Improved process representation of leaf phenology significantly shifts climate sensitivity of ecosystem carbon balance

Alexander J. Norton[1], A. Anthony Bloom[1], Nicholas C. Parazoo[1], Paul A. Levine[1], Shuang Ma[1], Renato K. Braghiere[2,1], and Luke T. Smallman[3]

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91101, USA
[2]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena 91125, USA
[3]School of GeoSciences and NCEO, University of Edinburgh, Edinburgh, UK

Correspondence: Alexander J. Norton (alexander.j.norton@jpl.nasa.gov), A. Anthony Bloom (alexis.a.bloom.jpl.nasa.gov)

Abstract.

Terrestrial carbon cycle models are routinely used to determine the response of the land carbon sink under expected future climate change, yet these predictions remain highly uncertain. Increasing the realism of processes in these models may help with predictive skill but any such addition should be confronted with observations and evaluated in the context of the aggregate behavior of the carbon cycle. Here, two formulations for leaf area index (LAI) phenology are coupled to the same terrestrial biosphere model, one is climate agnostic and the other incorporates direct environmental controls on both timing and growth. Each model is calibrated simultaneously to observations of LAI, net ecosystem exchange (NEE), and biomass using the CARbon DAta-MOdel fraMework (CARDAMOM) and validated against withheld data, including eddy covariance estimates of gross primary productivity (GPP) and ecosystem respiration (R_e), across six ecosystems from the tropics to high-latitudes. Both model formulations show similar predictive skill for LAI and NEE. However, with the addition of direct environmental controls on LAI, the integrated model explains 22% more variability in GPP and R_e, and reduces biases in these fluxes by 58% and 77%, respectively, while also predicting more realistic annual litterfall rates, due to changes in carbon allocation and turnover. We extend this analysis to evaluate the inferred climate sensitivity of LAI and NEE with the new model and show that the added complexity shifts the sign, magnitude, and seasonality of NEE sensitivity to precipitation and temperature. This highlights the benefit of process complexity when inferring underlying processes from Earth observations and representing the climate response of the terrestrial carbon cycle.

1 Introduction

Terrestrial ecosystems play a critical role in the Earth's climate system due to their varied couplings to and feedbacks between carbon, water, and energy with the atmosphere. Improving our ability to quantify and predict the response of terrestrial ecosystems to climate is essential to advancing our understanding of these feedbacks and predicting future climate change (Booth et al., 2012). Despite their importance, there remains considerable uncertainty in our understanding of the terrestrial carbon cycle, undermining our ability to make accurate predictions of future carbon-climate feedbacks (Friedlingstein et al., 2014; Huntzinger et al., 2017; Piao et al., 2013).

Vegetation phenology is a key component of terrestrial ecosystem dynamics as it is directly linked to key processes in the carbon, water, and energy cycles (e.g. photosynthesis, autotrophic respiration, evapotranspiration, and surface albedo), making it an area of focus in understanding the climate response of ecosystems. Phenology refers to the timing of periodic events in plant development such as reproduction, bud burst, canopy senescence, activity-dormancy cycles, and carbon allocation. Quantitatively, the leaf area index (LAI) represents the one-sided surface area of leaves per area of ground surface. LAI is a cornerstone biophysical quantity for monitoring vegetation phenology, as it can be observed globally from space, and for representing the canopy in terrestrial biosphere models (TBMs), a key component of Earth system models (Sellers et al., 1997). LAI mediates the canopy interception of radiation, thus it directly controls processes such surface albedo and the rates of photosynthesis and transpiration. Indirectly, LAI also has significant impacts such as influencing how much precipitation reaches the soil surface altering plant available water and evaporation from the soil and canopy surfaces. It also represents the mass of foliar carbon in the canopy, coupling these processes to the cycling of carbon within the plant and litterfall that supplies carbon to the soil (Richardson et al., 2013).

A variety of concepts have been used to represent LAI dynamics in TBMs, including ecohydrological equilibrium (Yang et al., 2018), optimality principles such as the maximization of plant net carbon gain (Caldararu et al., 2014; Manzoni et al., 2015), direct carbon-supply (Xin et al., 2020), demand for growth (Schiestl-Aalto et al., 2015), and approaches that consider climate and biophysical controls more directly (Jolly et al., 2005; Knorr et al., 2010; Stöckli et al., 2008). Mechanistic modeling approaches are lacking, a reflection of our limited fundamental understanding of processes such as bud burst, leaf longevity, and canopy senescence and their variability across species (Cooke et al., 2012). Observations have shown that the dynamics of LAI are correlated with environmental variables (Clelend et al., 2007); in particular temperature, water availability, and photoperiod (Delpierre et al., 2016; Iio et al., 2014; Richardson et al., 2013), although variability also occurs within and across species for a given climate (Cole and Sheldon, 2017; Marchand et al., 2020). In the absence of mechanistic understanding, it is important that model formulations are generalized and calibrated to available observations (e.g. Wheeler and Dietze, 2021). Therefore, many TBMs use semi-empirical representations of LAI which depend upon an understanding of these correlations. While many of these models have shown fidelity in representing LAI dynamics (e.g. Stöckli et al., 2008), vegetation phenology remains a large source of uncertainty in models and is therefore an ongoing area of research (Migliavacca et al., 2012; Richardson et al., 2012; Seiler et al., 2022).

In the context of carbon-climate feedbacks, it is critical to understand the role of LAI phenology in mediating the carbon cycle, particularly net ecosystem exchange of carbon (NEE) (Richardson et al., 2012). However, few studies have investigated whether a more complex model representation of LAI can actually improve predictions of NEE or how these improvements affect the sensitivity of the terrestrial carbon balance to climate. These additional steps are needed to evaluate how the representation of specific processes in models ultimately affect the integrated response of NEE to climate (Fisher and Koven, 2020). Neglecting these steps runs the risk of biased predictions of future carbon-climate feedbacks.

In this study we use a Bayesian parameter data assimilation system to generate a data-informed representation of LAI, its coupling to NEE, and the climate sensitivity of both LAI and NEE. Data assimilation or model-data fusion (MDF) provides a framework for systematically combining observations with a model (Rayner et al., 2019). For understanding a complex system

like the terrestrial carbon cycle, MDF is a useful approach to improve model performance and mechanistic understanding by constraining the diverse set of processes, and their interactions, contributing to carbon exchange (Fisher and Koven, 2020; MacBean et al., 2016).

For the present study, we consider two key aspects of model uncertainty: (i) that the model formulation must accurately represent the main processes that govern LAI, NEE, and their response to climate (Schwalm et al., 2019), and (ii) that the parameters of the model must be appropriately assigned (Prentice et al., 2015). For a given model, MDF can provide a parameterization that is statistically consistent with observations and their uncertainties. When applied equally across multiple model structures, MDF can be used to evaluate different model structures and their impact on the data-informed processes (e.g. Famiglietti et al., 2021). Here, we investigate two model formulations for LAI and perform MDF at six flux tower sites distributed across diverse ecosystems from the tropics to high-latitudes (Baldocchi, 2008). We implement a prognostic, climate-sensitive LAI model into a TBM and benchmark this against an empirical diagnostic LAI model used in a previous version of the same TBM (Bloom and Williams, 2015; Quetin et al., 2020). We constrain both models using multiple observations of carbon states and fluxes, and then use these data-informed models to infer the climate sensitivity of NEE. The main objectives of this study are to: (i) Investigate the impact of a more complex process representation of LAI on predicting LAI and NEE dynamics in a MDF system, (ii) evaluate the impact of a more complex process representation of LAI on inferring the processes underlying NEE: GPP and R_e, and (iii) evaluate how a change in LAI process representation affects the climate sensitivity of the terrestrial carbon cycle at seasonal and annual time-scales.

