarios, with the exception of stochastic information content-related runs. In some cases, for example, for the FR-LBr site, when SIF (both synthetic and observational data) was assimilated, it resulted in divergence in some model parameters (e.g., Decomposition rate, Fraction of GPP respired, Leaf Lifespan, Canopy efficiency), however, when both SIF and VOD data were assimilated, all runs converged.

3.1 Can SIF and VOD observations substitute NBE?

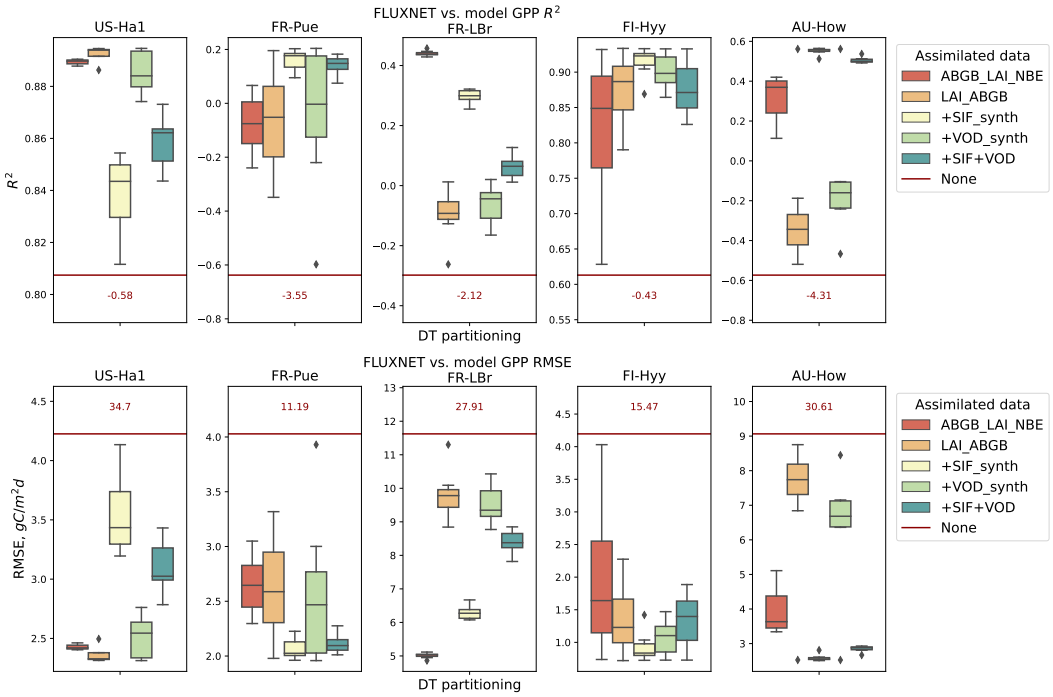


Figure 1. Effect of assimilating different data combinations on determination coefficient R^2 and RMSE between median CARDAMOM results and FLUXNET GPP data. The 'none' case (no data assimilated) is denoted as a red line not in scale with the corresponding value of the parameter denoted.

We start by comparing the model performance in matching FLUXNET daytime GPP for all five sites. Figure 1 illustrates the base case 'ABGB_LAI_NBE' in compar-

ison to the case without NBE data directly assimilated ('LAI_ABGB') and with SIF, VOD, and both SIF and VOD consequently added to the set of the assimilated observations. Comparing the first two cases, we notice that NBE data is only crucial for FR-LBr and AU-How sites. For the rest of the sites, without NBE, CARDAMOM can converge to the same or even better solution. When SIF and VOD are added, it further improves the results for FR-Pue, FR-LBr, and AU-How sites. For US-Ha1, the '+SIF+VOD' case results are worse than the 'ABGB_LAI_NBE' case, however, the fit is still tight - $R^2 = 0.86$ in comparison to the initial $R^2 = 0.89$. For FI-Hyy site, there was no apparent effect related to the SIF and VOD inclusion, which is likely due to this Finnish evergreen needle-leaved forest site with cool summer and no dry season experiencing little to no water stress. In all cases, significant improvements in the model performance are observed with any data assimilated in comparison to no data assimilated (the 'none' case). In all cases except for the US-Ha1 site, adding SIF and VOD, especially together, can successfully substitute NBE data.

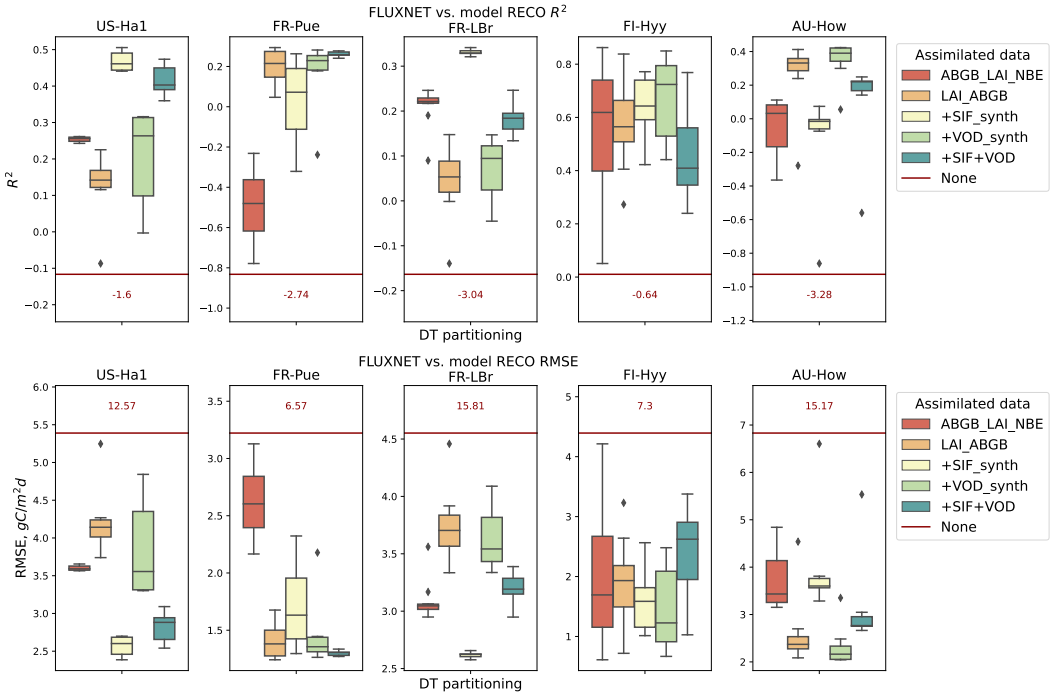


Figure 2. Effect of assimilating different combinations of data on determination coefficient R^2 and RMSE between median CARDAMOM results and FLUXNET RECO data. The 'none' case (no data assimilated) is denoted as a red line not in scale with the corresponding value of the parameter denoted.

Next, we compare the model performance in the same setup to match the FLUXNET estimated ecosystem respiration. Figure 2 illustrates the change in R^2 and RMSE for ecosystem respiration when assimilating different combinations of observational datasets for all five sites. Here, NBE was improving the model results only for US-Ha1 and FR-LBr sites. From Table 3, we notice that there are less available data points for SIF and VOD observational data in comparison to NBE and LAI. However, even with that, assimilating SIF and VOD results in higher R^2 and lower RMSE than for 'ABGB_LAI_NBE' and/or 'LAI_ABGB' cases for all sites except for FI-Hyy. For example, for respiration at US-Ha1, R^2 increases from 0.25 for 'ABGB_LAI_NBE' and 0.16 for 'LAI_ABGB' to 0.42 with inclusion of SIF and VOD. Correspondingly, respiration RMSE reduces from 3.6 and 4.2 to 2.8 gC/m^2d. For the FR-Pue site, assimilating LAI and NBE results in the worst performance among all cases, whereas the '+SIF+VOD' shows the best and most consistent results with the narrowest interquartile interval for both R^2 and RMSE, indicating model convergence across optimized parameters. This evergreen oak forest is located in a Mediterranean climate with long dry summers and has large interannual variation (Rambal et al., 2004), which can explain why additional observational data can improve the model results. For FR-LBr, '+SIF+VOD' cannot achieve the results obtained for the 'ABGB_LAI_NBE' case, however, the SIF and VOD addition significantly improves the results in comparison to the 'LAI_ABGB' case. In other words, SIF and VOD cannot fully substitute NBE for this site but still bring improvement in contrast to the case without either. The Finish site again does not show any improvements with respect to different combinations of data assimilated. Finally, the AU-How results show that the assimilation of SIF and VOD improves the results compared to the 'ABGB_LAI_NBE' case but not the 'LAI_ABGB' case. That might be due to the peculiarities of this site and LAI and VOD formulations in the model that will be discussed separately in **Section 3.1.1. Peculiar case of Australian site**.

For all sites, SIF and VOD data can either substitute or improve the results obtained with NBE, however, the degree of improvement and the effect of a particular combination of the assimilated data varies from site to site, depending on the local conditions and, potentially, data availability and quality.

Overall, the results measured with R^2 and RMSE are similar to the result based on histogram intersection (Fig. 3) and interannual anomaly metric ratio (Fig. 4). For the histogram intersection, the '+SIF+VOD' case improves the results in comparison

to the 'ABGB_LAI_NBE' case for US-Ha1 both respiration and GPP, FR-Pue and AU-How respiration, and for the rest of the cases, improves the results in comparison to 'LAI_ABGB'. The only exception is the FI-Hyy site, where the addition of SIF and VOD degrades the respiration histogram intersection. The fact that SIF and VOD assimilation can improve both the metrics related to the model's median output (like R^2 and RMSE) and the model's distribution output is an important result. Improving on just the median metrics can indicate overfitting of the model, whereas improvement on histogram intersection alone can signal about the model underfitting. The improvement on the two types of metrics indicates the model indeed matched the observational distributions better. However, the '+SIF+VOD' case never reached the interannual variability, measured as the IAM ratio, achieved for the 'ABGB_LAI_NBE' case except for Harvard Forest respiration and FI-Hyy site. At the same time, assimilating SIF and VOD consistently improves the interannual variability compared to the 'LAI_ABGB'. In other words, while SIF and VOD cannot fully substitute the degree of variability introduced by NBE, they still improve the results in comparison to the case when NBE is simply not used.

3.1.1 Peculiar case of Australian site

Another reason why VOD was less effective for the Australian (AU-How) site might lie in the peculiarity of the VOD and LAI dynamic for this site. As shown in Fig. S7, in this particular case, the synthetic and observational VOD data seem decoupled. That is because we defined synthetic VOD as being proportional to aboveground biomass (Eq. 4). Considering that LAI is also linearly proportional to the leaf mass, by definition, VOD and LAI will always be coupled in the model. Yet, Tian et al. (2018) showed that VOD and LAI can be decoupled for some regions. In particular, this pattern was found for African tropical woodlands (centered at 11.5°S, 18.5°E). While the Australian site is classified as a woody savanna, it is located at the same latitude as the African study region and may exhibit similar dynamics. In addition, grass and trees in woody savannas have different seasonal dynamics, with LAI dynamics driven mainly by the strong seasonal cycle of grass, whereas the dynamics of VOD are driven by trees. Due to the way LAI and VOD are represented in the model, CARDAMOM is unable to reproduce such a decoupling pattern. The tradeoff between model complexity (and, correspondingly, its ability to replicate a wide variety of natural phenomena and, here, multiple plant functional

types within a single pixel) and model fidelity should be considered when carbon models like CARDAMOM are used.

3.2 Which of the SIF and VOD is more useful in this task?

Further, different configurations of the assimilated data can help answer the research question: Which of the SIF and VOD is more useful in this task, and how does it depend on the local conditions? Supplementary figures S1–S4 illustrate the same metric as discussed above for assimilating either synthetic or observational SIF and VOD. For all sites, including SIF in data assimilation immediately improves the model results for GPP and respiration. For US-Ha1, FR-LBr, and FI-Hyy, the best results were achieved with the assimilation of synthetic SIF data, which were outperformed by the observational SIF only for FR-Pue and AU-How. Note that from Table 3, these are the two sites with the most SIF data available, suggesting that more data can help further improve the model skill in constraining respiration. The difference between the two versions of the observational SIF data (GOME SIF and normalized GOME SIF) is minimal and was only evident for the FR-LBr site. Synthetic and observational VOD performed surprisingly similarly when assessed by respiration R^2 and RMSE, except FR-Pue and AU-How sites that had slightly better performance with the synthetic VOD. GLAB-VOD data and VODCA VOD data are very different in the mean and the amplitude of year-to-year variation (for illustration, see Fig. S5 and S7). Despite that, there is an apparent similarity of assimilating either of the VOD datasets on the model performance. We can assume some level of flexibility in this model configuration, and note that the effect of assimilating VOD should be more pronounced when looking at the simulated pools, rather than fluxes (like GPP and respiration). At the same time, assimilating both SIF and VOD together leads to improved results across sites and on an aggregate basis across metrics.

3.3 Can respiration flux and carbon pools be reliably inferred from data assimilation, given indirect observational constraints?

Figure 5 illustrates the effect of assimilating LAI, ABGB, SIF, and VOD ('+SIF+VOD' case) in comparison to 'NBE_LAI_ABGB' case and FLUXNET data on GPP and respiration. For all time series, the uncertainty is presented. The median output for GPP in the '+SIF+VOD' case slightly underestimates the summer peak. Still, the model output has a more prominent uncertainty range than the 'NBE_LAI_ABGB' case that in-

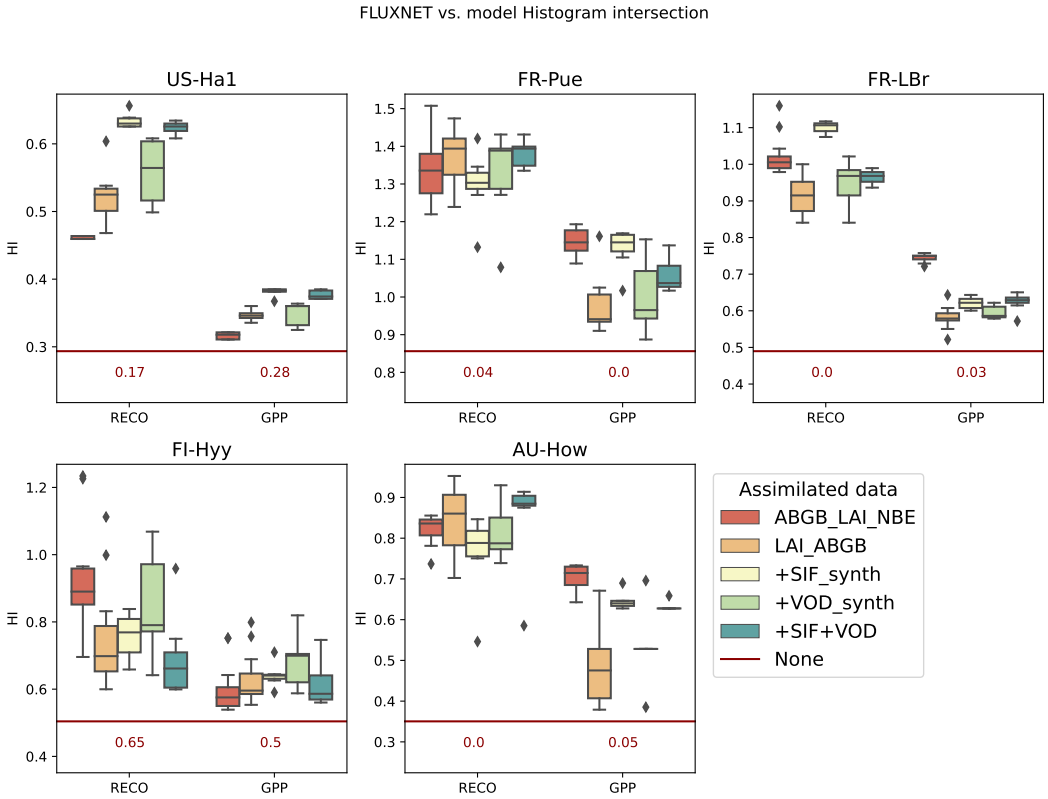


Figure 3. Effect of assimilating different combinations of data on Histogram intersection between CARDAMOM results distribution and FLUXNET RECO and GPP data. The 'none' case (no data assimilated) is denoted as a red line not in scale with the corresponding value of the parameter denoted.

cludes the observational data. In turn, the FLUXNET respiration time series are better matched with the SIF and VOD assimilation. Figure 6 sheds some light on the potential reasons behind the shift from the 'NBE_LAI_ABGB' case to the '+SIF+VOD' case. It shows the posterior distribution for 10 sample model parameters, reflecting the changes in model dynamics (parameters Fraction of GPP respired - "Frac GPP resp"; temperature sensitivity Q10 - "q10"; canopy efficiency - "Canopy eff"; leaf mass carbon per area, gC/m^2 - "LMCA," moisture factor) and initial values for five carbon pools ("C labile" to "C SOM," gC/m^2). Most parameters converge equally well for both 'NBE_LAI_ABGB' and '+SIF+VOD' cases with the posterior distributions clearly defined. While the fraction of GPP respired varies across a wider interval for the '+SIF+VOD' case, it takes more realistic values around 50% (Van Oijen et al., 2010), than less than 25% obtained in 'NBE_LAI_ABGB' case. At the same time, the posterior distributions for Q10 and