2 Methods

2.1 Study Sites

The study is focused at six sites distributed from the tropics to high-latitudes (Fig. 1) that are part of a global network of eddy covariance flux sites, FLUXNET (Pastorello et al., 2020). These sites cover a range of climate zones and phenological strategies (Table 1), allowing for more robust model evaluation and climate sensitivity analysis in the global context. Following Famiglietti et al. (2021), we selected these sites based on the following criteria: (i) meteorological forcing data availability, (ii) observational data availability including repeat woody biomass observations and eddy covariance measurements of carbon dioxide and water vapor fluxes, (iii) temporal coverage of at least ten years, (iv) no intensive human management (e.g., agriculture or logging), and (v) vegetation is dominated by C3 photosynthetic pathway.

2.2 Model-data fusion

To quantitatively evaluate the impacts of the process representation of LAI on the net carbon balance and its climate sensitivity, we utilized the CARbon DAta-MOdel fraMework (CARDAMOM, Bloom and Williams, 2015; Bloom et al., 2016). CARDAMOM is a Bayesian MDF system, used to retrieve time-invariant parameters and initial conditions for the Data Assimilation Linked ECosystem model (DALEC) and has been used widely as a diagnostic tool to infer stocks and fluxes of carbon
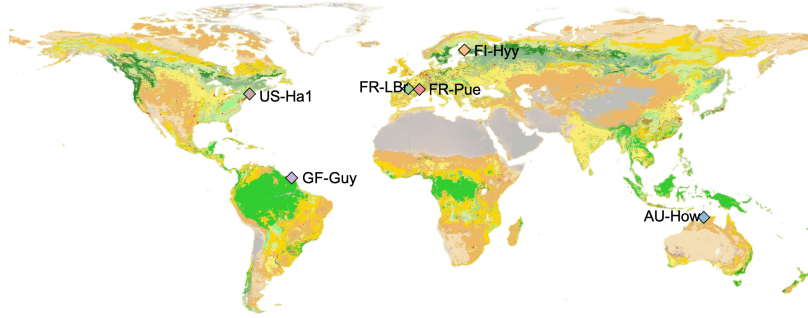
Figure 1. Study site locations overlaid onto the land cover type map from the Moderate Resolution Imaging Spectroradiometer, MODIS, 2016 (MCD12C1, Friedl and Sulla-Menashe, 2015).

Table 1. Description of the study sites including the site name, FLUXNET code, dominant plant functional type (PFT), location, mean annual temperature (MAT) and mean annual total precipitation (MAP).

Site	Id	Dominant PFT	Lat, Lon (°)	MAT (°C)	MAP (mm)	Study Period
Howard Springs	AU-How	Tropical woody savanna	-12.49, 131.15	27.0	1449	2001-2014
Hyytiala	FI-Hyy	Boreal evergreen needleleaf	61.85, 24.29	3.8	709	1999-2014
Le Bray	FR-LBr	Temperate evergreen needleleaf	44.72, -0.77	13.6	900	1998-2008
Puechabon	FR-Pue	Temperate evergreen broadleaf	43.74, 3.60	13.5	883	2000-2014
Guyaflux	GF-Guy	Tropical evergreen broadleaf	5.28, -52.92	25.7	3041	2004-2018
Harvard Forest	US-Ha1	Temperate deciduous broadleaf	42.54, -72.17	6.2	1071	1998-2012

90   and water (Bloom et al., 2020; Quetin et al., 2020; Yin et al., 2020; Smallman et al., 2021). CARDAMOM has the capability to assimilate a diverse range of observations (Bloom et al., 2020; Famiglietti et al., 2021) and shows comparable performance to more complex terrestrial biosphere models when it is constrained by observations (Quetin et al., 2020). It has advantages over other MDF frameworks as it does not rely on definitions of plant functional types or on steady-state assumptions.

    CARDAMOM is also capable of utilizing different DALEC model formulations (Famiglietti et al., 2021). A number of
95   DALEC model versions have been developed for various purposes (Williams et al., 2005; Famiglietti et al., 2021). Recent developments have incorporated more processes, as CARDAMOM is increasingly being used to diagnose climate sensitivity of terrestrial carbon and water cycles (Bloom et al., 2020; Ge et al., 2022; Smallman and Williams, 2019; Yang et al., 2022). With this in mind, we implemented a climate-sensitive LAI phenology model into a version of DALEC, which is described further below.

2.3 Observations and Model Forcing

Multiple types of observations were used to constrain the processes relevant to LAI, NEE, and their interactions. This helps prevent over-fitting and ensures a consistent view of the terrestrial carbon cycle is achieved between model and data (Kaminski et al., 2013). Observations used include monthly LAI, monthly NEE, and annual woody biomass. Observations of LAI were retrieved from the Earth Observation Copernicus 1 km gridded product over each site, which includes a time-varying uncertainty estimate that we utilized (Fuster et al., 2020; Verger et al., 2014). Observations of NEE using the eddy covariance technique are from the FLUXNET2015 database (Pastorello et al., 2020). The time-varying uncertainty estimate for NEE was based upon propagating instrumentation error (0.58 g C m^{-2} d^{-1} Hill et al., 2012) and temporal aggregation due to missing sub-monthly time steps. Uncertainty due to temporal aggregation was estimated based on site specific statistical models derived from sub-sampling time periods without missing values. Aggregation and instrumentation uncertainty was then combined assuming uncertainties are fully correlated. The in-situ woody biomass observations were converted into estimates of above- and below-ground biomass (ABGB) using allometric scaling based on principle investigator advice at each site. Further details on all of the observations used can be found in Famiglietti et al. (2021).

The study period at each site ranges from 11 to 16 years (Table 1). The data was split into two periods, a training window (calibration) and a prediction window (validation). The first five years was used for calibration and the remaining data was used for validation. Climate forcing data for the model consisted of downward shortwave radiation, air temperature (average 2 m minimum and maximum), precipitation, vapor pressure deficit, and atmospheric carbon dioxide concentration.

To support model evaluation, we used gross primary productivity (GPP) and ecosystem respiration (R_e) fluxes from the FLUXNET partitioned NEE (Pastorello et al., 2020). We selected GPP and R_e estimates derived using night-time partitioning, i.e. R_e is determined as a night time R_e fitted to a function of temperature extrapolated into the day time, thus, GPP is estimated as the residual between R_e and NEE. This data was withheld from the model calibration step, thus providing a stringent metric of model skill in representing the processes governing carbon cycling. There are three primary motivations for withholding this data for validation: (i) These are model-based products, (ii) the NEE observations and the GPP and R_e estimates are not wholly independent, and (ii) we only assimilate observations that can be produced directly from Earth observations to permit global application of this framework in future work.

2.4 Model Description

The DALEC model version used here has been fully described elsewhere (Bloom and Williams, 2015; Bloom et al., 2016; Quetin et al., 2020; Yang et al., 2022; Yin et al., 2020). We describe, in brief, the representation of the carbon cycle in DALEC. Full description of the water balance can be found in Bloom et al. (2020), which includes a plant-available and a plant-unavailabale soil water pool. Following this, we describe the two separate implementations of LAI phenology used in this study that are linked to same representation of carbon and water cycles. Common model parameters between the two models are shown in Table A1.

2.4.1 Carbon Balance

The carbon cycle in DALEC consists of six carbon pools (labile, foliar, wood, fine root, litter, and soil) and simulates pool transfers using ordinary differential equations. The NEE of an undisturbed ecosystem is calculated as:

135
$$NEE = (R_h + R_a) - GPP = R_e - GPP \tag{1}$$

Where GPP is the gross primary productivity, R_h is the heterotrophic respiration from litter and soil carbon, R_a is the autotrophic respiration, and R_e is the ecosystem respiration representing the sum of R_h and R_a. The representation of GPP is based on the Aggregated Canopy Model (Williams et al., 1997, ACM), with the specific implementation described in (Bloom et al., 2016). The ACM model is a parsimonious approach for representing GPP with calibration (Williams et al., 1997). It
140 requires inputs of temperature, carbon dioxide concentration, downward shortwave radiation, and LAI. The GPP simulated by ACM is then scaled by a soil moisture limitation factor that is calculated using the plant available soil water and a parameter for the wilting point.

2.4.2 LAI Phenology Models

The focus of this study is upon the representation of LAI phenology. We implemented a new model for LAI phenology into
145 DALEC, based upon the model of Knorr et al. (2010). As a benchmark, we also utilized a diagnostic LAI phenology model commonly used in CARDAMOM studies, the Combined Deciduous-Evergreen Analytical model (CDEA). These two DALEC model formulations are denoted DALEC$_{Knorr}$ and DALEC$_{CDEA}$, respectively.

CDEA Model

The CDEA model is a relatively simple model for simulating the phenology, growth, and turnover of LAI, with full descriptions
150 described elsewhere (Bloom and Williams, 2015; Famiglietti et al., 2021; Quetin et al., 2020), with the specific formulation the same as that of Bloom et al. (2020). In brief, the CDEA model computes leaf onset and leaf fall factors that govern the flux of carbon from the plant labile carbon pool (C_{lab}) to the foliar carbon pool (C_{fol}) and flux of carbon from C_{fol} to the litter carbon pool (C_{lit}), respectively. Carbon can also be supplied directly from GPP via a fractional allocation parameter. The leaf onset and leaf fall factors are based on a day-of-year approach that govern the timing of phenological events, including peak day of
155 year for labile turnover (supporting leaf growth) and for foliar turnover (controlling litterfall). A parameter that governs the leaf longevity determines how much of the canopy is turned over each year. The CDEA model is a relatively simple and generic representation of LAI phenology and consists of eight parameters as well as two initial condition parameters for C_{lab} and C_{fol} (Table B1). There is no representation of direct environmental control on the timing of phenological events (e.g. spring onset, fall senescence), hence these are fixed from year to year. However, environmental effects on GPP can propagate through to
160 LAI via changes in carbon allocation, thus allowing for the magnitude of LAI to be indirectly sensitive to climate via changes in carbon supply.

Knorr Model

The new LAI phenology model implemented into CARDAMOM is a development of the model by Knorr et al. (2010). This is a prognostic model governed by environmental constraints on the timing and growth of the canopy. Full description of the model can be found in Knorr et al. (2010). Here, we briefly describe the model and its novel implementation in DALEC which includes coupling to the carbon and water cycles. The Knorr LAI model as implemented in DALEC consists of ten parameters as well as four initial condition parameters (Table C1).

The prevailing understanding of the dynamics of LAI across global biomes is that there are three primary environmental controls: temperature, photoperiod, and water availability (Richardson et al., 2012). The Knorr model considers all three as potentially limiting factors, with the specific dynamics governed by climate and model parameterization. In addition, we couple LAI phenology to the plant carbon balance and incorporate a function for carbon supply limitation on LAI growth thus providing a fourth potentially limiting factor.

Representing Activity-Dormancy Triggers in a Population

A common approach to modeling leaf phenology is to use one or more growth triggers (or thresholds) that transition a plant into or out of an active growth state. This is problematic as it is often modeled using a discrete, binary formulation, which makes these functions unrealistic when representing a population of individuals (e.g. within a model grid cell). In reality, plants within a given population do not reach these thresholds simultaneously (Cooke et al., 2012), thus a distribution of threshold values to represent the population of individuals is more realistic, and this is likely to result in a relatively smooth transition toward the new growth state when integrated over the population. A discrete formulation is also non-differentiable, which is problematic for derivative-based MDF techniques. Knorr et al. (2010) developed a convenient solution to this problem by representing threshold parameters with a normal distribution in space.

Two temperature thresholds, one for temperature and one for photoperiod, are each represented by a cumulative normal distribution function (Φ). The multiplication of these two cumulative normal distributions gives the fraction of individuals within the population that are in an active growth state, f, as:

$$f = \Phi\left(\frac{T - T_\phi}{T_r}\right) \cdot \Phi\left(\frac{t_d - t_c}{t_r}\right) \tag{2}$$

Where T represents the air temperature memory, analogous to the growing degree days concept (Eq. 20, Knorr et al., 2010), T_ϕ is a parameter representing the mean temperature threshold for leaf onset, T_r is a parameter representing the one sigma spatial range of T_ϕ, t_d is the day length, t_c is a parameter representing the mean day length threshold for leaf senescence, t_r is a parameter representing the one sigma spatial range of t_c.

Temporal Dynamics of LAI

The temporal evolution of LAI is represented by the following ordinary differential equation:

$$\frac{dLAI(t)}{dt} = \xi \cdot \left(LAI_{max}(t) - \text{LAI(t)}\right) \cdot f - \frac{\text{LAI(t)}}{\tau_L} \cdot (1 - f) \tag{3}$$

where τ_L is a parameter describing the longevity of leaves during senescence and $\mathrm{LAI_{max}}$ is the maximum potential LAI computed as:

$$LAI_{max}(t) = \nu(\hat{\Lambda}, LAI_W) \tag{4}$$

Where ν represents a quadratically smoothed minimum function (see Eq. C2), $\hat{\Lambda}$ is a parameter describing the maximum LAI (as limited by factors such as structure), and LAI_W is the LAI based upon water availability, computed by:

$$LAI_W(t) = \frac{W \cdot \mathrm{LAI(t)}}{E \cdot \tau_W} \tag{5}$$

Where W represents the plant available soil water, E is the evapotranspiration rate, and τ_W is a parameter representing the expected length of water deficit periods tolerated before leaf shedding. We note that this differs from the original formulation for water limitation such that we use E instead of transpiration as in Knorr et al. (2010), considering this version of DALEC does not differentiate the two.

Coupling to the Carbon Balance

While the fundamentals of the Knorr model are grounded in biophysical concepts for activity-dormancy of individuals in a population of plants, it only predicts the net change in LAI. Coupling LAI dynamics to the carbon cycle requires additional assumptions which were not defined in the original model description (Knorr et al., 2010). First, in both DALEC$_{\mathrm{CDEA}}$ and DALEC$_{\mathrm{Knorr}}$, LAI is related to C$_{\mathrm{fol}}$ via a parameter for the leaf mass of carbon per area, LMA, as follows:

$$C_{fol} = \mathrm{LAI} \cdot LMA \tag{6}$$

Second, we must consider that inputs to C$_{\mathrm{fol}}$ come from plant carbon allocation and outputs go to C$_{\mathrm{lit}}$, with the rate of change in C$_{\mathrm{fol}}$ represented by:

$$\frac{dC_{fol}(t)}{dt} = F_{C,lab2fol}(t) - F_{C,fol2lit}(t) \tag{7}$$

Where $F_{C,lab2fol}$ represents the flux of carbon from C$_{\mathrm{lab}}$ to C$_{\mathrm{fol}}$, and $F_{C,fol2lit}$ represents the flux of carbon from C$_{\mathrm{fol}}$ to C$_{\mathrm{lit}}$.

The carbon supplied via net primary productivity (i.e. GPP - R$_{\mathrm{a}}$) into C$_{\mathrm{lab}}$ provides the substrate to grow new leaves. To represent $F_{C,lab2fol}$ we consider both the supply and demand of carbon for new foliar growth. The supply of labile carbon is the sum of new labile production at time t (Eq. C1) and C$_{\mathrm{lab}}$ at end of the previous time step (expressed as a flux over the time step, Δt), represented by:

$$F_{C,fol,supply}(t) = F_{labprod}(t) + \frac{\mathrm{C_{lab}}(t-1)}{\Delta t} \tag{8}$$

This formulation implies that the entire C_{lab} pool is available for foliar growth at any given time step which is consistent with findings that C_{lab} does not follow first-order decay kinetics (Martínez-Vilalta et al., 2016). We do not consider constraints on the release of stored labile carbon such as the phloem loading rate (e.g. Trugman et al., 2018).