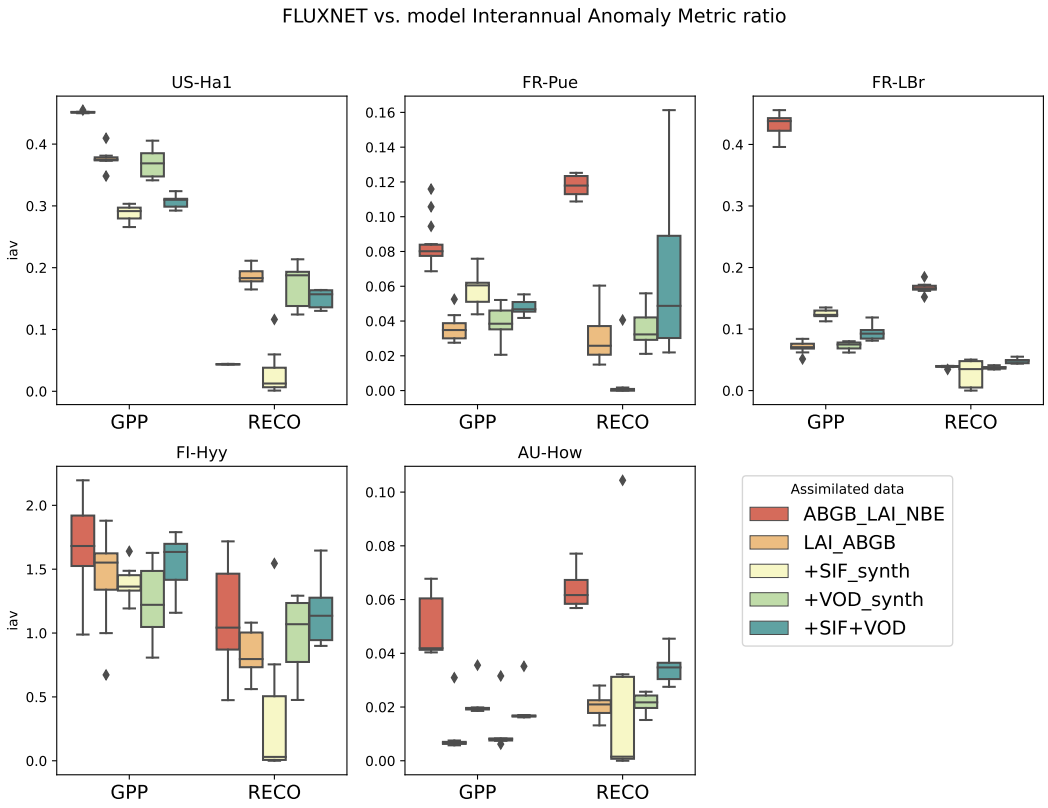


Figure 4. Effect of assimilating different combinations of data on relative interannual anomaly metric between median CARDAMOM results and FLUXNET GPP data. The 'none' case (no data assimilated) is denoted as a red line not in scale with the corresponding value of the parameter denoted.

canopy efficiency (that is used to calculate GPP using aggregated canopy model (ACM)) have broader ranges for the '+SIF+VOD' case, which is expected since NBE is removed from the assimilated data, and hence, GPP-related parameters are less constrained. There is also a shift in LMCA. For deciduous forests (e.g., at US-Ha1), the typical value of leaf mass per area is 75 g/m^2 (Poorter et al., 2009), which translates into a mean leaf carbon mass per area 37.5 gC/m^2, that corresponds better with the '+SIF+VOD' case results. A slight shift in moisture factor – a parameter used in CARDAMOM to scale decomposition rate based on water availability – in the '+SIF+VOD' case shows that precipitation influences the decomposition rate more than in the 'NBE_LAI_ABGB' case. Another apparent shift is in initial carbon pool partitioning – the initial leaf carbon pool is smaller, with the initial value for the wood carbon pool being larger for the '+SIF+VOD' case in comparison to the previous case. At the same time, the '+SIF+VOD' case sug-

gests very low initial values for the root carbon pool, which might be infeasible. Fig. S6 shows the same 10 parameters for the remaining four sites. For all sites, there is a significant difference in distribution for at least some of the parameters, especially for the fraction of GPP respired, Q10, canopy efficiency, and wood carbon pool. This suggests that assimilating SIF and VOD positively affects the model and, indeed, favors better constraining respiration and carbon pools.

Overall, assimilating SIF and VOD instead of NBE does not deteriorate the model results and even seems to improve them at some sites, leading to a reduction in respiration RMSE for US-Ha1, FR-Pue, and AU-How. Furthermore, the changes in the model results are likely due to improvement of the model parameters, which shift towards more realistic values. Note that good metrics and a decent fit for the 'NBE_LAI_ABGB' case were achieved with a very different combination of CARDAMOM parameters than in the '+SIF+VOD' case as shown in Fig. 6.

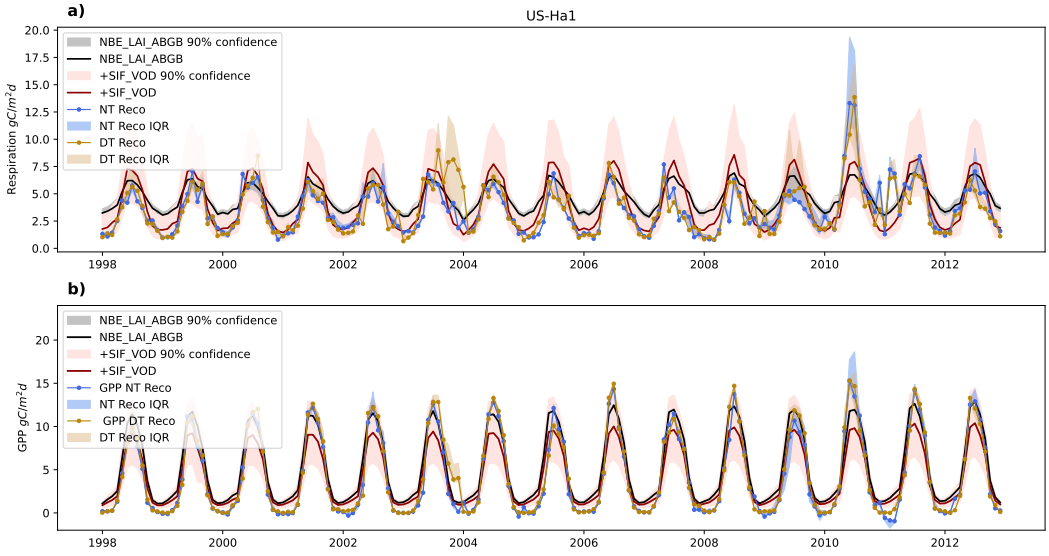


Figure 5. GPP and Respiration time series comparing FLUXNET data and median model outputs for 'NBE_LAI_ABGB' and '+SIF+VOD' cases with inter-quantile range for all variables

3.4 Constraining carbon pools

Assimilating SIF and VOD allows for constraining the carbon pools. Fig. 7 represents the temporal dynamics of the Harvard Forest carbon pools for the 'NBE_LAI_ABGB' and '+SIF+VOD' cases. The median and the interquantile range correspond to the full

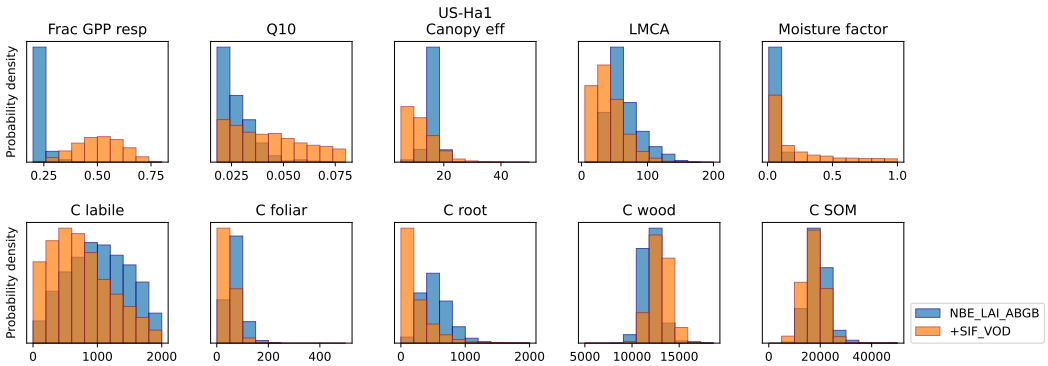


Figure 6. The effect of assimilating SIF and VOD in comparison to the 'NBE_LAI_ABGB' case on a selection of CARDAMOM parameters for the Harvard Forest site. All parameters are given in the ranges of their prior, except for C foliar pool (the original prior is 1 – 2000 gC/m^2), C wood (the original prior is 1 – 100,000 gC/m^2), and C SOM (the original prior is 1 – 200,000 gC/m^2). The ranges were changed for these three parameters to highlight the visual differences between the two cases.

model outputs from all runs for a given scenario combined. Essentially, the difference in the carbon pool partitioning represents the effect of substituting NBE with SIF and VOD in data assimilation. This shift corresponds to the shift in the initial values of carbon pools discussed above. In addition to it, the '+SIF+VOD' case modifies the dynamics of all pools, effectively propagating the constraints through the model. The same effect persists for all sites examined in this study (See Supplementary Fig. S8 - S11). For the Harvard Forest site, SIF and VOD reduced the size of all pools except for the wood pool compared to the 'NBE_LAI_ABGB' case. For example, in the 'NBE_LAI_ABGB' case, the mean values for the root carbon pool are close to 500 gC/m^2, while in the '+SIF+VOD' case, the mean value is reduced to 150 gC/m^2. Additionally, for this pool, the seasonal amplitude is reduced from 260 gC/m^2 to 60 gC/m^2. Another interesting note is related to the change in the wood carbon pool dynamics. In the 'NBE_LAI_ABGB' case, it has a significant trend growing from 11,900 gC/m^2 on average in 1998 to 13,100 gC/m^2 by the end of 2012, whereas, for the '+SIF+VOD' case, the overall growth over these years is about 300 gC/m^2. The more stable size of the wood biomass and moderate variability of the root biomass are expected for a stable ecosystem of the US-Ha1 site (Finzi et al., 2020).

Another interesting result is revealed when above-ground biomass (ABG, calculated as the sum of foliar and wood biomass) is compared to VOD and the empirical range defined in Y. Y. Liu et al. (2015) as shown in Fig. 8. While the estimated biomass in the '+SIF+VOD' case still mainly lies outside the range defined in Y. Y. Liu et al. (2015), the VOD-AGB relationship is closer to the empirically determined bounds than the one obtained from the 'NBE_LAI_ABGB' case. By definition, in-model VOD is linear with respect to above-ground biomass and hence, will not follow the empirical relationship exactly. Yet, the shift in carbon pools distribution and size brings the system towards a different equilibrium than the 'NBE_LAI_ABGB' case. This new state is more aligned with the Y. Y. Liu et al. (2015) findings. Hence, substituting NBE with SIF and VOD allows constraining carbon pools more effectively.

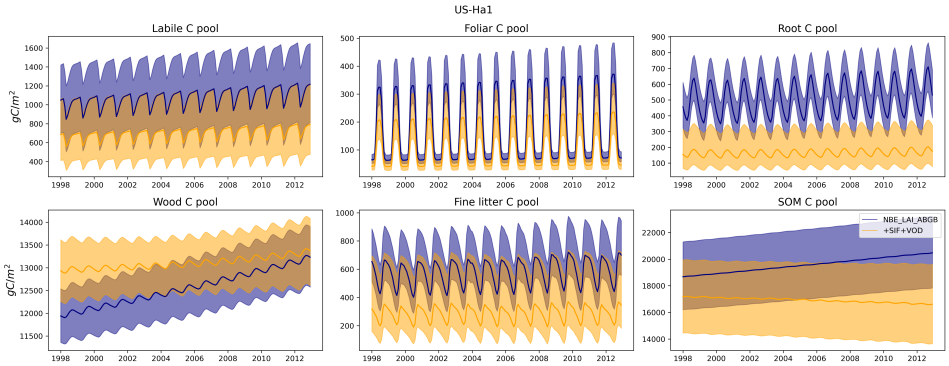


Figure 7. Carbon pools' temporal dynamic for 'NBE_LAI_ABGB' and '+SIF+VOD' cases for Harvard Forest. Median and interquantile range over the full output of all runs for a given scenario.

3.5 Information content

The real-world observational data can be sparse and uncertain. We noted this effect already when examining the VOD data availability for different sites. We further explore the effect of assimilating more or less data and data uncertainty more systematically here.

Fig. 9 shows how varying the number of assimilated data points and data uncertainty affect ecosystem respiration estimates in terms of R^2 and RMSE for our five FLUXNET sites. The behavior at the Harvard Forest site follows our expectations when reducing

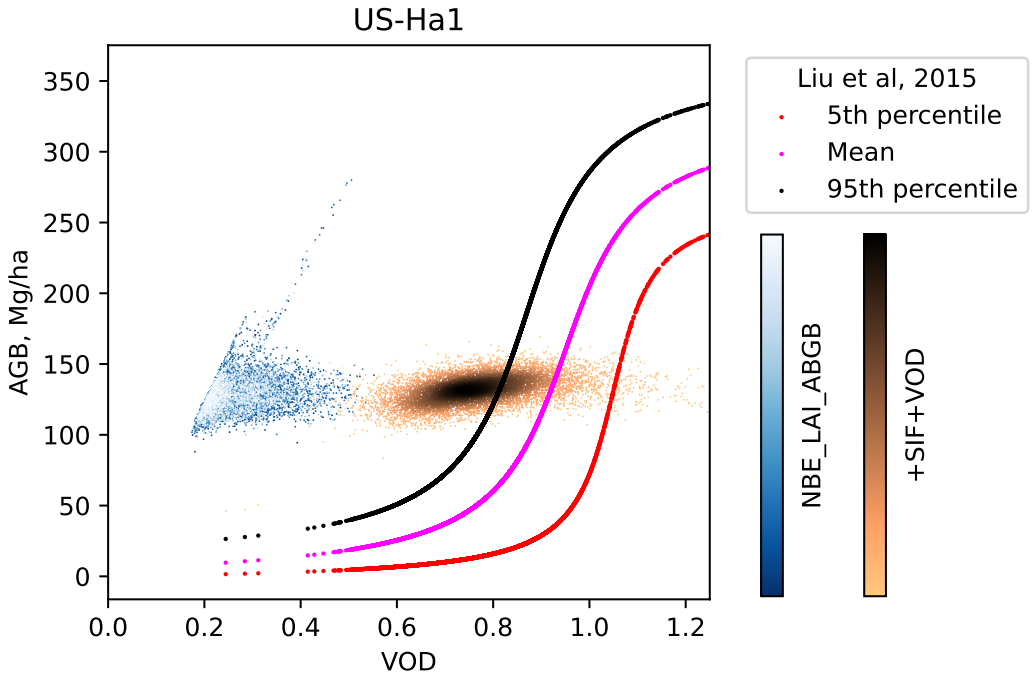


Figure 8. VOD vs. biomass relationship for 'NBE_LAI_ABGB' and '+SIF+VOD' cases in comparison to the empirical range defined in (Y. Y. Liu et al., 2015)

the amount of assimilated data deteriorated the model performance. Reducing the data to 40 and 20% of the initially available data is equivalent to doubling and quadrupling data uncertainty. Interestingly, reducing the data uncertainty for US-Ha1 improves the results for both metrics (R^2 0.52 vs. 0.42, RMSE 2.29 vs 2.82 gC/m^2d for the initial uncertainty), suggesting that the used uncertainty is not optimal and the model can extract more signal information from the assimilated data. For the FR-Pue site, the results changing the data amount follow a similar pattern; however, reducing and increasing data uncertainty deteriorates the results. Based on that, we can conclude that the initially used data uncertainty for this site is close to the optimal values. For the second French site, FR-LBr, the pattern of the results changes – using 80 and 60% of the initially available data improves the respiration metrics (followed by further deterioration of the results for even smaller amount of the used data that is also equivalent to quadrupling the data uncertainty). Given the overall scarcity of SIF and VOD data for this site and an apparent improvement of the results with the reduced data uncertainty (R^2 0.25 vs 0.18, RMSE 2.93 vs. 3.19 gC/m^2d for the initial uncertainty), we can assume that the model struggles to fully differentiate between the signal and the noise in the initial

configuration. Based on the results for the Finnish site, the model might be overfitted when all available data is used. Here, the best results (yet with quite a large uncertainty) are achieved for the smallest portion of the data – 20%. For this site, it was already noted that assimilating SIF and VOD does not improve the model results. Here, it is further confirmed – an attempt to impose more constraints for this evergreen needle-leaved forest site with little year-to-year variability deteriorates CARDAMOM predictions. Interestingly, the results for the AU-How site do not follow any pattern, most likely due to the previously mentioned issue of LAI and VOD varying in anti-phase. Since the data points for the assimilation are selected at random (Section 2.3 **Information content of observations**), this subset can either improve or deteriorate the model output depending on the selected points.