For the demand for new foliar growth, we make the assumption that there is a gross demand flux of carbon from C_{lab} to C_{fol} when the canopy LAI is in a net growth state. Conversely, when the canopy (LAI) is in a net senescent state, there is zero gross demand flux of carbon from C_{lab}. This is represented by:

$$F_{C,fol,demand}(t) = max(0, LMA \cdot \frac{dLAI(t)}{dt}) \tag{9}$$

Where $\frac{dLAI(t)}{dt}$ is computed from Eq. 3. The actual $F_{C,lab2fol}$ is computed as the smoothed minimum of the supply and demand fluxes as follows:

$$F_{C,lab2fol}(t) = \nu\big(F_{C,fol,supply}(t), F_{C,fol,demand}(t)\big) \tag{10}$$

Where ν represents a quadratically smoothed minimum function (Eq. C2). This formulation ensures that new foliar growth only occurs when carbon substrate is available.

The litterfall flux, $F_{C,fol2lit}$, also depends on whether the canopy is in a phase of growth or senescence. Here, we incorporate an additional term that is necessary to represent litterfall. We note that when the Knorr model is in a fully active growth phase (i.e. $f = 1$,) which may occur during canopy closure or in evergreen systems, the model (Eq. 3) would predict zero LAI loss and hence zero litterfall. Observations across the major global biome types show that litterfall never goes completely to zero (Zhang et al., 2014), as leaves are being turned over constantly at a rate governed by factors such as longevity, herbivory, and disturbance, even if its not evident from ecosystem scale LAI observations (Albert et al., 2019). To overcome this limitation, we add an additional term for loss of LAI via a nominal background turnover rate (θ_{fol}). Therefore, the litterfall flux is computed as:

$$F_{C,fol2lit}(t) = \begin{cases} \theta_{fol} \cdot C_{fol}, & \text{if } \frac{dLAI(t)}{dt} > 0. \\ LMA \cdot \frac{dLAI(t)}{dt} + \theta_{fol} \cdot C_{fol}, & \text{otherwise.} \end{cases} \tag{11}$$

This formulation ensures that some litter production occurs regardless of the growth-senescence state of the Knorr LAI model while ensuring conservation of mass.

2.4.3 Optimization Algorithm

Following previous CARDAMOM efforts, we jointly retrieve the probability distribution of DALEC time-invariant process parameters and initial state conditions (henceforth vector x) given the observational constraints (henceforth vector O) using a standard Bayesian inference formulation, where;

$$p(x|O) \; p(x)p(O|x); \tag{12}$$

$p(x)$ is the prior probability distribution of x, and $p(O|x)$ is proportional to the likelihood of x given observations O ($L(x|O)$). The prior probability of x, $p(x)$ is characterized as the product of (i) a log-uniform prior distribution based on ecologically plausible minimum and maximum values, and (ii) ecological and dynamical constraints (EDCs), where $p(x)$ is equal to 0 if DALEC parameter combinations or simulation outputs meet ecological conditions; these are described in (Bloom and Williams, 2015) and (Bloom et al., 2016).

For a given DALEC run, the likelihood, $L(x|O)$, is defined as:

$$L(x|O) = L_{LAI} \cdot L_{NEE} \cdot L_{biomass} \tag{13}$$

Where L_{LAI}, L_{NEE} and $L_{biomass}$ are the model-observation mismatches. Each likelihood term is derived as:

$$L* = exp(-\frac{1}{2}\sum((M_i - O_i)^2/U_i^2)) \tag{14}$$

Where M_i, O_i and U_i represent the model output, corresponding observation, and uncertainty, which represents the combined effects of model and observation error on model-data mismatch.

To sample $p(x|O)$, we used a Differential Evolution metropolis hastings Markov Chain Monte Carlo (DE-MCMC) algorithm (Ter Braak, 2006) with 200 walkers (Levine et al., 2023); in previous efforts we used a adaptive metropolis hastings MCMC (Haario et al., 2001). We found that the two algorithms overall give statistically similar results with comparable run times, however the DE-MCMC algorithm was found to be more stable and less likely to generate chains trapped in local minima.

2.5 Model Analysis and Diagnostics

2.5.1 Parameter Uncertainty Reduction

Following model calibration using CARDAMOM, it is useful to evaluate the constraint that the observations provide upon the model parameters. The prior probability density function (PDF) for the parameters is log uniform between the assumed minimum and maximum prior limits. The posterior PDFs for each parameter are represented by the sub-sampled solutions from the CARDAMOM optimization, hence the posterior PDF can take any form. Uncertainty reduction of the parameters was quantified using the relative change in interquartile range (IQR) from the prior to posterior, given by $100 \times (1 - \frac{IQR_{post}}{IQR_{prior}})$.

Note that this is calculated after transforming posterior parameter sub-samples into log space, to ensure consistency with the prior PDFs. The relative change metric is analogous to the relative uncertainty reduction calculated by the change in one sigma uncertainty used in other MDF studies (e.g. Knorr et al., 2010; Norton et al., 2019), but here the PDFs can be non-normal, therefore the IQR provides a simple and more representative metric of the uncertainty without the assumption of normality.

2.5.2 Model Performance

The model-data fit and predictive skill were evaluated using multiple statistical metrics. For the model output we used the ensemble median of CARDAMOM sub-samples at each time step. We used Pearson's correlation coefficient (r) to evaluate to the model skill at capturing the variability, the root mean squared error (RMSE) to evaluate the magnitude of the model-observed residuals, and mean bias (bias) to evaluate the model prediction bias. The best benchmark for model performance is whether it can predict data outside of the calibration period, therefore all model skill metrics presented are computed over the validation period.

The year-to-year variation, or interannual variability (IAV), in carbon cycle processes is better related to climate-carbon cycle relationships (Piao et al., 2020). Therefore, on top of the monthly variability, we report model skill at capturing IAV in LAI, NEE, GPP, and R_e over the validation period. We computed the IAV of the annual means, as well as on a seasonal basis. For the tropical savanna site, AU-How, we define the annual mean by the sites hydrological year which goes from September to the following August (Hutley and Beringer, 2010), and compute seasonal IAV based on Austral seasons.

Trend analyses were performed using linear regression over the entire simulation period (calibration and validation) to increase temporal coverage. Only months where observations are available were included to ensure a direct comparison between the modeled and observed trend.

2.5.3 Climate Sensitivity

A key aim of this study is to evaluate the climate sensitivity of the carbon cycle following MDF with the two model formulations. We focused on two climatic drivers, temperature and precipitation. Temperature can impact a number of carbon cycle processes including turnover rates of carbon pools and the physiological response of GPP and LAI. Precipitation impacts plant available soil water, W, and evapotranspiration, E. Hence, precipitation can impact GPP in both model formulations via a soil moisture factor, and Knorr LAI directly via the balance between W and E. We also computed the climate sensitivity to vapor pressure deficit, but this sensitivity was found to be multiple orders of magnitude smaller than temperature and precipitation so we did not include it in the analysis.

We used the finite difference method to compute the intrinsic climate sensitivity of LAI and NEE to precipitation and temperature. All simulations were performed using the forward model, M, and CARDAMOM posterior parameter set, p_{opt}. First, the model was run using the prescribed forcing data (F) and p_{opt} to generate the control simulation. Second, we perturbed the precipitation and temperature forcing data (denoted F'), independently, over the entire simulation period and ran the model forward to generate the perturbed simulations. The size of the forcing perturbation needs to be sufficiently small to avoid a non-linear response in the model. For the precipitation (P) perturbation we used $\delta F = \delta P = 1 \times 10^{-8} \ mm \ d^{-1}$ and for the

temperature (T) perturbation we use $\delta F = \delta T = 1 \times 10^{-5}\ ^\circ C$ (applied to both the minimum and maximum air temperature). With the control and perturbed simulations, we computed the derivative of the model output, LAI and NEE, with respect to the climate forcing variables, precipitation and temperature, by:

$$\frac{\delta X}{\delta F} = \frac{M(F', p_{opt}) - M(F, p_{opt})}{\delta F} \tag{15}$$

Where X represents the model output (LAI or NEE) and F represents either precipitation or temperature. This gave a time-series of the intrinsic precipitation and temperature sensitivities of LAI and NEE. We then decomposed these intrinsic sensitivities into (i) seasonal sensitivity by computing the monthly climatology over all simulation years, and (ii) average annual sensitivity by computing the average over the last n years of the simulation period.

To compare and evaluate the relative strength of precipitation sensitivity and temperature sensitivity, it was necessary to normalize the intrinsic sensitivities to a common unit. To do this, we scaled the intrinsic sensitivities by the respective climate variability. For the seasonal sensitivity analysis we multiplied the monthly average intrinsic sensitivity by the monthly interannual variability, computed as the standard deviation of each month in the simulation period. For the annual sensitivity analysis we multiplied the annual average intrinsic sensitivity by the interannual variability, computed as the standard deviation of the annual mean temperature or annual total precipitation. This is calculated using:

$$S_X^F = \frac{\delta X}{\delta F} \cdot \sigma(F) \tag{16}$$

Where S_X^F provides a measure of the climate sensitivity, S, of quantity X to the variability in forcing F. This generated four sensitivity metrics, two for LAI (S_{LAI}^T, S_{LAI}^P) and two for NEE (S_{NEE}^T, S_{NEE}^P), which are evaluated on seasonal and annual time-scales.

3 Results and Discussion

3.1 Model-Data Fit

The time-series of LAI and NEE at each site is shown in Fig. 2 for the calibration and validation periods, including the two models and the observations. This shows the CARDAMOM posterior PDF for both models at each site. Predictive skill over the validation period (r, RMSE, bias) at each site and for all site data combined is summarized in Fig. 3 and scatter plots in Appendix Fig. A1. Pearson's r shows that NEE temporal variability is better captured by the DALEC$_{Knorr}$ model at four of the six sites (AU-How, FR-LBr, FR-Pue, GF-Guy), while both models show comparable performance at the remaining two sites (FI-Hyy, US-Ha1), with equal correlations between the two models with all site data combined (r=0.83). The RMSE in NEE for both models is comparable for each site, with smaller residuals from DALEC$_{Knorr}$ at three sites (AU-How, FR-LBr, FR-Pue), larger residuals at two sites (FI-Hyy and US-Ha1), and equivalent residuals at GF-Guy. For all site data combined, RMSE=0.96 g C m^{-2} d^{-1} for both models. There is a small high bias in model NEE at most sites (<0.5 g C m^{-2} d^{-1}), with

DALEC$_{CDEA}$ showing a slightly lower bias for all site data combined (0.16 g C m^{-2} d^{-1}) compared to DALEC$_{Knorr}$ (0.21 g C m^{-2} d^{-1}), indicating the both models slightly underestimate net carbon uptake across sites.

Predictive skill for LAI (Fig. 3) shows that DALEC$_{Knorr}$ captures a larger proportion of the variability at one site (FR-Pue) and less at three sites (FI-Hyy, FR-LBr, GF-Guy), while both models perform similarly well at the remaining two sites (AU-How, US-Ha1). With all site data combined there is equal correlation with r=0.93. The RMSE for LAI shows that the DALEC$_{CDEA}$ residuals are smaller than DALEC$_{Knorr}$ at three sites (AU-How, FI-Hyy, US-Ha1) indicating better performance, while DALEC$_{Knorr}$ shows better performance at FR-LBr and FR-Pue, and finally both models show similar performance at GF-Guy. Across sites, there is a marginally better performance by DALEC$_{CDEA}$ (RMSE=0.55 m^2 m^{-2}) relative to DALEC$_{Knorr}$ (RMSE=0.57 m^2 m^{-2}). Both models tend to systematically underestimate LAI, as both models are biased low by 0.23 m^2 m^{-2} for all site data.

Both models capture the across site variability in ABGB with r=0.99, as shown in the Appendix Fig. A1. However, DALEC$_{CDEA}$ has smaller residuals for all site data combined, with a RMSE=472 g C m^{-2} and bias=-217 g C m^{-2}, versus a RMSE=952 g C m^{-2} and bias=-609 g C m^{-2} for DALEC$_{Knorr}$.

Observed and modeled trends for LAI and NEE are shown in Appendix Fig. A3. Observed LAI shows a significant trend at two sites, FR-LBr (p <0.05) and FR-Pue (p <0.001). DALEC$_{CDEA}$ also shows a significant trend at these two sites but overestimates the magnitude of the trend at FR-LBr by a factor of two and misrepresents the sign of the trend at FR-Pue. DALEC$_{CDEA}$ also produces a significant positive trend at FI-Hyy (p <0.05) and negative trend at GF-Guy (p <0.001), neither of which are shown by the observations. DALEC$_{Knorr}$ does not predict a significant trend in LAI at any site. No site shows a significant trend in observed NEE, which the DALEC$_{Knorr}$ is consistent with. DALEC$_{CDEA}$, however, predicts a significant positive NEE trend (p <0.05) suggesting a weakening carbon sink, consistent with the strong negative trend in modeled LAI. This suggests that DALEC$_{CDEA}$ can produce unrealistic trends in both LAI and NEE, whereas DALEC$_{Knorr}$ tends to be more stable at these time-scales.

Evaluation of the model-data fit to interannual variability (IAV) of LAI and NEE at annual and seasonal time-scales reveal distinct patterns (Appendix Fig. A2). On a seasonal basis, the IAV in LAI for winter, spring, and summer is represented similarly well by both models. However, fall IAV in LAI is captured better by DALEC$_{CDEA}$. This leads to a slightly better performance by DALEC$_{CDEA}$ in capturing LAI IAV on an annual basis. For NEE, DALEC$_{Knorr}$ performs slightly better at capturing IAV across sites with r=0.45 and RMSE=0.33 g C m^{-2} d^{-1}, while DALEC$_{CDEA}$ is r=0.33 and RMSE=0.35 g C m^{-2} d^{-1}, with the largest differences in fit occurring during fall.

CARDAMOM is fitting a global cost function which considers all observations in the MDF simultaneously. Therefore, trade-offs can occur between the fit to different observations, both in time and across the observation types (Kato et al., 2013). The results demonstrate that changing the model structure modifies how CARDAMOM converges to an optimal fit for the global cost function. There can therefore be compensatory effects between the fit to LAI and NEE observations. This occurs for the fall IAV in LAI and NEE, where DALEC$_{Knorr}$ better captures observed fall IAV in NEE, yet it performs worse at capturing observed fall IAV in LAI (Fig. A2). In other cases, a different model structure can lead to improved fit in both LAI and NEE, such as at FR-Pue for DALEC$_{Knorr}$, suggesting the integrated model structure is overall improved. In any case, assimilating multiple
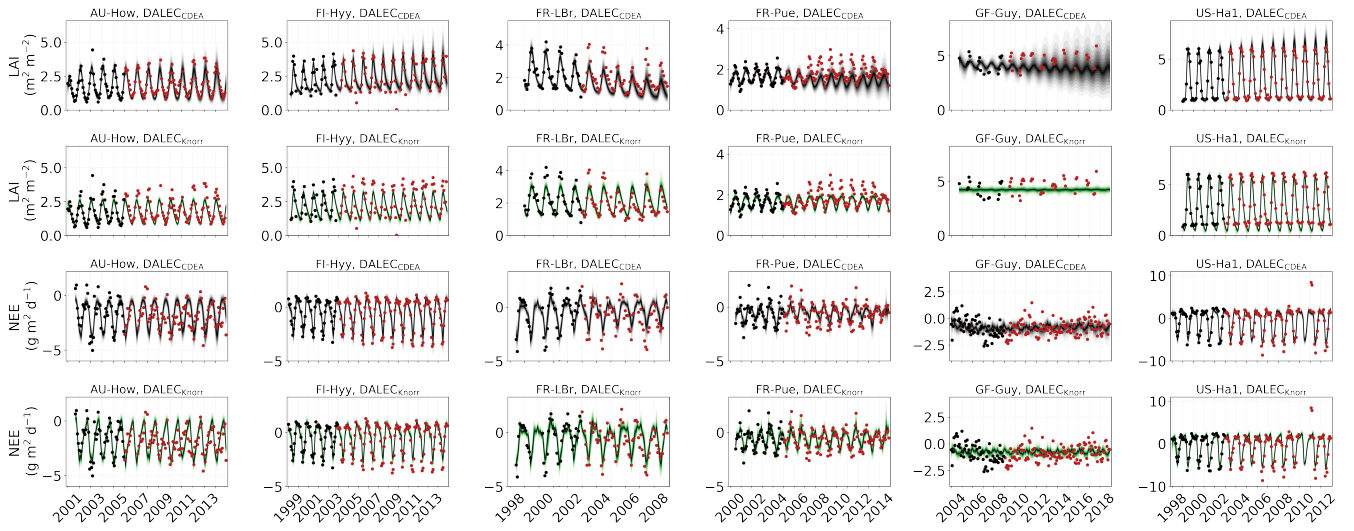
Figure 2. Model-data fit shown as time-series at each site and for each CARDAMOM LAI model formulation, for LAI (top panel) and NEE (bottom panel). Assimilated observations are shown as black markers (calibration) and withheld observations as red markers (validation). The gray shading shows the DALEC$_{CDEA}$ model and the green shading shows the DALEC$_{Knorr}$ model.