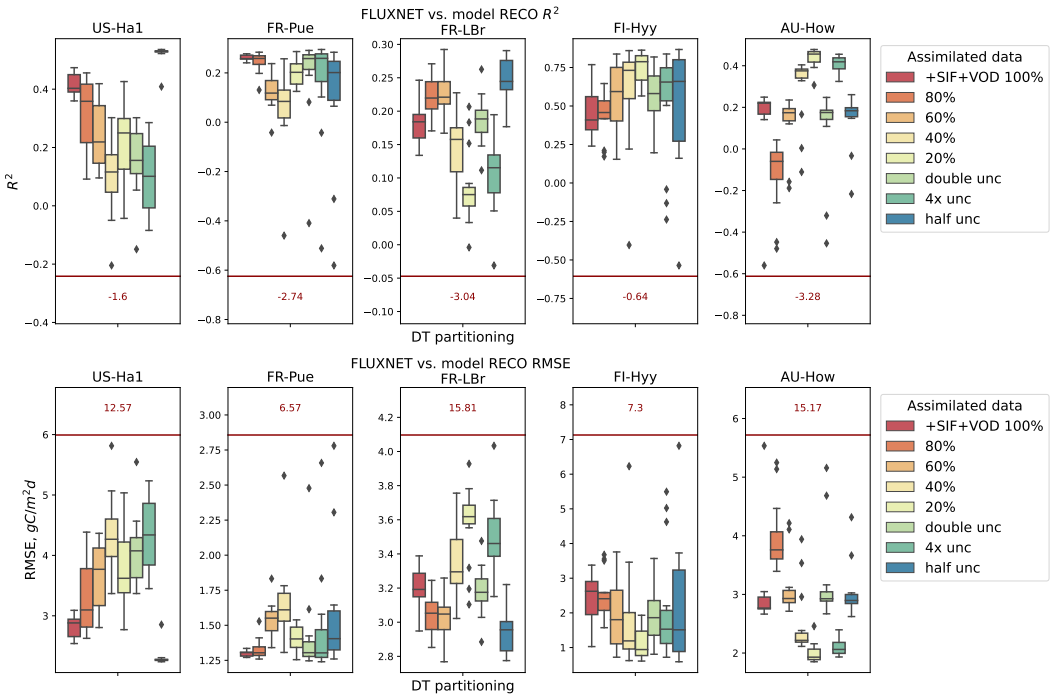


Figure 9. The effect of information content on CARDAMOM results: respiration R^2 and RMSE for all five sites

4 Conclusions

Process-based carbon cycle models like CARDAMOM can effectively model the carbon cycle and reconstruct carbon fluxes so that they match well with observations. The quality of the skill of the model in adequately simulating carbon pools and predicting

meaningful system behavior in contrafactual scenarios is based on the model structure and how well it can reproduce natural phenomena with model equations – and model parameters that are tuned through assimilation to observational data. Given the absence of direct observations for respiration, photosynthesis, and net primary productivity and the scarcity of biomass data on the global scale, this study investigates CARDAMOM's ability to effectively utilize physical model constraints and indirect observations connected to those processes, namely Solar-Induced Fluorescence (SIF) (as a proxy for photosynthesis) and Vegetation optical Depth (VOD) (used as a proxy for biomass). The results indicate that these indirect observations can lead to improvements in the estimates of carbon pools and respiration flux, often better than the ones obtained using net ecosystem exchange observations. This effect is even more witnessable when we consider the scales of the observations since these results were achieved when comparing and assimilating FLUXNET site level net ecosystem exchange with a 25 km grid cell remote sensing-based SIF and VOD. The discrepancy between the site-level and remote sensing level scales as a potential source of inconsistencies between CARDAMOM results and data was also previously noted by A. J. Norton et al. (2023).

Moreover, the better performance achieved when assimilating net ecosystem exchange (e.g., higher R^2 for US-Ha1 GPP) can be due to overfitting the model to follow net ecosystem exchange variability. Bacour et al. (2023) used a different carbon cycle model with a different combination of observational data assimilated, yet came to similar conclusions. Assimilation of more data, in our case, VOD, and SIF, leads to more consistent results. Yet, similarly to Bacour et al. (2023), observational errors of multiple observations should be addressed to ensure an adequate signal-to-noise ratio and avoid under- and overfitting.

The effect to which SIF and VOD observations improve the model estimates of carbon pools and respiration flux depends on various factors. Among them are data availability, data uncertainty and local conditions (interannual variability, stresses). Informational content analysis reveals that for the locations that experience seasonal stresses, more data helps better constrain the model. However, even for the sites with more stable climate conditions, assimilating any data was beneficial compared to not assimilating any SIF and VOD. The potential of SIF for carbon cycle modeling was noted earlier (MacBean et al., 2018; Bacour et al., 2019; MacBean et al., 2022; A. J. Norton et al., 2023). Indeed, for FR-LBr, FI-Hyy, and AU-How, SIF not only successfully substi-

tuted but outperformed NBE (the performance measures as R^2 and RMSE for GPP). VOD, on the other hand, while was proposed for data assimilation in carbon cycle models (Scholze et al., 2017), and was assimilated in other instances (Kaminski et al., 2018; Smith et al., 2020), was not previously used in models like CARDAMOM. We show that assimilating VOD improves model performance in matching respiration flux and constraining carbon pools. Indeed, respiration and carbon pool sizes are related (Ma et al., 2022), and connected through the model parameters like respiration rate and canopy efficiency, that were better constrained with the new assimilated data.

Additionally, we show that SIF and VOD assimilation improves metrics like R^2 and RMSE, without deteriorating metrics like histogram intersection – a metric that assesses the output distribution. Nevertheless, SIF and VOD could not achieve the same level of interannual variability in the results that were previously achieved with net ecosystem exchange. This is expected since the site-level net ecosystem exchange comprises more signal than averaged over the grid cell observational SIF and VOD. However, given less reliable global net ecosystem exchange assessments (e.g., Peylin et al., 2013; Cui et al., 2022), this finding can be wavered on a global scale. Satellite-based observational SIF and VOD may be expected to further outperform the globally available version of net ecosystem exchange obtained from CO_2 inversion.

Future work should focus on extending the study's geographical area to include other climate zones. Observational data can have varying quality depending on the ecosystem specifics (e.g., dense or sparse vegetation, frozen ground, and complex topography can pose technical challenges for space-born sensors) or even human activity (e.g., radio-frequency interference contaminating the signal). Further, model parameters' sensitivity to the assimilated data and its uncertainty can be spatially variable (Ma et al., 2022). As we have already shown here, for the sites not subjected to stresses, the model can be overfitted to data so that assimilating less data leads to better model performance. Finding more regions prone to this behavior is essential to properly utilizing carbon cycle model-data assimilation on the global scale. In addition to this, more attention should be dedicated to the process representation in the model. As it was shown for LAI in A. J. Norton et al. (2023), process representation plays a crucial role in the model's ability to effectively extract information from the assimilated data to constrain model parameters. In this study,we chose a simple VOD representation with minimal new parameters introduced into the model. Even in this configuration, VOD successfully constrained carbon pools,

including aboveground biomass, and improved model performance in comparison to assimilating just SIF. Future work should include VOD and SIF within a broad range of datastreams to best constrain the different carbon fluxes and pools.

Open Research Section

The data and scripts are publicly available at `https://github.com/os2328/CARDAMOM_SIF_VOD`

Acknowledgments

The authors were funded by the NASA Science Utilization of SMAP program under grant 80NSSC20K1792. AGK and CAF were also funded by the NSF DEB grant 1942133. AGK was also funded by the Alfred P. Sloan Foundation. PG and OS acknowledge funding from Schmidt Future LEMONTREE project.

References

Arneth, A., Sitch, S., Pongratz, J., Stocker, B. D., Ciais, P., Poulter, B., ... others (2017). Historical carbon dioxide emissions caused by land-use changes are possibly larger than assumed. *Nature Geoscience*, *10*(2), 79–84.

Bacour, C., MacBean, N., Chevallier, F., Léonard, S., Koffi, E. N., & Peylin, P. (2023). Assimilation of multiple datasets results in large differences in regional- to global-scale nee and gpp budgets simulated by a terrestrial biosphere model. *Biogeosciences*, *20*(6), 1089–1111.

Bacour, C., Maignan, F., MacBean, N., Porcar-Castell, A., Flexas, J., Frankenberg, C., ... Bastrikov, V. (2019). Improving estimates of gross primary productivity by assimilating solar-induced fluorescence satellite retrievals in a terrestrial biosphere model using a process-based sif model. *Journal of Geophysical Research: Biogeosciences*, *124*(11), 3281–3306.

Barr, A., Richardson, A., Hollinger, D., Papale, D., Arain, M., Black, T., ... others (2013). Use of change-point detection for friction–velocity threshold evaluation in eddy-covariance studies. *Agricultural and forest meteorology*, *171*, 31–45.

Berbigier, P., Bonnefond, J.-M., & Mellmann, P. (2001). Co2 and water vapour fluxes for 2 years above euroflux forest site. *Agricultural and Forest Meteorology*, *108*(3), 183–197.

Beringer, J., Hutley, L. B., Tapper, N. J., & Cernusak, L. A. (2007). Savanna fires and their impact on net ecosystem productivity in north australia. *Global Change Biology*, *13*(5), 990–1004.

Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in water resources*, *16*(1), 41–51.

Beven, K., & Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the glue methodology. *Journal of hydrology*, *249*(1-4), 11–29.

Bloom, A., & Williams, M. (2015). Constraining ecosystem carbon dynamics in a data-limited world: integrating ecological" common sense" in a model–data fusion framework. *Biogeosciences*, *12*(5), 1299–1315.

Bloom, A. A., Exbrayat, J.-F., Van Der Velde, I. R., Feng, L., & Williams, M. (2016). The decadal state of the terrestrial carbon cycle: Global retrievals of terrestrial carbon allocation, pools, and residence times. *Proceedings of the National Academy of Sciences*, *113*(5), 1285–1290.

Bloom, A. A., Liu, J., Bowman, K. W., Konings, A., Saatchi, S., Worden, J. R., ... Schimel, D. (2018). Observing the tropical carbon balance sensitivity to memory and climate extremes. In *Agu fall meeting abstracts* (Vol. 2018, pp. B51G–2016).

Chevallier, F., Remaud, M., O'Dell, C. W., Baker, D., Peylin, P., & Cozic, A. (2019). Objective evaluation of surface-and satellite-driven carbon dioxide atmospheric inversions. *Atmospheric Chemistry and Physics*, *19*(22), 14233–14251.

Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A., & Totterdell, I. J. (2000). Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, *408*(6809), 184–187.

Cui, Y. Y., Zhang, L., Jacobson, A. R., Johnson, M. S., Philip, S., Baker, D., ... others (2022). Evaluating global atmospheric inversions of terrestrial net ecosystem exchange co2 over north america on seasonal and sub-continental scales. *Geophysical Research Letters*, *49*(18), e2022GL100147.

De Martonne, E. (1923). Aridité et indices d'aridité. *Académie des Sciences. Comptes Rendus*, *182*(23), 1935–1938.

Deng, F., & Chen, J. (2011). Recent global co 2 flux inferred from atmospheric co 2

observations and its regional analyses. *Biogeosciences*, *8*(11), 3263–3281.

Denissen, J. M., Teuling, A. J., Pitman, A. J., Koirala, S., Migliavacca, M., Li, W., ... Orth, R. (2022). Widespread shift from ecosystem energy to water limitation with climate change. *Nature Climate Change*, *12*(7), 677–684.

Famiglietti, C. A., Smallman, T. L., Levine, P. A., Flack-Prain, S., Quetin, G. R., Meyer, V., ... others (2021). Optimal model complexity for terrestrial carbon cycle prediction. *Biogeosciences*, *18*(8), 2727–2754.

Finzi, A. C., Giasson, M.-A., Barker Plotkin, A. A., Aber, J. D., Boose, E. R., Davidson, E. A., ... others (2020). Carbon budget of the harvard forest long-term ecological research site: Pattern, process, and response to global change. *Ecological Monographs*, *90*(4), e01423.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Bakker, D. C., Hauck, J., ... others (2022). Global carbon budget 2021. *Earth System Science Data*, *14*(4), 1917–2005.

Fuster, B., Sánchez-Zapero, J., Camacho, F., García-Santos, V., Verger, A., Lacaze, R., ... Smets, B. (2020). Quality assessment of proba-v lai, fapar and fcover collection 300 m products of copernicus global land service. *Remote Sensing*, *12*(6), 1017.

Gasser, T., Crepin, L., Quilcaille, Y., Houghton, R. A., Ciais, P., & Obersteiner, M. (2020). Historical co 2 emissions from land use and land cover change and their uncertainty. *Biogeosciences*, *17*(15), 4075–4101.

Haario, H., Saksman, E., & Tamminen, J. (2001). An adaptive metropolis algorithm. *Bernoulli*, 223–242.

Hartmann, H., Bahn, M., Carbone, M., & Richardson, A. D. (2020). Plant carbon allocation in a changing world–challenges and progress: introduction to a virtual issue on carbon allocation. *New Phytologist*, *227*(4), 981–988.

Jackson, T., & Schmugge, T. (1991). Vegetation effects on the microwave emission of soils. *Remote Sensing of Environment*, *36*(3), 203–212.

Jian, J., Steele, M. K., Thomas, R. Q., Day, S. D., & Hodges, S. C. (2018). Constraining estimates of global soil respiration by quantifying sources of variability. *Global Change Biology*, *24*(9), 4143–4159.

Joiner, J., Yoshida, Y., Koehler, P., Frankenberg, C., & Parazoo, N. (2023). *L2 daily solar-induced fluorescence (sif) from metop-a gome-2, 2007-2018. ornl daac,*

oak ridge, tennessee, usa.

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., ... Reichstein, M. (2019). The fluxcom ensemble of global land-atmosphere energy fluxes. *Scientific data*, *6*(1), 74.

Kaminski, T., Scholze, M., Knorr, W., Vossbeck, M., Wu, M., Ferrazzoli, P., ... others (2018). Constraining terrestrial carbon fluxes through assimilation of smos products. In *Igarss 2018-2018 ieee international geoscience and remote sensing symposium* (pp. 1455–1458).

Kennedy, D., Swenson, S., Oleson, K. W., Lawrence, D. M., Fisher, R., Lola da Costa, A. C., & Gentine, P. (2019). Implementing plant hydraulics in the community land model, version 5. *Journal of Advances in Modeling Earth Systems*, *11*(2), 485–513.

Kirschbaum, M. U., & McMillan, A. (2018). Warming and elevated co2 have opposing influences on transpiration. which is more important? *Current forestry reports*, *4*(2), 51–71.

Konings, A. G., Holtzman, N. M., Rao, K., Xu, L., & Saatchi, S. S. (2021). Interannual variations of vegetation optical depth are due to both water stress and biomass changes. *Geophysical Research Letters*, *48*(16), e2021GL095267.

Konings, A. G., Piles, M., Das, N., & Entekhabi, D. (2017). L-band vegetation optical depth and effective scattering albedo estimation from SMAP. *Remote Sensing of Environment*, *198*, 460–470.

Konings, A. G., Saatchi, S. S., Frankenberg, C., Keller, M., Leshyk, V., Anderegg, W. R., ... others (2021). Detecting forest response to droughts with global observations of vegetation water content. *Global change biology*, *27*(23), 6005–6024.

Li, X., Liu, F., & Fang, M. (2020). Harmonizing models and observations: Data assimilation in earth system science. *Science China Earth Sciences*, *63*, 1059–1068.

Li, X., Ma, H., Ran, Y., Wang, X., Zhu, G., Liu, F., ... Huang, C. (2021). Terrestrial carbon cycle model-data fusion: Progress and challenges. *Science China Earth Sciences*, *64*(10), 1645–1657.

Liu, L., Gudmundsson, L., Hauser, M., Qin, D., Li, S., & Seneviratne, S. I. (2020). Soil moisture dominates dryness stress on ecosystem production globally. *Na-*

ture communications, *11*(1), 1–9.

Liu, Y. Y., Van Dijk, A. I., De Jeu, R. A., Canadell, J. G., McCabe, M. F., Evans, J. P., & Wang, G. (2015). Recent reversal in loss of global terrestrial biomass. *Nature Climate Change*, *5*(5), 470–474.

López-Blanco, E., Exbrayat, J.-F., Lund, M., Christensen, T. R., Tamstorf, M. P., Slevin, D., ... Williams, M. (2019). Evaluation of terrestrial pan-arctic carbon cycling using a data-assimilation system. *Earth System Dynamics*, *10*(2), 233–255.

Ma, R., Xiao, J., Liang, S., Ma, H., He, T., Guo, D., ... Lu, H. (2022). Pixel-level parameter optimization of a terrestrial biosphere model for improving estimation of carbon fluxes with an efficient model–data fusion method and satellite-derived lai and gpp data. *Geoscientific Model Development*, *15*(17), 6637–6657.

MacBean, N., Bacour, C., Raoult, N., Bastrikov, V., Koffi, E., Kuppel, S., ... others (2022). Quantifying and reducing uncertainty in global carbon cycle predictions: Lessons and perspectives from 15 years of data assimilation studies with the orchidee terrestrial biosphere model. *Global Biogeochemical Cycles*, *36*(7), e2021GB007177.

MacBean, N., Maignan, F., Bacour, C., Lewis, P., Peylin, P., Guanter, L., ... Disney, M. (2018). Strong constraint on modelled global carbon uptake using solar-induced chlorophyll fluorescence data. *Scientific reports*, *8*(1), 1973.

Moesinger, L., Dorigo, W., de Jeu, R., van der Schalie, R., Scanlon, T., Teubner, I., & Forkel, M. (2020). The global long-term microwave vegetation optical depth climate archive (VODCA). *Earth System Science Data*, *12*(1), 177–196.

Moore, A. M., Arango, H. G., Broquet, G., Edwards, C., Veneziani, M., Powell, B., ... Robinson, P. (2011). The regional ocean modeling system (roms) 4-dimensional variational data assimilation systems: part ii–performance and application to the california current system. *Progress in Oceanography*, *91*(1), 50–73.

Munger, W., & Wofsy, S. (2014). Biomass inventories at harvard forest ems tower since 1993. *Harvard Forest Data Archive: HF069*.

Nathaniel, J., Liu, J., & Gentine, P. (2023). Metaflux: Meta-learning global carbon fluxes from sparse spatiotemporal observations. *Scientific Data*, *10*(1), 440.

Norton, A., Bloom, A. A., Parazoo, N., Levine, P. A., Ma, S., & Braghiere, R. K. (2021). The impacts of a climate-sensitive leaf phenology model on predicting the terrestrial carbon cycle using the cardamom framework. In *Agu fall meeting 2021.*

Norton, A. J., Bloom, A. A., Parazoo, N. C., Levine, P. A., Ma, S., Braghiere, R. K., & Smallman, T. L. (2023). Improved process representation of leaf phenology significantly shifts climate sensitivity of ecosystem carbon balance. *Biogeosciences*, *20*(12), 2455–2484.

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., ... others (2020). The fluxnet2015 dataset and the oneflux processing pipeline for eddy covariance data. *Scientific data*, *7*(1), 1–27.

Peylin, P., Law, R., Gurney, K., Chevallier, F., Jacobson, A., Maki, T., ... others (2013). Global atmospheric carbon budget: results from an ensemble of atmospheric co 2 inversions. *Biogeosciences*, *10*(10), 6699–6720.

Piao, S., Huang, M., Liu, Z., Wang, X., Ciais, P., Canadell, J. G., ... others (2018). Lower land-use emissions responsible for increased net land carbon sink during the slow warming period. *Nature Geoscience*, *11*(10), 739–743.

Poorter, H., Niinemets, Ü., Poorter, L., Wright, I. J., & Villar, R. (2009). Causes and consequences of variation in leaf mass per area (lma): a meta-analysis. *New phytologist*, *182*(3), 565–588.

Quetin, G. R., Bloom, A. A., Bowman, K. W., Diffenbaugh, N. S., Liu, J., & Konings, A. G. (2019). Attribution of historical terrestrial carbon uptake due to climate change. In *Agu fall meeting abstracts* (Vol. 2019, pp. B11B–03).

Quetin, G. R., Bloom, A. A., Bowman, K. W., & Konings, A. G. (2020). Carbon flux variability from a relatively simple ecosystem model with assimilated data is consistent with terrestrial biosphere model estimates. *Journal of Advances in Modeling Earth Systems*, *12*(3), e2019MS001889.

Rambal, S., Joffre, R., Ourcival, J., Cavender-Bares, J., & Rocheteau, A. (2004). The growth respiration component in eddy co2 flux from a quercus ilex mediterranean forest. *Global Change Biology*, *10*(9), 1460–1469.

Raupach, M. R., Rayner, P. J., Barrett, D. J., DeFries, R. S., Heimann, M., Ojima, D. S., ... Schmullius, C. C. (2005). Model–data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications.

Global Change Biology, *11*(3), 378–397.

Scholze, M., Buchwitz, M., Dorigo, W., Guanter, L., & Quegan, S. (2017). Reviews and syntheses: Systematic earth observations for use in terrestrial carbon cycle data assimilation systems. *Biogeosciences*, *14*(14), 3401–3429.

Skulovich, O., Li, X., Wigneron, J.-P., & Gentine, P. (2024). GLAB-VOD: Global L-band AI-BasedVegetation Optical Depth dataset based on machine learning and remote sensing. *Scientific Data (submitted)*. Retrieved from `https://zenodo.org/doi/10.5281/zenodo.10306094`

Smith, W. K., Fox, A. M., MacBean, N., Moore, D. J., & Parazoo, N. C. (2020). Constraining estimates of terrestrial carbon uptake: New opportunities using long-term satellite observations and data assimilation. *New Phytologist*, *225*(1), 105–112.

Song, J., Wan, S., Piao, S., Knapp, A. K., Classen, A. T., Vicca, S., ... others (2019). A meta-analysis of 1,119 manipulative experiments on terrestrial carbon-cycling responses to global change. *Nature ecology & evolution*, *3*(9), 1309–1320.

Suni, T., Rinne, J., Reissell, A., Altimir, N., Keronen, P., Rannik, U., ... Vesala, T. (2003). Long-term measurements of surface fluxes above a scots pine forest in hyytiala, southern finland, 1996-2001. *Boreal Environment Research*, *8*(4), 287–302.

Talagrand, O. (1997). Assimilation of observations, an introduction (gtspecial issueltdata assimilation in meteology and oceanography: Theory and practice). *Journal of the Meteorological Society of Japan. Ser. II*, *75*(1B), 191–209.

Tharammal, T., Bala, G., Narayanappa, D., & Nemani, R. (2019). Potential roles of co 2 fertilization, nitrogen deposition, climate change, and land use and land cover change on the global terrestrial carbon uptake in the twenty-first century. *Climate Dynamics*, *52*, 4393–4406.

Tian, F., Wigneron, J.-P., Ciais, P., Chave, J., Ogée, J., Peñuelas, J., ... others (2018). Coupling of ecosystem-scale plant water storage and leaf phenology observed by satellite. *Nature ecology & evolution*, *2*(9), 1428–1435.

Van Oijen, M., Schapendonk, A., & Höglind, M. (2010). On the relative magnitudes of photosynthesis, respiration, growth and carbon storage in vegetation. *Annals of Botany*, *105*(5), 793–797.

Wang, H., Jiang, F., Wang, J., Ju, W., & Chen, J. M. (2019). Terrestrial ecosystem carbon flux estimated using gosat and oco-2 xco 2 retrievals. *Atmospheric Chemistry and Physics*, *19*(18), 12067–12082.

Wang, Y.-P., Trudinger, C. M., & Enting, I. G. (2009). A review of applications of model–data fusion to studies of terrestrial carbon fluxes at different scales. *Agricultural and forest meteorology*, *149*(11), 1829–1842.

Wigneron, J.-P., Jackson, T., O'neill, P., De Lannoy, G., de Rosnay, P., Walker, J., ... others (2017). Modelling the passive microwave signature from land surfaces: A review of recent results and application to the l-band smos & smap soil moisture retrieval algorithms. *Remote Sensing of Environment*, *192*, 238–262.

Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., ... others (2009). Improving land surface models with fluxnet data. *Biogeosciences*, *6*(7), 1341–1359.

Wood, J. D., Griffis, T. J., Baker, J. M., Frankenberg, C., Verma, M., & Yuen, K. (2017). Multiscale analyses of solar-induced florescence and gross primary production. *Geophysical Research Letters*, *44*(1), 533–541.

Wu, C., Niu, Z., Tang, Q., Huang, W., Rivard, B., & Feng, J. (2009). Remote estimation of gross primary production in wheat using chlorophyll-related vegetation indices. *Agricultural and Forest Meteorology*, *149*(6-7), 1015–1021.

Xu, Z., Jiang, Y., & Zhou, G. (2015). Response and adaptation of photosynthesis, respiration, and antioxidant systems to elevated co2 with environmental stress in plants. *Frontiers in plant science*, *6*, 701.

Yang, X., Tang, J., Mustard, J. F., Lee, J.-E., Rossini, M., Joiner, J., ... Richardson, A. D. (2015). Solar-induced chlorophyll fluorescence that correlates with canopy photosynthesis on diurnal and seasonal scales in a temperate deciduous forest. *Geophysical Research Letters*, *42*(8), 2977–2987.

Yang, Y., Bloom, A. A., Ma, S., Levine, P., Norton, A., Parazoo, N. C., ... others (2022). Cardamom-fluxval version 1.0: a fluxnet-based validation system for cardamom carbon and water flux estimates. *Geoscientific Model Development*, *15*(4), 1789–1802.

Yang, Y., Saatchi, S., Bloom, A. A., & Yu, Y. (2019). Validation of cms carbon flux and stock datasets across the fluxnet eddy covariance flux tower network. In

Agu fall meeting abstracts (Vol. 2019, pp. B13F–2435).

Yin, Y., Bloom, A. A., Worden, J., Saatchi, S., Yang, Y., Williams, M., ... others (2020). Fire decline in dry tropical ecosystems enhances decadal land carbon sink. *Nature communications*, *11*(1), 1–7.

Zhang, Y., Guanter, L., Berry, J. A., Joiner, J., van der Tol, C., Huete, A., ... Köhler, P. (2014). Estimation of vegetation photosynthetic capacity from space-based measurements of chlorophyll fluorescence for terrestrial biosphere models. *Global change biology*, *20*(12), 3727–3742.

5 Supplemental

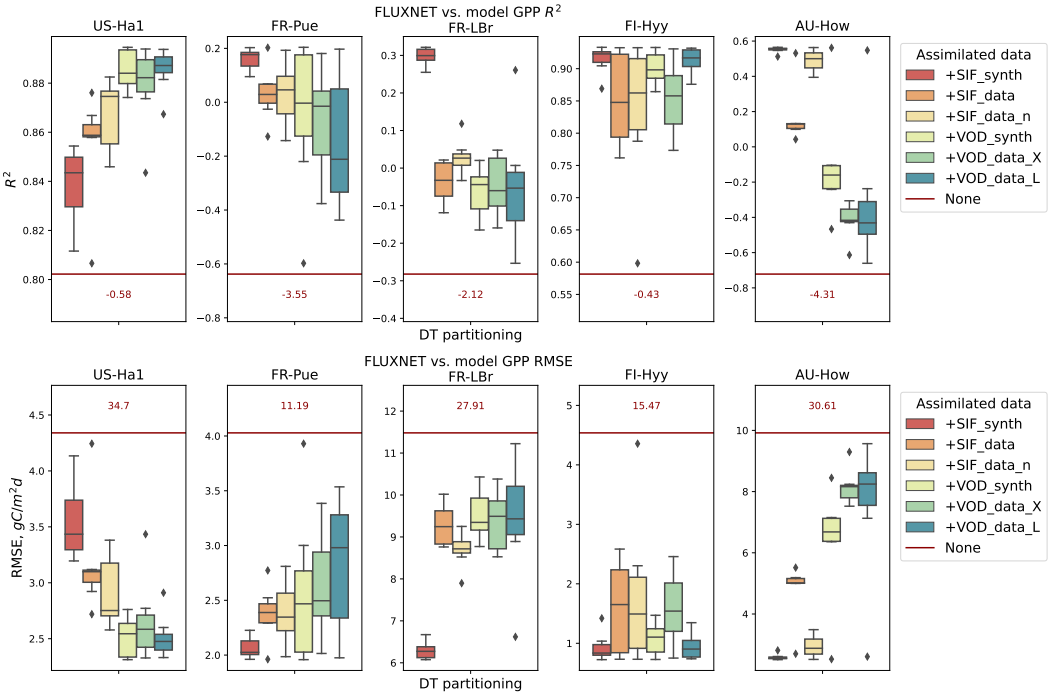


Figure S1. Effect of assimilating different data combinations on determination coefficient R^2 and RMSE between median CARDAMOM results and FLUXNET GPP data. The 'none' case (no data assimilated) is denoted as a red line not in scale with the corresponding value of the parameter denoted.

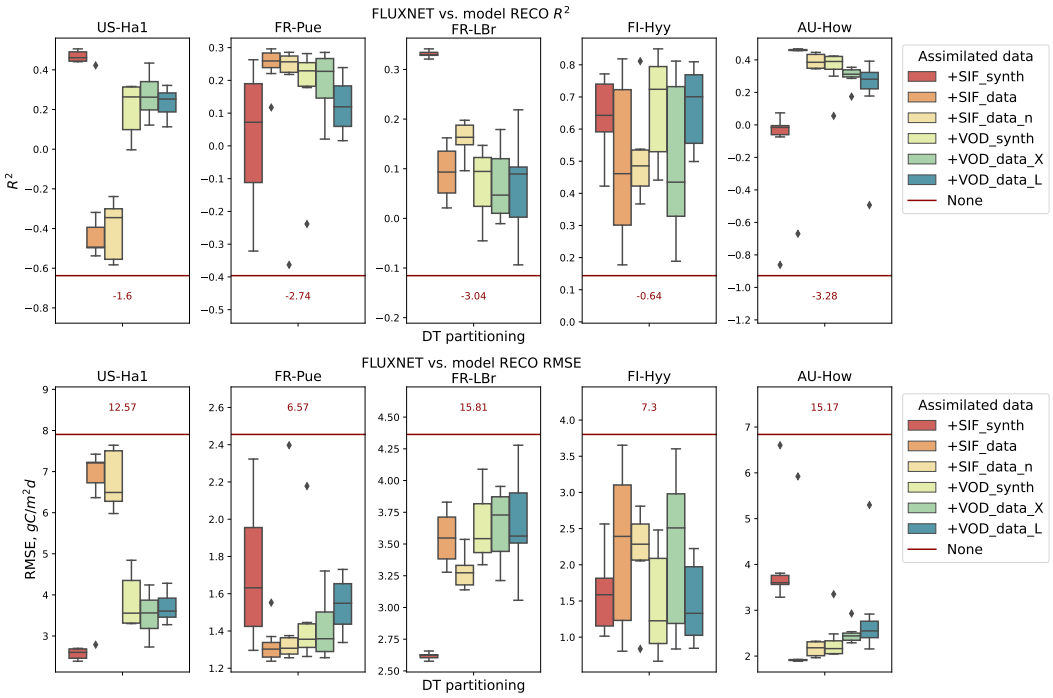


Figure S2. Effect of assimilating different combinations of data on determination coefficient R^2 and RMSE between median CARDAMOM results and FLUXNET RECO data. The 'none' case (no data assimilated) is denoted as a red line not in scale with the corresponding value of the parameter denoted.

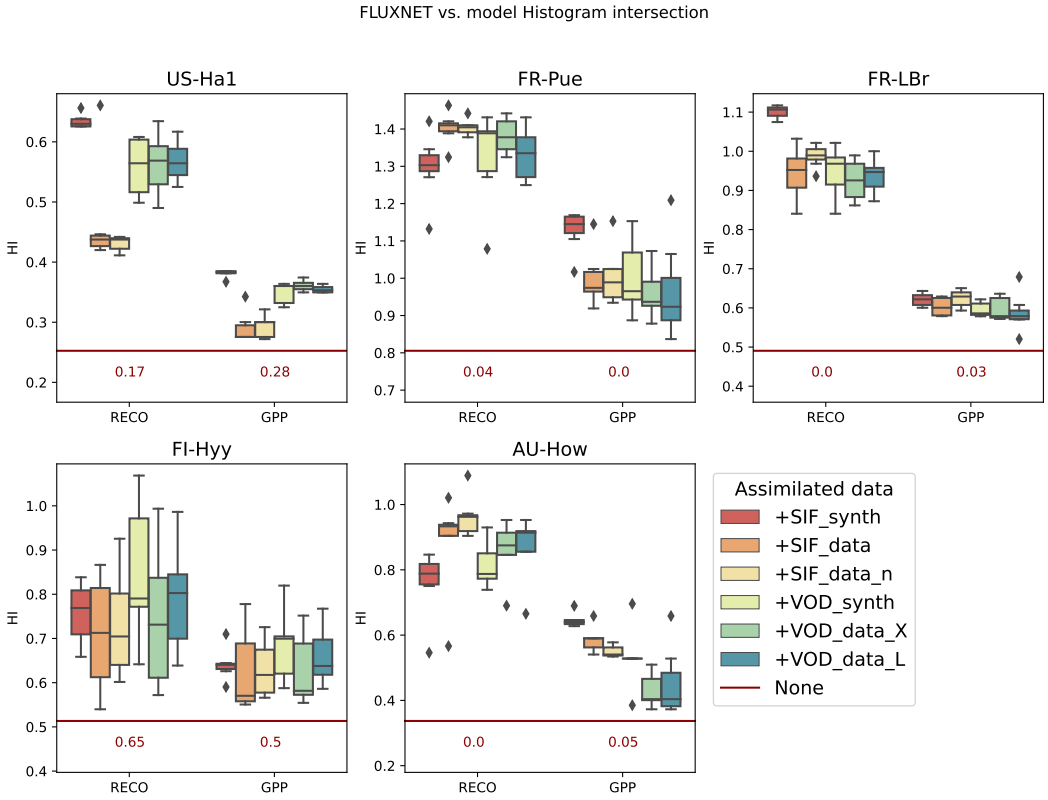


Figure S3. Effect of assimilating different combinations of data on Histogram intersection between CARDAMOM results distribution and FLUXNET RECO and GPP data. The 'none' case (no data assimilated) is denoted as a red line not in scale with the corresponding value of the parameter denoted.

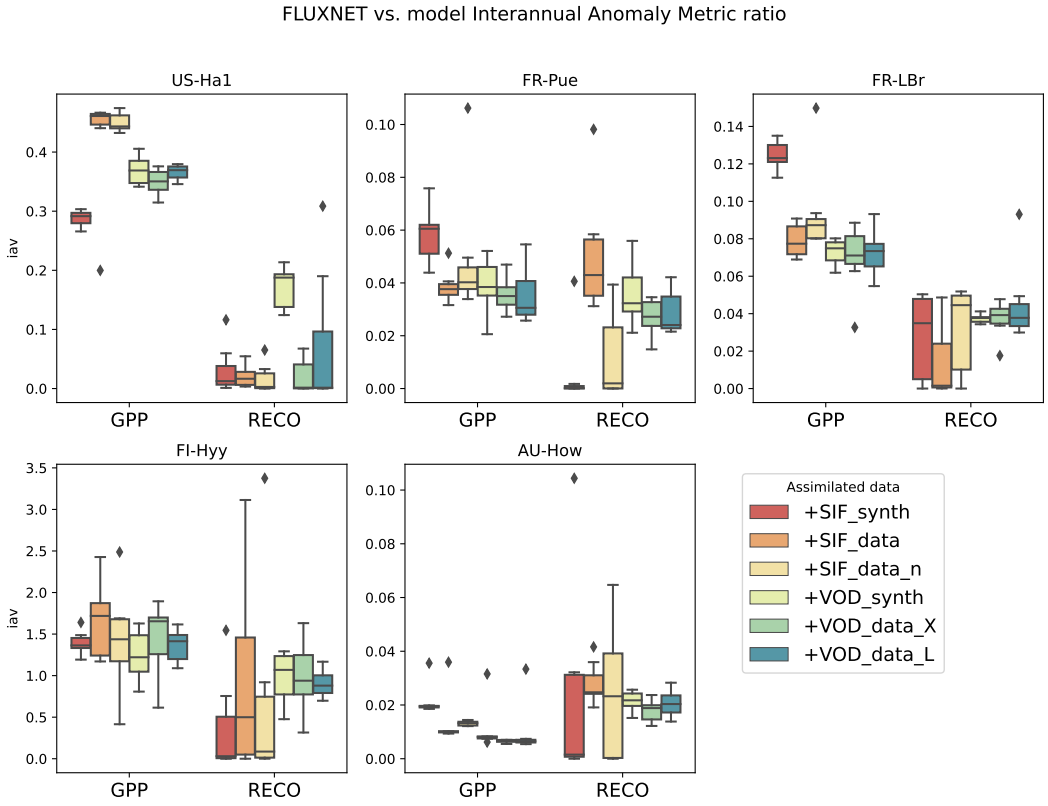


Figure S4. Effect of assimilating different combinations of data on relative interannual anomaly metric between median CARDAMOM results and FLUXNET GPP data. The 'none' case (no data assimilated) is denoted as a red line not in scale with the corresponding value of the parameter denoted.