observational data streams simultaneously has benefits over assimilating data streams in separate, sequential steps (Kaminski et al., 2013). A sequential approach requires all uncertainties from each step to be propagated to the next, and this can be challenging when dealing with non-linear models. By assimilating multiple data streams simultaneously, the complementarity of the observations can be exploited and a more consistent view of the terrestrial carbon cycle can be achieved.

3.2 Underlying Parameters and Process Constraints

Here, we describe the estimated model parameters and their uncertainty reduction. First, we focus on the Knorr LAI phenology parameters, which govern the link between local climate and phenology, then follow up with a comparison of the remaining shared parameters of the two models.

The parameter uncertainty reduction, calculated as the relative reduction in IQR from the prior to posterior (see methods), for the Knorr LAI phenology parameters differs across sites (Appendix Fig. A5). There are nine process parameters in the Knorr LAI model (excluding initial conditions) across six sites, giving a total of 54 estimated LAI phenology parameters. Of these, 14 parameters show an uncertainty reduction of more than 80%, 15 show an uncertainty reduction between 50-80%, and 8 show an uncertainty reduction between 20-50%. The temperate deciduous forest site, US-Ha1, sees the strongest uncertainty reductions in LAI phenology parameters, with seven of the nine parameters being constrained by more than 50%. The weakest uncertainty reductions in LAI phenology parameters occur at the tropical evergreen forest site, GF-Guy, with just two parameters showing uncertainty reductions greater than 50%.
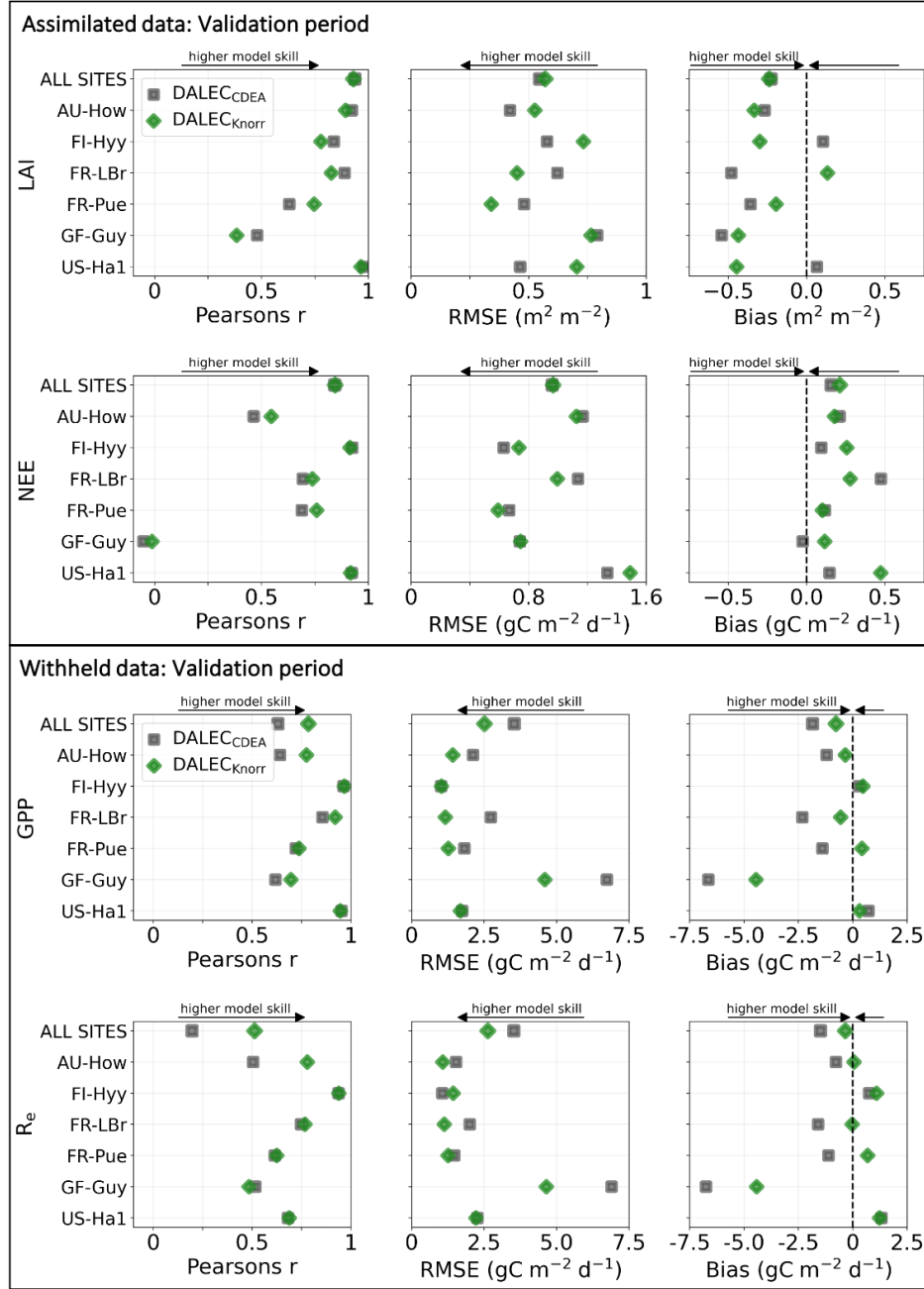
Figure 3. Top panel shows the model-data fit statistics to the assimilated observational data streams, LAI (top) and NEE (bottom), over the validation period. Bottom panel shows the model-data fit statistics to the withheld observational data streams, GPP (top) and R_e (bottom), over the validation period. Markers show the Pearson's correlation coefficient (r), RMSE, and bias, per study site and for all site data combined. The corresponding time-series showing the model-data fits are shown in Fig. 2 for LAI and NEE, and Fig. A4 for GPP and R_e. All site data combined scatter plots are shown in the Appendix Fig. A1.

Across all sites, the parameter representing the structural maximum LAI, $\hat{\Lambda}$, is well-constrained, showing an approximate normal posterior PDF and uncertainty reduction of more than 80%, indicating that $\hat{\Lambda}$ is well characterized by the MDF regardless of the site. The parameter for the mean temperature threshold at leaf onset, T_ϕ, shows strong uncertainty reductions of
385   66-80% at the four cooler sites (FI-Hyy, FR-LBr, FR-Pue, US-Ha1) and weaker uncertainty reductions of 35-39% at the two warmer tropical sites (AU-How, GF-Guy). This highlights the stronger role of temperature on LAI phenology at the cooler sites. The parameter for the mean photoperiod at leaf senescence, t_c, shows strong uncertainty reductions at all sites except the tropical evergreen forest (GF-Guy). There are small reductions in uncertainty of 5-18% for the parameter governing water limitation on LAI (i.e. drought deciduousness), τ_W, with the strongest reductions (15-18%) at AU-How, FI-Hyy, and US-Ha1. The
390   parameter τ_W is a proxy for the drought-deciduous behavior of phenology, with larger values indicating a stronger drought-deciduous strategy. The FR-LBr, FR-Pue, and GF-Guy sites show the strongest drought-deciduous behavior, while FI-Hyy and US-Ha1 sites show the weakest. The leaf growth rate parameter, ξ, the leaf longevity parameter, $\mathrm{k_{leaf}}$, and the background foliar turnover rate, θ_{foliar}, tend to show moderate uncertainty reductions across sites, although GF-Guy shows lower constraint on ξ and $\mathrm{k_{leaf}}$.