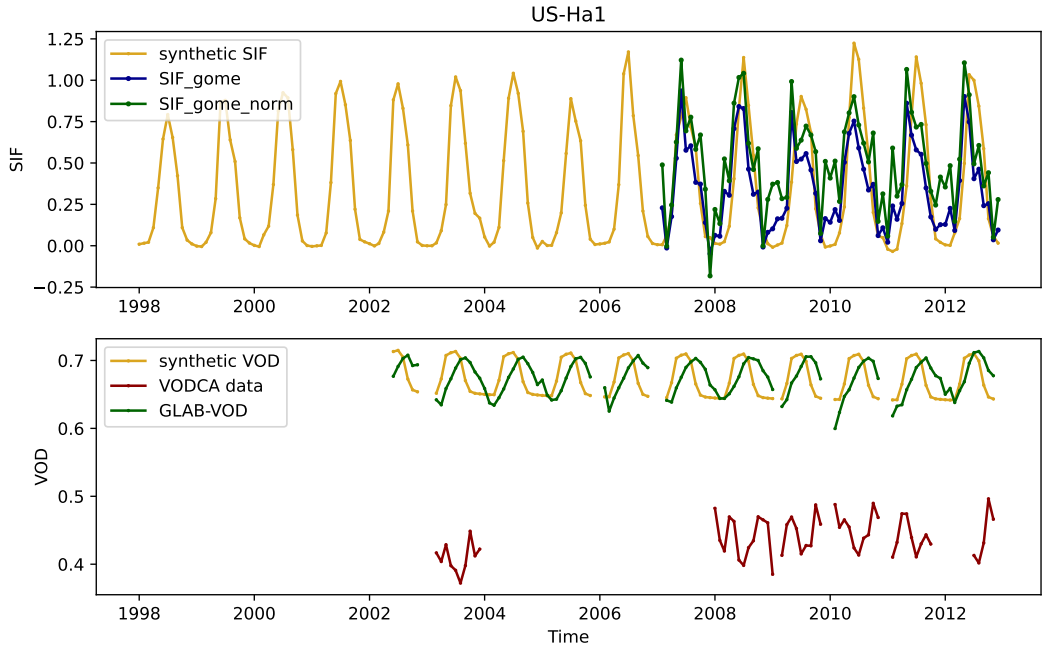


Figure S5. Synthetic and observational SIF and VOD data for US-Ha1 site

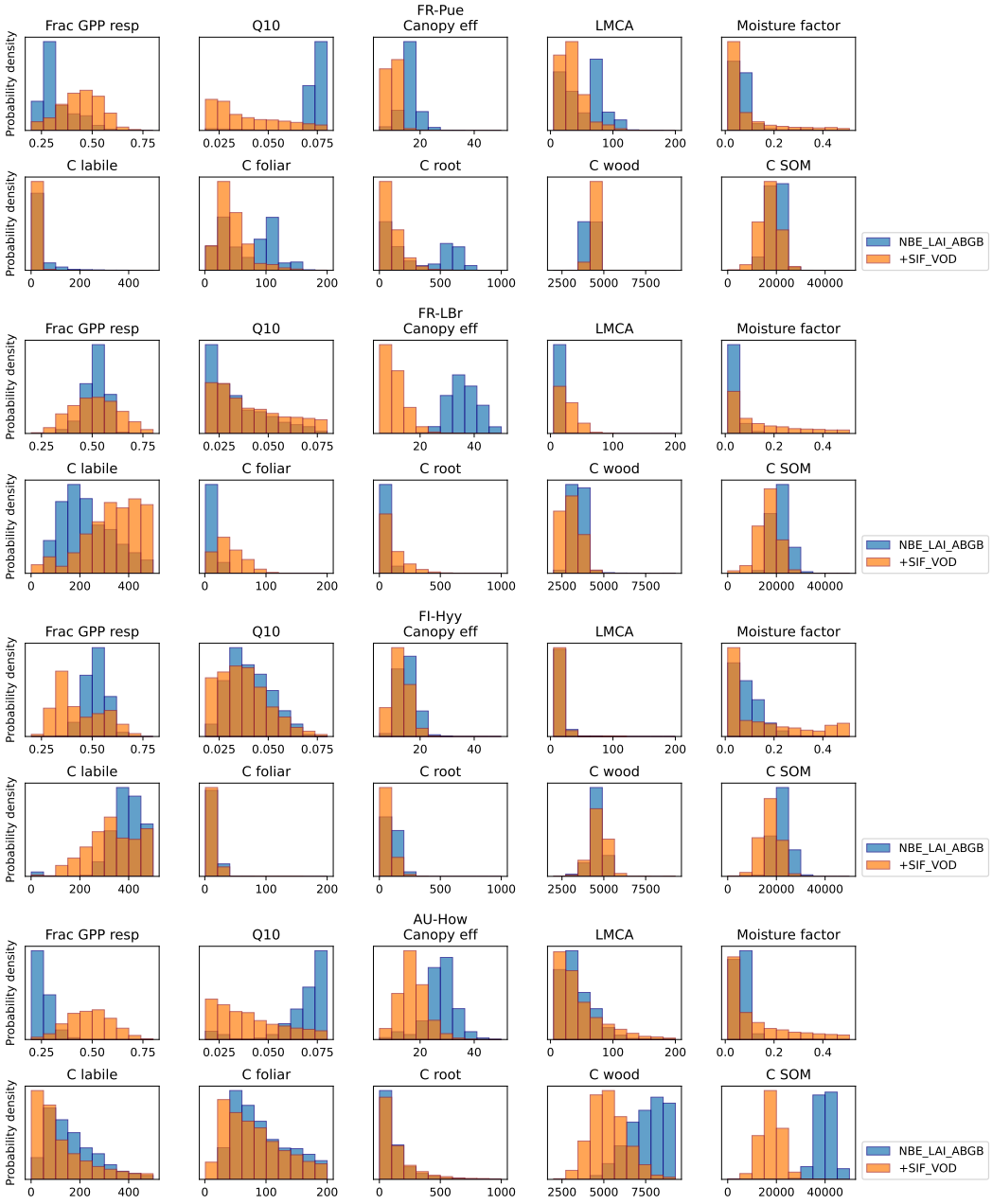


Figure S6. The effect of assimilating SIF and VOD in comparison to the "NBE_LAI_ABGB" case on a selection of CARDAMOM parameters for FR-Pue, FR-LBr, FI-Hyy, and AU-How sites.

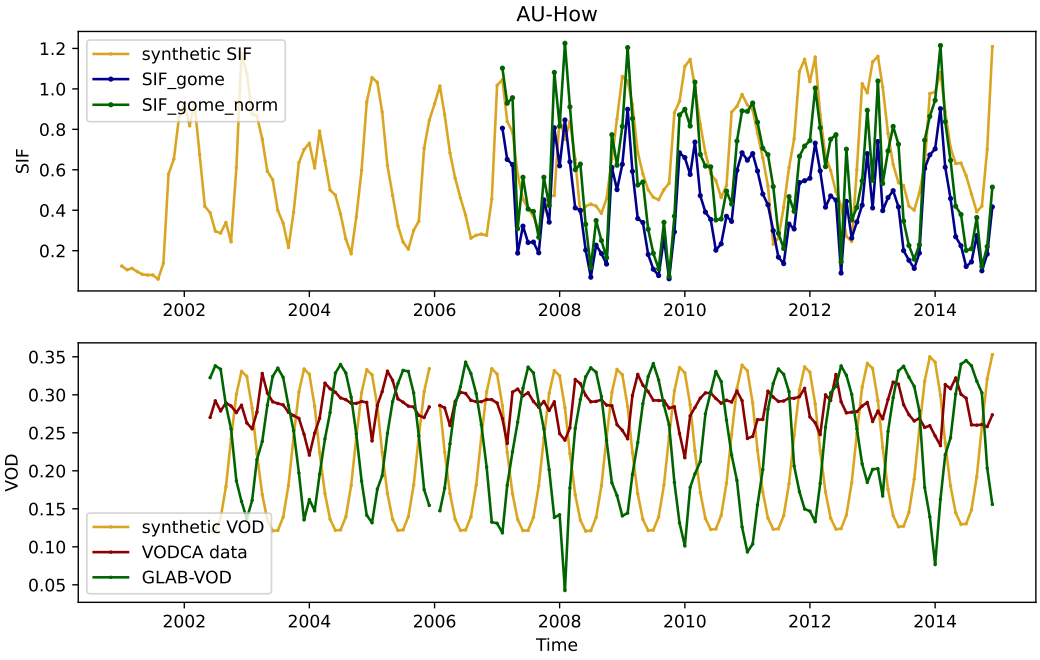


Figure S7. Synthetic and observational SIF and VOD data for AU-How site

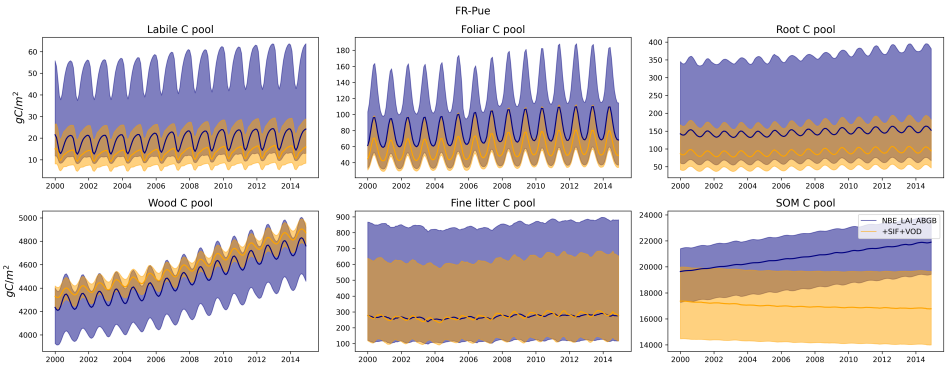


Figure S8. Carbon pools temporal dynamic for 'NBE_LAI_ABGB' and '+SIF+VOD' cases for the FR-Pue site. Median and interquantile range over the full output of all runs for a given scenario.

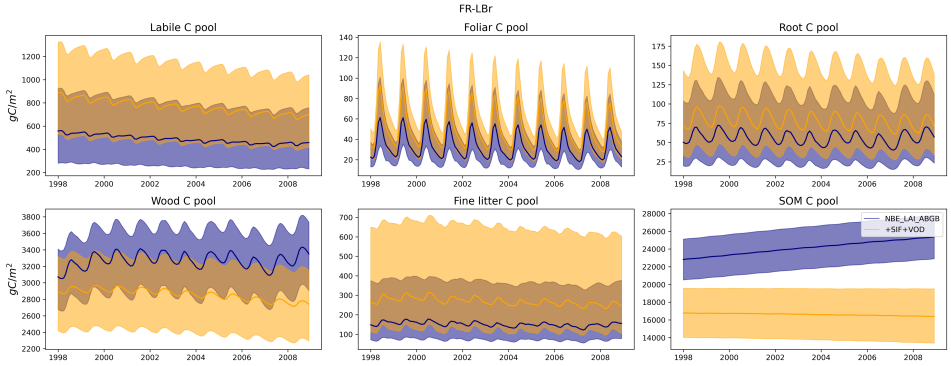


Figure S9. Carbon pools temporal dynamic for 'NBE_LAI_ABGB' and '+SIF+VOD' cases for the FR-LBr site. Median and interquantile range over the full output of all runs for a given scenario.

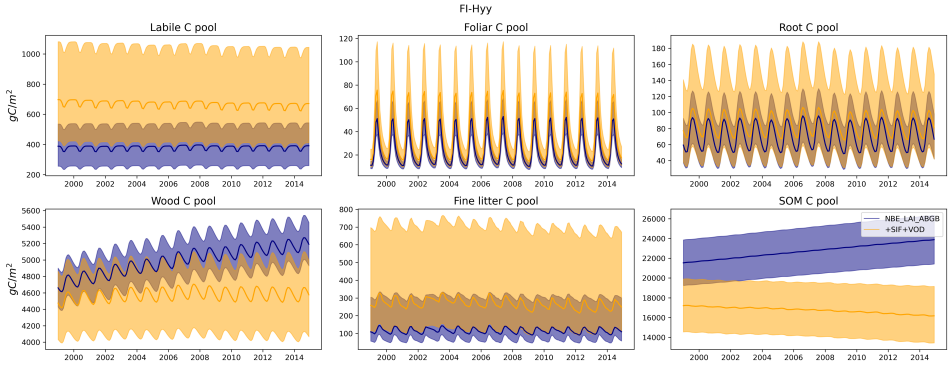


Figure S10. Carbon pools temporal dynamic for 'NBE_LAI_ABGB' and '+SIF+VOD' cases for the FI-Hyy site. Median and interquantile range over the full output of all runs for a given scenario.

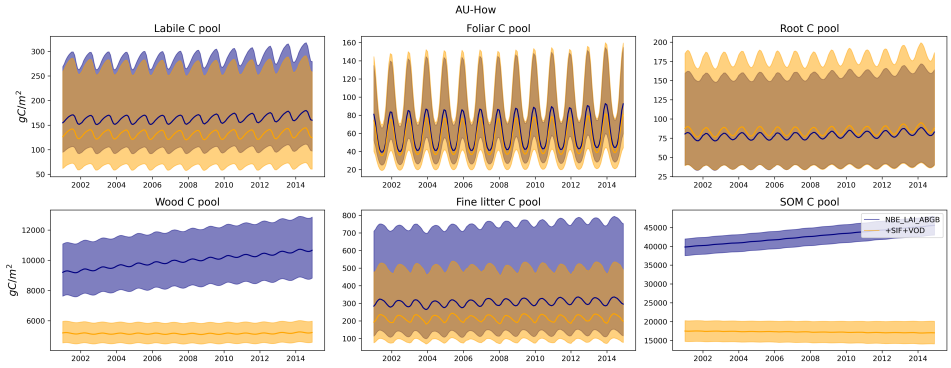


Figure S11. Carbon pools temporal dynamic for 'NBE_LAI_ABGB' and '+SIF+VOD' cases for the AU-How site. Median and interquantile range over the full output of all runs for a given scenario.

Figure 1.

FLUXNET vs. model GPP R^2

US-Ha1 | FR-Pue | FR-LBr | FI-Hyy | AU-How

R^2

-0.58 | -3.55 | -2.12 | -0.43 | -4.31

DT partitioning

Assimilated data
- ABGB_LAI_NBE
- LAI_ABGB
- +SIF_synth
- +VOD_synth
- +SIF+VOD
- None

FLUXNET vs. model GPP RMSE

US-Ha1 | FR-Pue | FR-LBr | FI-Hyy | AU-How

RMSE, gC/m^2d

34.7 | 11.19 | 27.91 | 15.47 | 30.61

DT partitioning

Assimilated data
- ABGB_LAI_NBE
- LAI_ABGB
- +SIF_synth
- +VOD_synth
- +SIF+VOD
- None

Figure 2.

FLUXNET vs. model RECO R^2

US-Ha1 | FR-Pue | FR-LBr | FI-Hyy | AU-How

R^2

-1.6 | -2.74 | -3.04 | -0.64 | -3.28

DT partitioning

Assimilated data
- ABGB_LAI_NBE
- LAI_ABGB
- +SIF_synth
- +VOD_synth
- +SIF+VOD
- None

FLUXNET vs. model RECO RMSE

US-Ha1 | FR-Pue | FR-LBr | FI-Hyy | AU-How

RMSE, gC/m^2d

12.57 | 6.57 | 15.81 | 7.3 | 15.17

DT partitioning

Assimilated data
- ABGB_LAI_NBE
- LAI_ABGB
- +SIF_synth
- +VOD_synth
- +SIF+VOD
- None

Figure 3.