395   These results are distinguished from previous MDF studies that used the same phenology model, where phenological types and strategies were differentiated *a priori* (Kaminski et al., 2012; Kato et al., 2013; Knorr et al., 2010). In these studies, each plant functional type used a different prior PDF, while certain parameters were set as constants, effectively switching off some environmental factors that govern the growth triggers and senescence of LAI. As outlined in Knorr et al. (2010), this has advantages if there is sufficient prior knowledge about the species at each study site, however in most cases these
400   details are based on limited evidence or entirely unknown. Extending these priors beyond well-characterized sites or for global scale analyses with satellite data, requires further assumptions, often by classifying plant functional types which has significant shortcomings (see Van Bodegom et al., 2012). Over confidence in prior knowledge can be problematic, considering the prior PDFs have a significant impact on the inferred parameters and, therefore, the climatic controls of both LAI and NEE. Here, we applied equivalent prior PDFs at all sites, thus allowing CARDAMOM to infer the controls based only on the observations and
405   local climate. The posterior LAI phenology parameters generally show moderate to strong uncertainty reductions, indicating the ability of CARDAMOM to improve knowledge on phenological controls. The advantage of this approach is that environmental controls on LAI phenology emerge from the MDF system, exemplified by the stronger temperature control of spring leaf onset in cooler climate forests compared to warmer tropical sites. Furthermore, Knorr et al. (2010) fixed *a priori* the LAI water limitation parameter, τ_W, to be zero for cooler forest plant functional types so that LAI is never impacted by water availability.
410   We find that even for the cooler forest sites, the posterior τ_W is non-zero, indicating that water limitation plays a role on LAI dynamics, a finding that is also supported by empirical evidence (Buermann et al., 2018; Zhang et al., 2020). We note that switching off the photoperiod regulation process for canopy senescence, as was done in Knorr et al. (2010) for some PFTs, would likely lead to much larger τ_W values than those in our study, and this would have implications on the link between water availability and LAI dynamics.

415   There are a number of common parameters between DALEC$_{\mathrm{Knorr}}$ and DALEC$_{\mathrm{CDEA}}$ (Table A1). The posterior PDFs for these shared parameters are shown in Appendix Fig. A6. The canopy efficiency, a key parameter for GPP, has a higher median in

DALEC$_{Knorr}$ at four of the six sites, while it is the same at FR-Pue and lower at GF-Guy. The inferred carbon-use efficiency for DALEC$_{Knorr}$, equal to the ratio of NPP to GPP (defined by 1-f_{auto}), is approximately the same at three sites (FR-LBr, GF-Guy, US-Ha1), higher at two sites (AU-How, FR-Pue) and lower at one site (FI-Hyy). The inferred LMA at each site is systematically lower in DALEC$_{Knorr}$. Parameters that govern plant carbon allocation and turnover show distinct differences between DALEC$_{CDEA}$ and DALEC$_{Knorr}$. The DALEC$_{Knorr}$ model shows lower fractional allocation of NPP to the foliar pool across sites, consistent with the lower LMA as less carbon is required to maintain the same LAI. DALEC$_{Knorr}$ also shows systematically lower fractional allocation to roots yet a higher fractional allocation to wood. Despite this, the woody residence time is about 60-80% lower in DALEC$_{Knorr}$, as the turnover rate of woody biomass is significantly higher at all sites. Combined, these differences lead to a lower carbon residence time in vegetation in the DALEC$_{Knorr}$ model.

A key process linked with LAI phenology is litterfall production. We find that the litterfall rates predicted by DALEC$_{CDEA}$ and DALEC$_{Knorr}$ models are substantially different. In DALEC$_{Knorr}$, average annual litterfall rates range between 5.0-12.9 Mg C m^{-2} yr^{-1}. These fall well within the range of litterfall rates reported across global forest ecosystems, which range from 1-14 Mg C m^{-2} yr^{-1} (Zhang et al., 2014). The DALEC$_{CDEA}$ model predicts significantly higher litterfall rates at four sites (AU-How, FI-Hyy, FR-LBr, US-Ha1), ranging between 15.6-37.0 Mg C m^{-2} yr^{-1}, which greatly exceeds observed rates documented by Zhang et al. (2014). DALEC$_{Knorr}$ infers litterfall rates at GF-Guy, the tropical evergreen site, that are well within the range observed for that ecosystem type, while DALEC$_{CDEA}$ predicts almost a factor two lower litterfall, which is below the range observed for tropical forests. Overall, this implies that the carbon allocation, turnover, and, subsequently, litterfall is more realistic in DALEC$_{Knorr}$.

Parameters governing the water cycle are generally poorly constrained by the observations and both models tend to infer similar median parameter values. With no hydrology observations used here, there is no expectation of strong constraint on these processes. Despite this, it is notable that DALEC$_{Knorr}$ provides relatively strong constraint ($\sim$80%) on the initial water pool size (W^i) across sites, perhaps a consequence of the process link between W and the water-limited LAI (Eqs. 4, 5).

3.3 Validation of Inferred GPP and R$_e$ Fluxes

The two model structures implemented in CARDAMOM lead to differences in simulated GPP and R$_e$, the component fluxes of NEE. Comparison of the model inferred GPP and R$_e$ to the FLUXNET partitioned GPP and R$_e$ over the validation period is shown in Fig. A4 and summary statistics in Fig. 3. We reiterate that no GPP or R$_e$ data was used during calibration. Almost invariably, DALEC$_{Knorr}$ has a higher correlation, lower RMSE, and lower bias in GPP and R$_e$ relative to DALEC$_{CDEA}$. For GPP, with all site data combined, the model-data fit of DALEC$_{Knorr}$ gives r=0.78, RMSE=2.5 g C m^{-2} d^{-1}, bias=-0.8 g C m^{-2} d^{-1}, while DALEC$_{CDEA}$ gives r=0.63, RMSE=3.5 g C m^{-2} d^{-1}, and bias=-1.8 g C m^{-2} d^{-1}. Qualitatively, the fit to GPP shows similar performance to another recent CARDAMOM study that assimilated GPP and ET data (Smallman and Williams, 2019). The inferred R$_e$ shows a similar RMSE and bias as the GPP model-data fit, although the correlations tend to be lower than for GPP in both models, with r=0.51 for DALEC$_{Knorr}$ and r=0.20 for DALEC$_{CDEA}$. These results indicate that use of DALEC$_{Knorr}$ in CARDAMOM leads to a better representation of the component fluxes underlying NEE, despite a similar fit to assimilated data streams NEE and LAI.

The improvement in predictive skill of GPP by DALEC$_{Knorr}$ occurs predominantly at four of the six sites: AU-How, FR-LBr, FR-Pue, and GF-Guy. Both models simulate GPP as a function of the local climate, LAI, the canopy efficiency parameter, and the wilting point parameter that scales with W to impose soil moisture limitation. The difference in inferred GPP between the two models can be traced to the latter three of these factors. We find that the improved prediction of GPP by DALEC$_{Knorr}$ occurs due to a higher canopy efficiency at AU-How and FR-LBr, and a weaker soil moisture limitation at FR-Pue and GF-Guy. For the FI-Hyy site, even though the GPP predictions are similar between Knorr and DALEC$_{CDEA}$, the underlying mechanisms are different as DALEC$_{Knorr}$ has a 46% higher canopy efficiency, which compensates for the lower LAI. At US-Ha1, the inferred mechanisms controlling GPP are very similar between the two models.