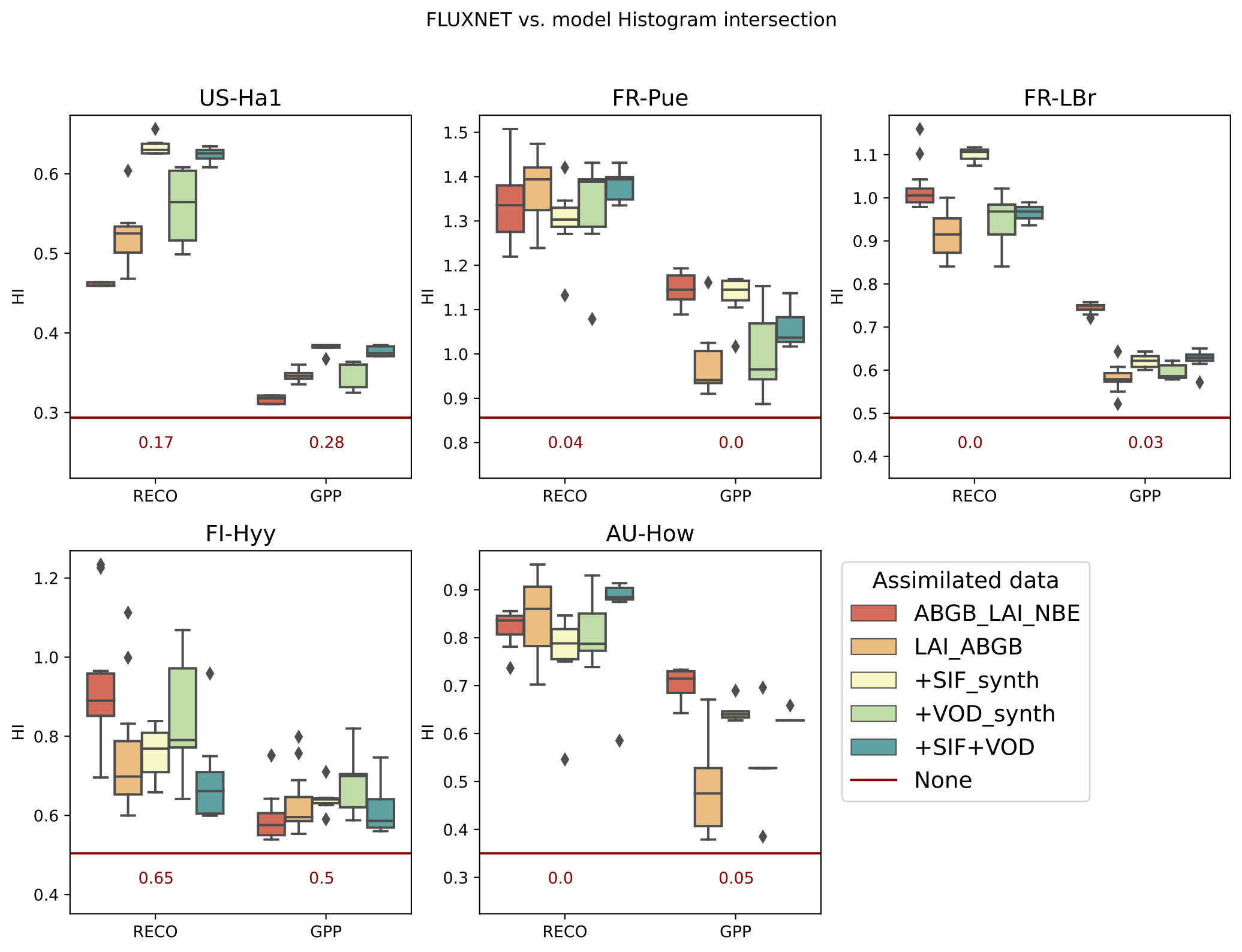

FLUXNET vs. model Histogram intersection
US-Ha1
FR-Pue
FR-LBr
FI-Hyy
AU-How
HI
RECO
GPP
0.17
0.28
0.04
0.0
0.0
0.03
0.65
0.5
0.0
0.05
Assimilated data
ABGB_LAI_NBE
LAI_ABGB
+SIF_synth
+VOD_synth
+SIF+VOD
None

Figure 4.

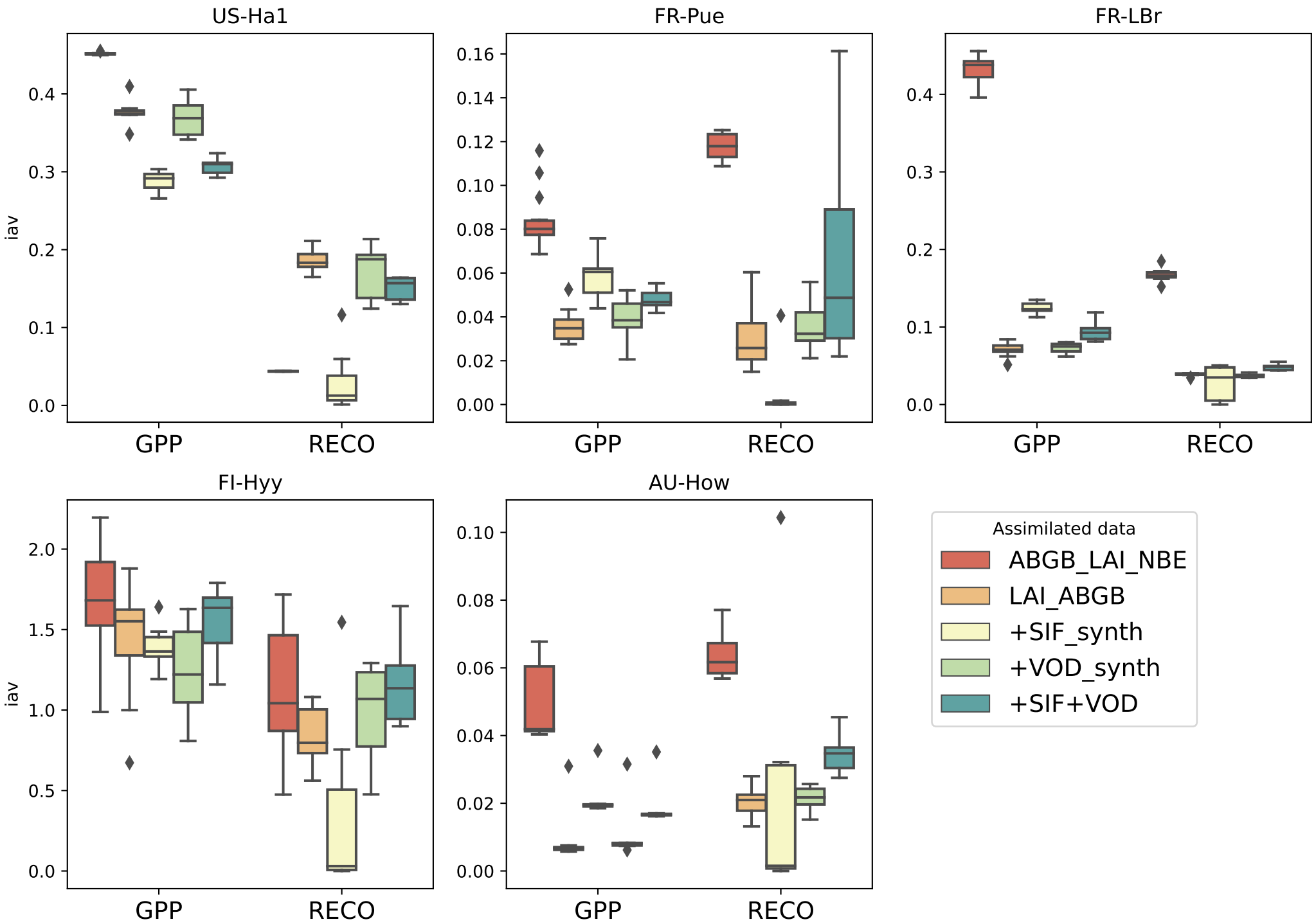

FLUXNET vs. model Interannual Anomaly Metric ratio
US-Ha1
FR-Pue
FR-LBr
FI-Hyy
AU-How
iav
GPP
RECO
Assimilated data
ABGB_LAI_NBE
LAI_ABGB
+SIF_synth
+VOD_synth
+SIF+VOD

Figure 5.

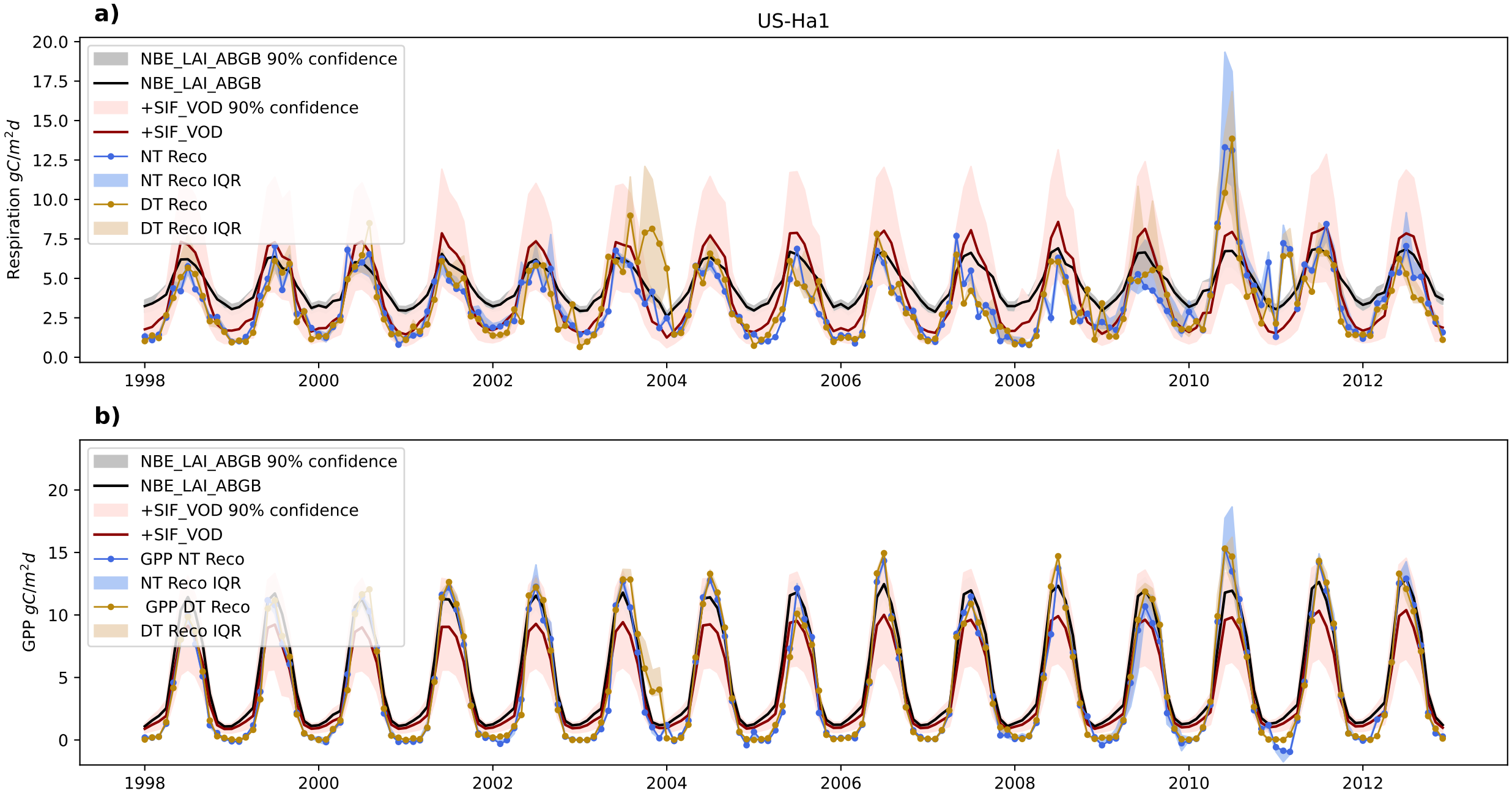

a)
US-Ha1
NBE_LAI_ABGB 90% confidence
NBE_LAI_ABGB
+SIF_VOD 90% confidence
+SIF_VOD
NT Reco
NT Reco IQR
DT Reco
DT Reco IQR
Respiration gC/m²d
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
1998
2000
2002
2004
2006
2008
2010
2012
b)
NBE_LAI_ABGB 90% confidence
NBE_LAI_ABGB
+SIF_VOD 90% confidence
+SIF_VOD
GPP NT Reco
NT Reco IQR
GPP DT Reco
DT Reco IQR
GPP gC/m²d
20
15
10
5
0
1998
2000
2002
2004
2006
2008
2010
2012

Figure 6.

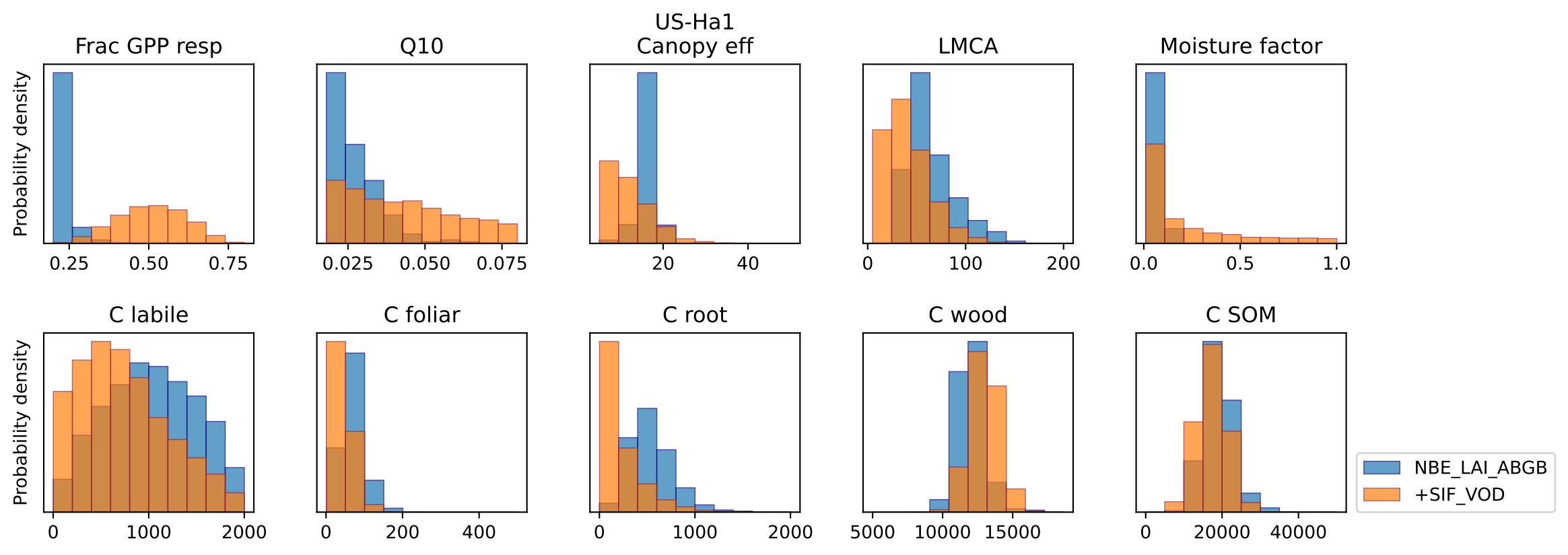

US-Ha1
Frac GPP resp
Q10
Canopy eff
LMCA
Moisture factor
C labile
C foliar
C root
C wood
C SOM
Probability density
Probability density
0.25
0.50
0.75
0.025
0.050
0.075
20
40
0
100
200
0.0
0.5
1.0
0
1000
2000
0
200
400
0
1000
2000
5000
10000
15000
0
20000
40000
NBE_LAI_ABGB
+SIF_VOD

Figure 7.

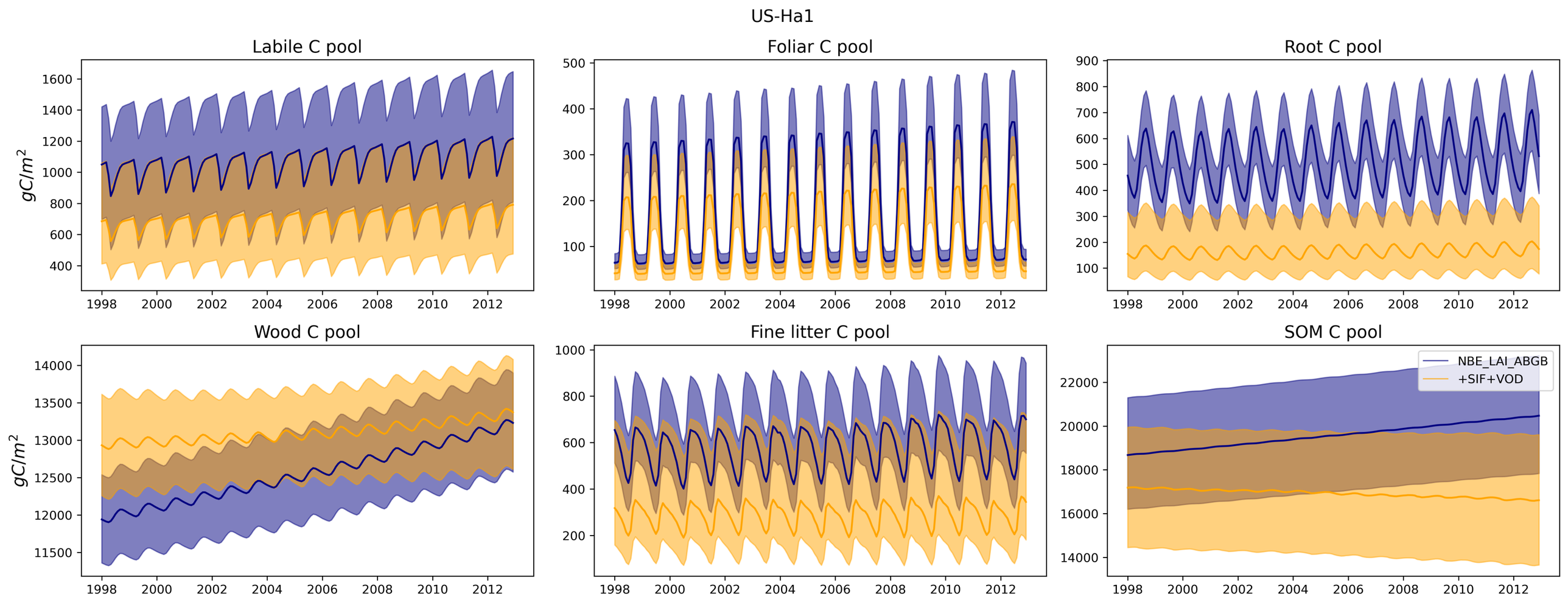

US-Ha1
Labile C pool
Foliar C pool
Root C pool
Wood C pool
Fine litter C pool
SOM C pool
gC/m2
NBE_LAI_ABGB
+SIF+VOD

Figure 8.

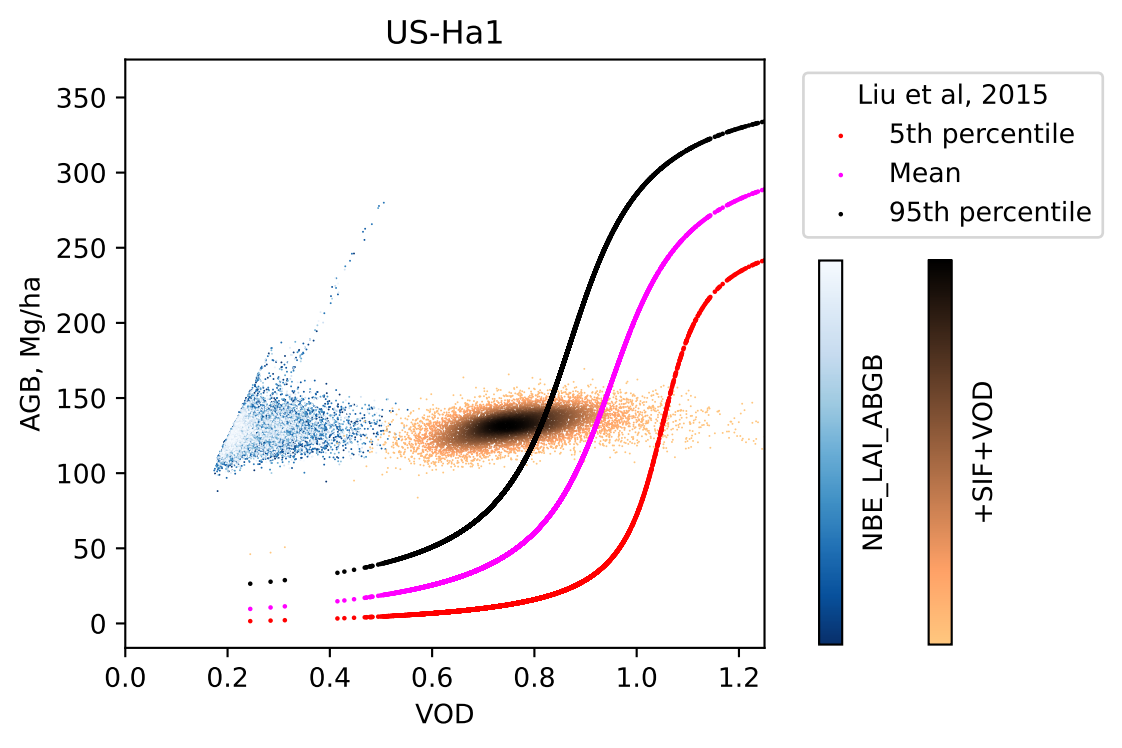

US-Ha1
AGB, Mg/ha
VOD
0
50
100
150
200
250
300
350
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Liu et al, 2015
5th percentile
Mean
95th percentile
NBE_LAI_ABGB
+SIF+VOD

Figure 9.

FLUXNET vs. model RECO R^2

US-Ha1 FR-Pue FR-LBr FI-Hyy AU-How

R^2

-1.6 -2.74 -3.04 -0.64 -3.28

DT partitioning

Assimilated data
+SIF+VOD 100%
80%
60%
40%
20%
double unc
4x unc
half unc

FLUXNET vs. model RECO RMSE

US-Ha1 FR-Pue FR-LBr FI-Hyy AU-How

RMSE, gC/m^2d

12.57 6.57 15.81 7.3 15.17

DT partitioning

Assimilated data
+SIF+VOD 100%
80%
60%
40%
20%
double unc
4x unc
half unc

Supplementary Figure 1.

FLUXNET vs. model GPP R^2

US-Ha1 | FR-Pue | FR-LBr | FI-Hyy | AU-How

R^2

-0.58 | -3.55 | -2.12 | -0.43 | -4.31

DT partitioning

Assimilated data
+SIF_synth
+SIF_data
+SIF_data_n
+VOD_synth
+VOD_data_X
+VOD_data_L
None

FLUXNET vs. model GPP RMSE

US-Ha1 | FR-Pue | FR-LBr | FI-Hyy | AU-How

RMSE, gC/m^2d

34.7 | 11.19 | 27.91 | 15.47 | 30.61

DT partitioning

Assimilated data
+SIF_synth
+SIF_data
+SIF_data_n
+VOD_synth
+VOD_data_X
+VOD_data_L
None

Supplementary Figure 2.

FLUXNET vs. model RECO R^2

US-Ha1 | FR-Pue | FR-LBr | FI-Hyy | AU-How

R^2

-1.6 | -2.74 | -3.04 | -0.64 | -3.28

DT partitioning

Assimilated data
+SIF_synth
+SIF_data
+SIF_data_n
+VOD_synth
+VOD_data_X
+VOD_data_L
None

FLUXNET vs. model RECO RMSE

US-Ha1 | FR-Pue | FR-LBr | FI-Hyy | AU-How

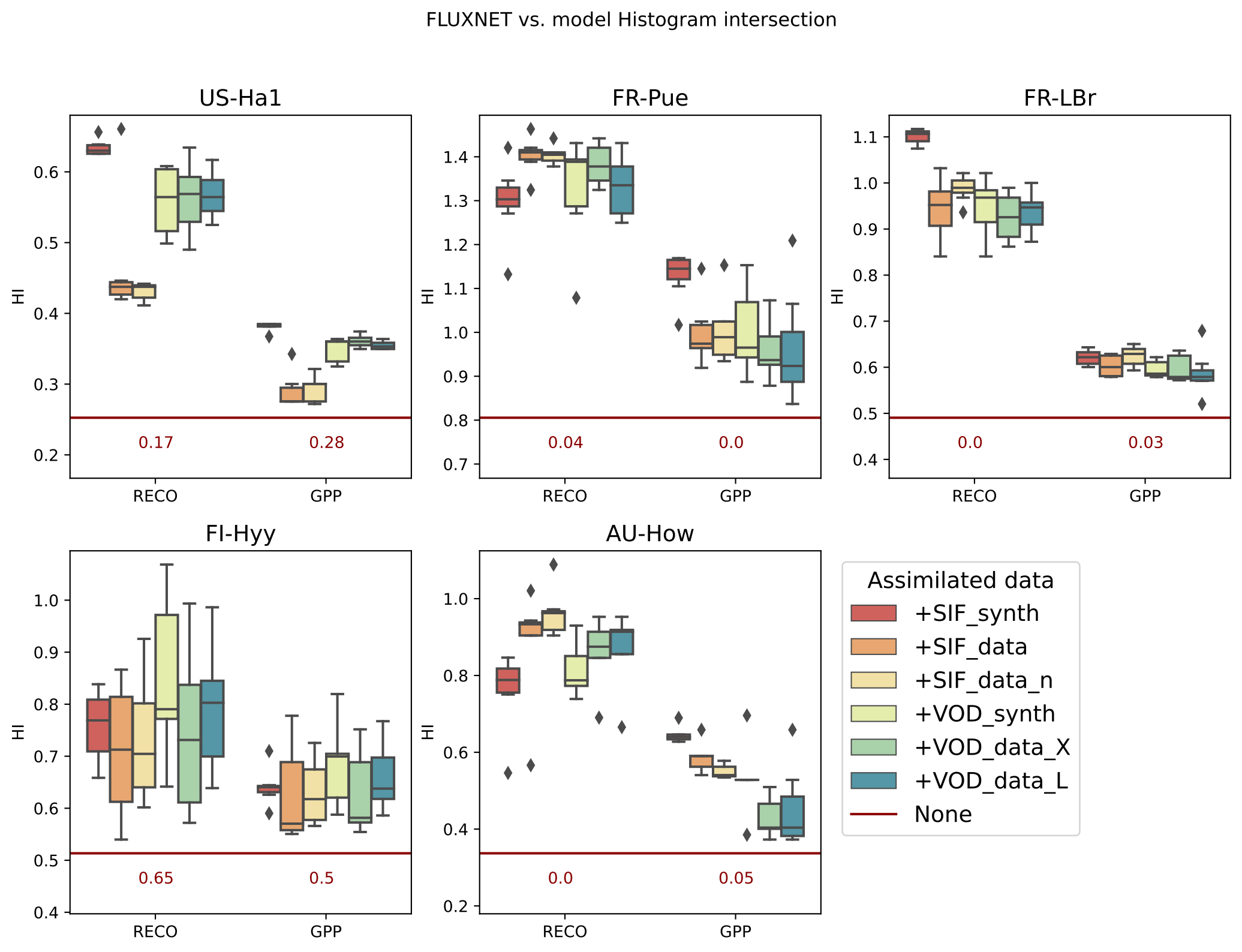
RMSE, gC/m^2d

12.57 | 6.57 | 15.81 | 7.3 | 15.17

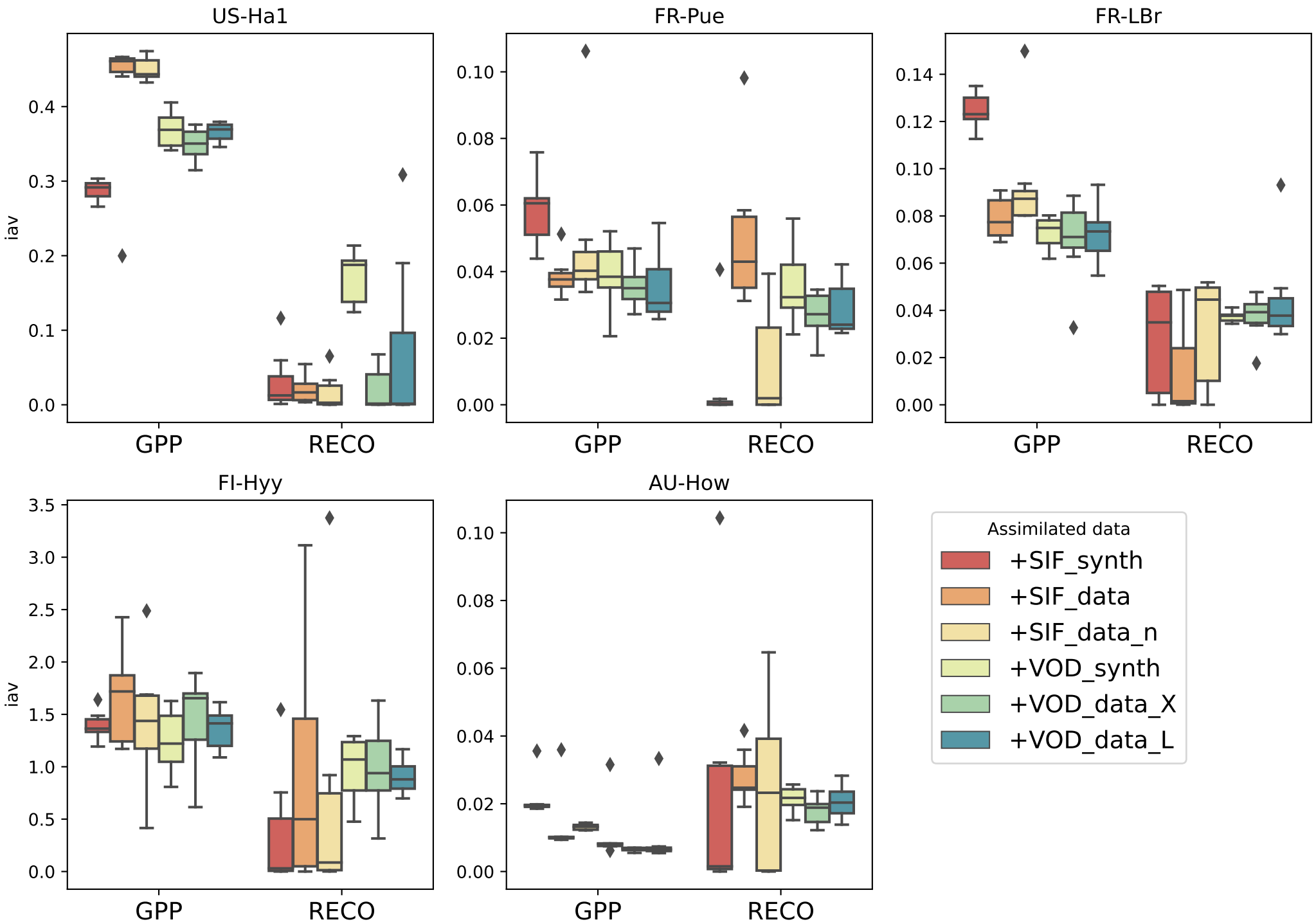
DT partitioning

Assimilated data
+SIF_synth
+SIF_data
+SIF_data_n
+VOD_synth
+VOD_data_X
+VOD_data_L
None

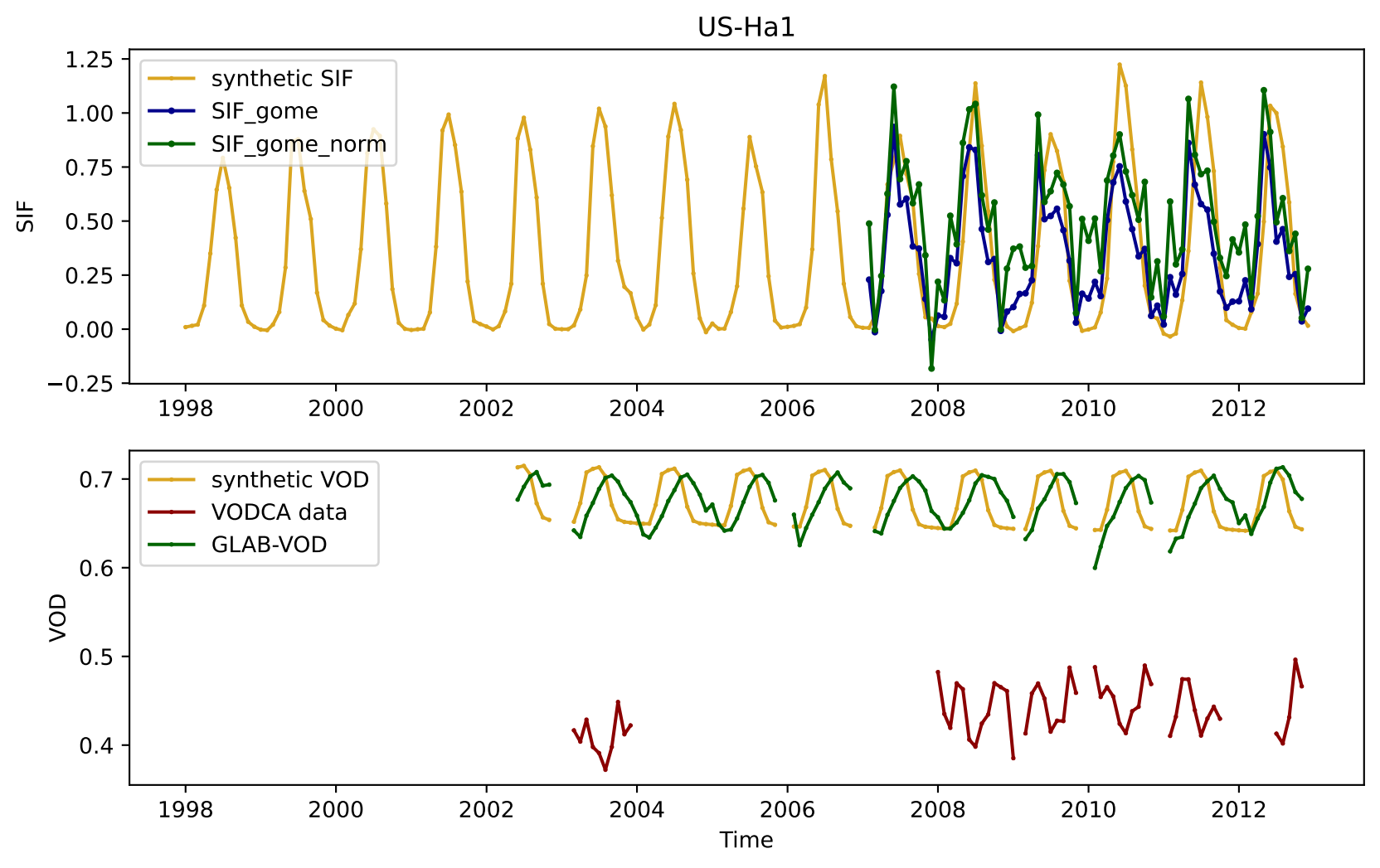
Supplementary Figure 3.


FLUXNET vs. model Histogram intersection
US-Ha1
FR-Pue
FR-LBr
FI-Hyy
AU-How
HI
RECO
GPP
0.17
0.28
0.04
0.0
0.0
0.03
0.65
0.5
0.0
0.05
Assimilated data
+SIF_synth
+SIF_data
+SIF_data_n
+VOD_synth
+VOD_data_X
+VOD_data_L
None

Supplementary Figure 4.


FLUXNET vs. model Interannual Anomaly Metric ratio
US-Ha1
FR-Pue
FR-LBr
FI-Hyy
AU-How
iav
GPP
RECO
Assimilated data
+SIF_synth
+SIF_data
+SIF_data_n
+VOD_synth
+VOD_data_X
+VOD_data_L

Supplementary Figure 5.


US-Ha1
synthetic SIF
SIF_gome
SIF_gome_norm
SIF
1.25
1.00
0.75
0.50
0.25
0.00
−0.25
1998
2000
2002
2004
2006
2008
2010
2012
synthetic VOD
VODCA data
GLAB-VOD
VOD
0.7
0.6
0.5
0.4
Time

Supplementary Figure 6.

FR-Pue

Frac GPP resp | Q10 | Canopy eff | LMCA | Moisture factor

C labile | C foliar | C root | C wood | C SOM

FR-LBr

Frac GPP resp | Q10 | Canopy eff | LMCA | Moisture factor

C labile | C foliar | C root | C wood | C SOM

FI-Hyy

Frac GPP resp | Q10 | Canopy eff | LMCA | Moisture factor

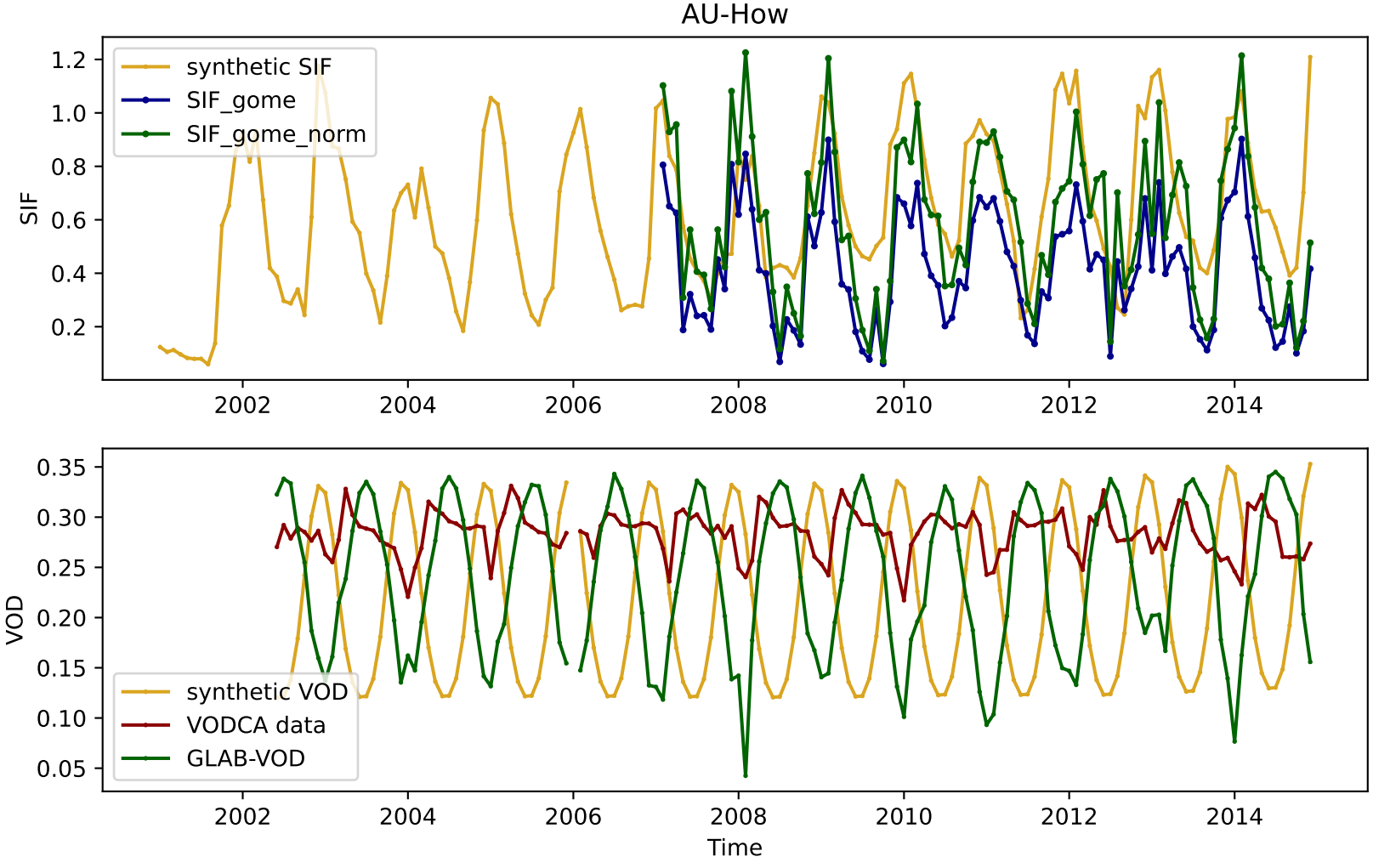
C labile | C foliar | C root | C wood | C SOM

AU-How

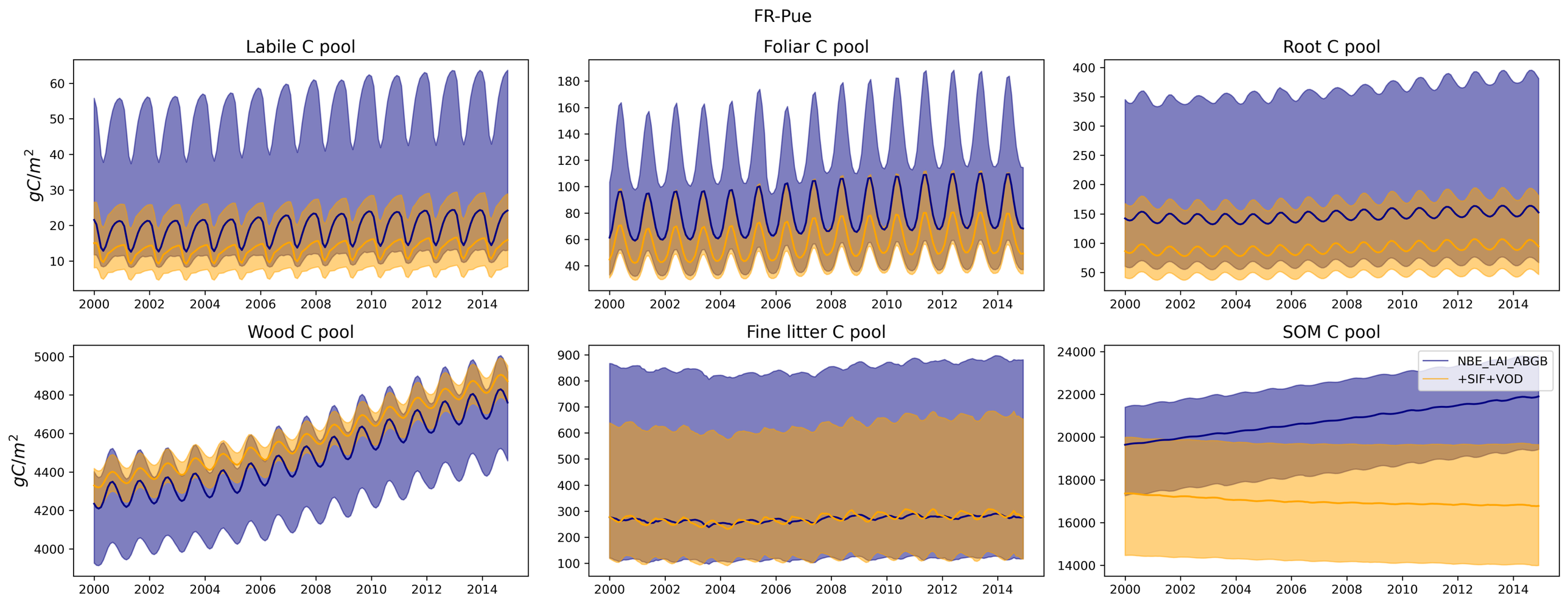
Frac GPP resp | Q10 | Canopy eff | LMCA | Moisture factor

C labile | C foliar | C root | C wood | C SOM

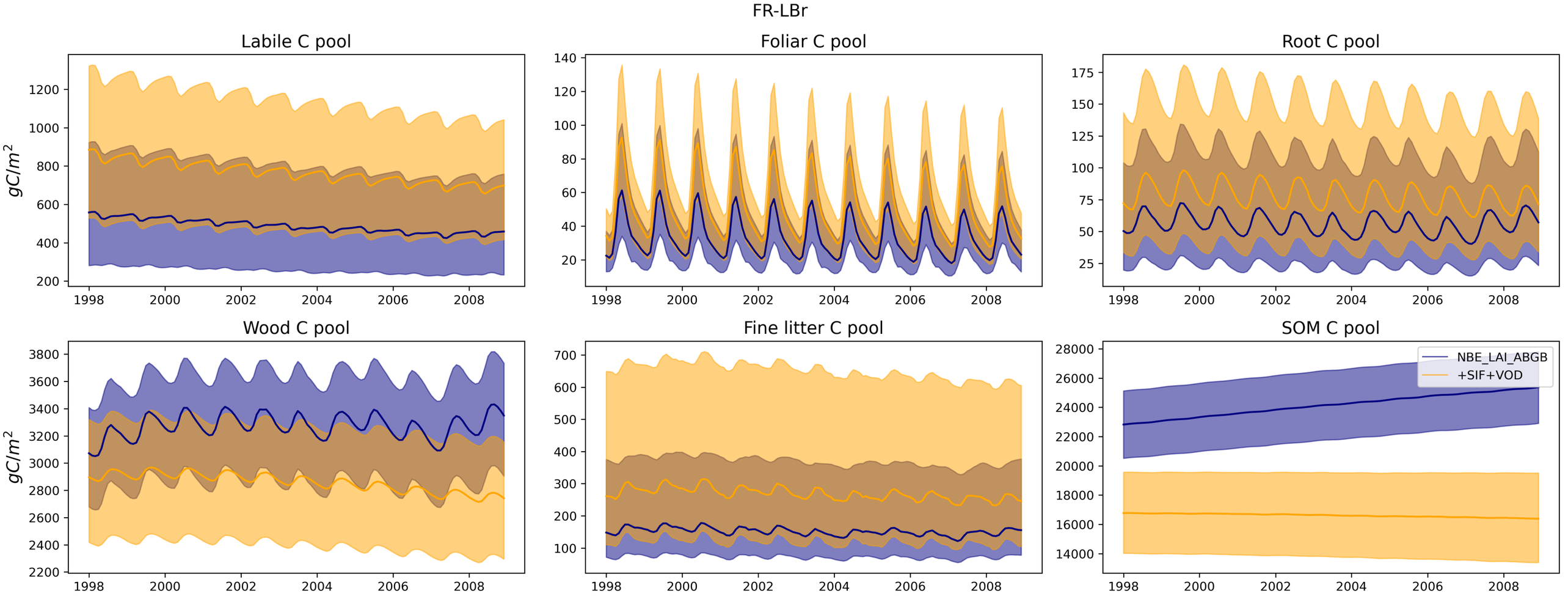
Supplementary Figure 7.


AU-How
synthetic SIF
SIF_gome
SIF_gome_norm
SIF
1.2
1.0
0.8
0.6
0.4
0.2
2002
2004
2006
2008
2010
2012
2014
synthetic VOD
VODCA data
GLAB-VOD
VOD
0.35
0.30
0.25
0.20
0.15
0.10
0.05
Time

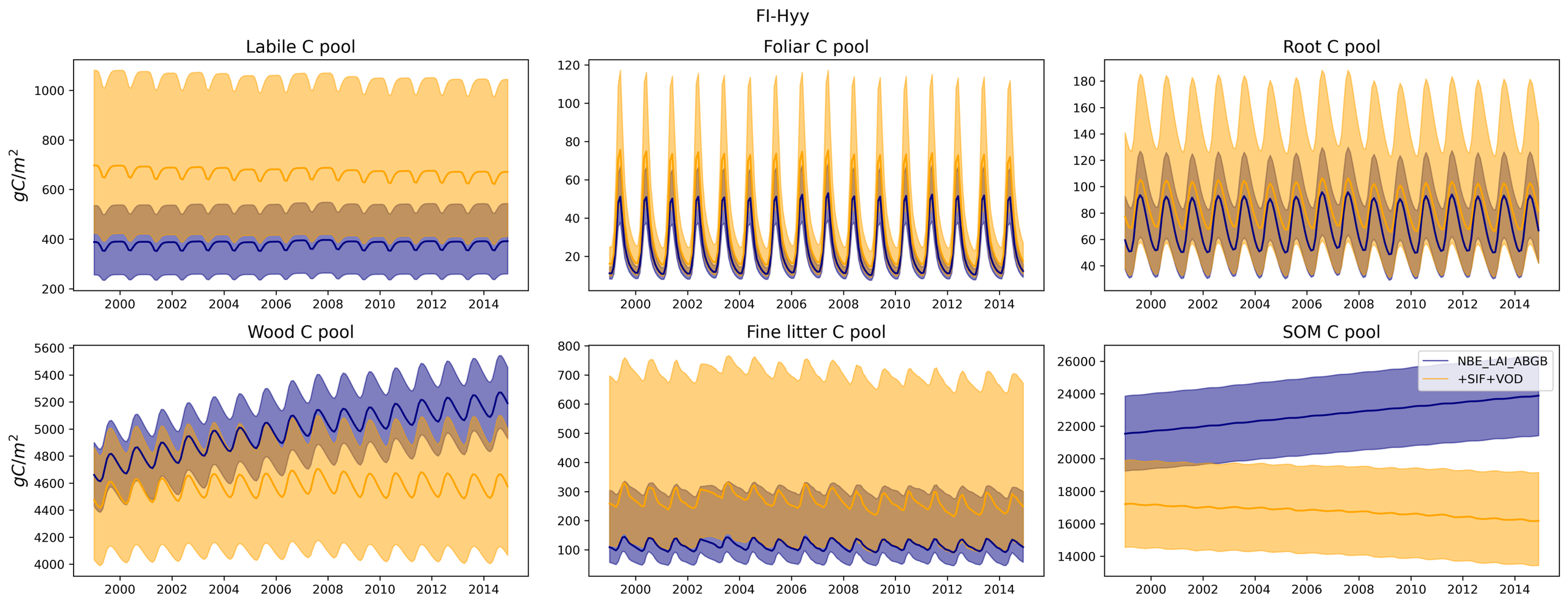
Supplementary Figure 8.


FR-Pue
Labile C pool
Foliar C pool
Root C pool
Wood C pool
Fine litter C pool
SOM C pool
gC/m2
NBE_LAI_ABGB
+SIF+VOD

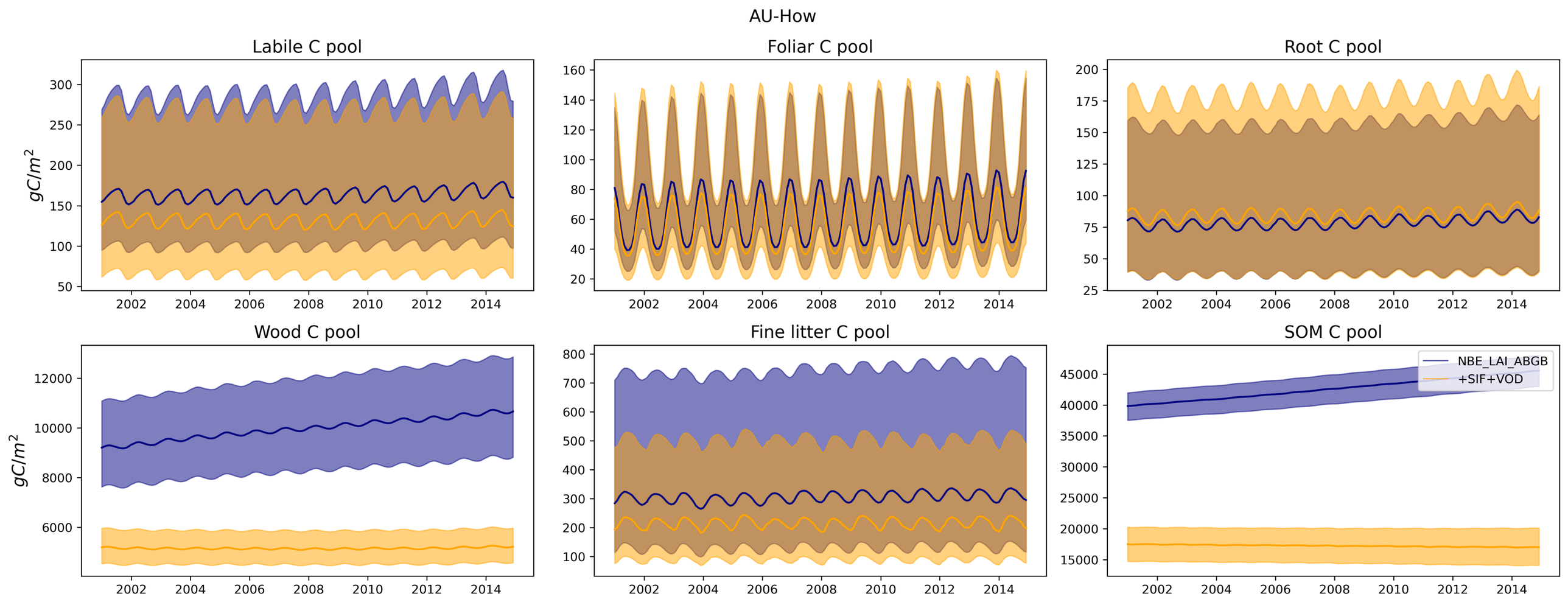
Supplementary Figure 9.


FR-LBr
Labile C pool
Foliar C pool
Root C pool
Wood C pool
Fine litter C pool
SOM C pool
gC/m2
NBE_LAI_ABGB
+SIF+VOD

Supplementary Figure 10.


FI-Hyy
Labile C pool
Foliar C pool
Root C pool
Wood C pool
Fine litter C pool
SOM C pool
gC/m2
NBE_LAI_ABGB
+SIF+VOD

Supplementary Figure 11.


AU-How
Labile C pool
Foliar C pool
Root C pool
Wood C pool
Fine litter C pool
SOM C pool
gC/m^2
NBE_LAI_ABGB
+SIF+VOD