At the two cooler climate forest sites, FI-Hyy and US-Ha1, both DALEC$_{Knorr}$ and DALEC$_{CDEA}$ show similar performance against FLUXNET GPP and R$_e$ data. This suggests that DALEC$_{CDEA}$ is adequate at inferring NEE component fluxes at these sites. While difficult to trace precisely why this occurs, it may relate to the development of the ACM GPP model which was originally formulated to represent cool climate forests (Williams et al., 1997). It is notable that the model inferred GPP and R$_e$ performs worst at the tropical evergreen site, GF-Guy, with low correlation and the highest residuals (RMSE and bias) of any site. This follows from the relatively poor model-data fit to LAI and NEE at GF-Guy (Fig. 3). At this site the seasonal variability in LAI and NEE is small relative to the data uncertainty, so seasonal variability carries little weight in the MDF system, potentially making it difficult for CARDAMOM to resolve the seasonal dynamics or its controls. Modeling GPP and LAI at tropical sites has been a longstanding challenge. Evidence has shown that the coupling of leaf age-dependent changes in photosynthetic capacity drives seasonal dynamics GPP, suggesting that models need to separate LAI into cohorts to better represent these leaf demography processes (Wu et al., 2016), a level of complexity that is not currently considered in any version of the DALEC model.

3.4 Climate Sensitivity of LAI and NEE

Here, we present and discuss the inferred climate sensitivities of LAI and NEE for both models. The focus is upon DALEC$_{Knorr}$ as it provides the relatively more process-based representation of LAI, however, we also explore the results from the DALEC$_{CDEA}$ model as it provides a useful test case for the effect of model structure on inferred climate sensitivities.

3.4.1 Temperature Sensitivities

The median S_{LAI}^T on an annual timescale ranges from zero to strongly positive depending upon the model and site. At all six sites, DALEC$_{CDEA}$ shows a larger S_{LAI}^T than DALEC$_{Knorr}$, perhaps indicative of the strong dependence of the CDEA formulation on temperature via carbon supply and a lack of direct climate controls. Both the DALEC$_{Knorr}$ and DALEC$_{CDEA}$ models infer the largest annual S_{LAI}^T at the two colder forest sites (FI-Hyy, US-Ha1), which is consistent with theoretical understanding of limiting factors on LAI (Caldararu et al., 2014; Richardson et al., 2013). DALEC$_{Knorr}$ infers a low annual S_{LAI}^T at the remaining four sites, with a median annual S_{LAI}^T of near zero for AU-How, FR-Pue, and GF-Guy, while there is a small positive median annual S_{LAI}^T at the FR-LBr site. DALEC$_{CDEA}$, however, infers moderately strong S_{LAI}^T at the remaining

four sites, and a median S^T_{LAI} for the two warm tropical sites (AU-How, GF-Guy) that exceed the S^T_{LAI} of two cooler temperate forest sites (FR-LBr, FR-Pue). In neither model does the LAI show a negative sensitivity to temperature.

485   The seasonality of the inferred S^T_{LAI} from DALEC$_{CDEA}$ and Knorr models show distinct differences (Figs. 4-5). DALEC$_{CDEA}$ tends to show a strong positive S^T_{LAI} year-round, often with a peak during the middle or late part of the growing season. DALEC$_{CDEA}$ also infers strong positive values during winter, even at the winter-dormant forest sites (FI-Hyy, US-Ha1) which goes against ecophysiological understanding at these sites (Richardson et al., 2013). For DALEC$_{Knorr}$, the seasonality of S^T_{LAI} is centered on spring onset, with very low S^T_{LAI} at other times of the year. These seasonal patterns of S^T_{LAI} suggest that

490   DALEC$_{Knorr}$ is more consistent with empirical understanding of temperature effects on LAI that show the temperature sensitivities is strongest during spring onset (Richardson et al., 2012; Piao et al., 2019).

Generally, at sites with a strong positive S^T_{LAI} at either annual or seasonal timescales, there is also a strong negative S^T_{NEE} (stronger carbon uptake with increased temperature), indicating the impact of LAI climate sensitivity on the sensitivity of NEE and the interrelated nature of these two processes. For DALEC$_{CDEA}$, the larger positive S^T_{LAI} generally leads to more negative

495   S^T_{NEE}, evident at seasonal timescales and by the negative annual S^T_{NEE} at across all sites. At four sites, the inferred median S^T_{NEE} by the two models differs in sign and magnitude, highlighting the impact of changes to the model formulation of LAI on the inferred temperature sensitivity of NEE. DALEC$_{Knorr}$ shows more variable annual S^T_{NEE} across sites, as the inferred median annual S^T_{NEE} can be either positive or negative depending upon the site. Seasonally, when there is a strong positive S^T_{LAI} (e.g. spring at FI-Hyy, FR-LBr, and US-Ha1) there is a concomitant negative S^T_{NEE}, demonstrating how a positive LAI temperature

500   sensitivity leads to stronger carbon uptake and that there is a strong seasonal dependence of this link in DALEC$_{Knorr}$. At these same three sites, other times of the year show a positive S^T_{NEE}, suggesting that LAI plays less of a role in in governing S^T_{NEE} outside of spring (Figs. A7-A8).

3.4.2   Precipitation Sensitivities

At all sites, S^P_{LAI} in DALEC$_{CDEA}$ is effectively zero (AU-How, FI-Hyy) or highly uncertain (FR-LBr, FR-Pue, GF-Guy, US-

505   Ha1). This may reflect the weak process link between water availability and LAI dynamics in DALEC$_{CDEA}$, considering the connection is mediated via the soil moisture limitation of GPP and subsequent changes in allocation of carbon to the foliar pool, which itself is buffered by the labile carbon pool, making LAI relatively insensitive to changes in water availability. In DALEC$_{Knorr}$, there is often a weak but consistently positive S^P_{LAI} with a more tightly constrained uncertainty, which implies that increases in precipitation lead to small increases in LAI. This is consistent with the Knorr model formulation for water-

510   limitation on LAI (Eq. 5), where changes in evapotranspiration (E) and/or plant available soil water (W) can mediate LAI directly via the τ_W parameter. More specifically, an increase in the ratio of W to E can increase LAI as there is less water limitation. The opposite is therefore also true, where a decline in the ratio of W to E will lead to a reduction in LAI, for example due to a decline in precipitation. The inferred S^P_{LAI} from DALEC$_{Knorr}$ is strongest at the AU-How and FR-LBr sites, with smaller S^P_{LAI} values at the three temperate and boreal forest sites, and effectively zero S^P_{LAI} at GF-Guy. Despite the lower

515   τ_W inferred at AU-How across sites, there is a relatively stronger S^P_{LAI} which may be due to higher evaporative demand at this tropical site causing larger imbalances between E and W.

Table 2. Annual temperature sensitivity per site, variable (LAI, NEE), and model in CARDAMOM. The values represent the median annual temperature sensitivity, with the 25th to 75th percentile range in brackets. * LAI sensitivities are scaled by 10^3.

		Annual Temperature Sensitivity per Site					
		AU-How	FI-Hyy	FR-LBr	FR-Pue	GF-Guy	US-Ha1
S_{LAI}^T *	DALEC$_{CDEA}$	17.7 (11.7 to 26.3)	25.2 (18.1 to 34.6)	2.9 (1.4 to 5.4)	5.3 (2.3 to 9.5)	7.2 (4.2 to 12.7)	24.1 (19.5 to 29.2)
	DALEC$_{Knorr}$	0 (0 to 0.3)	12.4 (1.4 to 25.7)	0.6 (0.0 to 7.8)	0 (0 to 1.9)	0 (0 to 2.6)	17.4 (1.7 to 49.5)
S_{NEE}^T	DALEC$_{CDEA}$	-8.5 (-11.2 to -6.4)	-3.1 (-4.4 to -1.8)	-0.8 (-1.6 to 0.1)	-0.6 (-1.3 to 0.0)	-0.8 (-1.5 to -0.4)	-0.8 (-2.7 to 2.2)
	DALEC$_{Knorr}$	0.5 (-1.4 to 3.0)	-1.8 (-5.5 to 0.9)	1.5 (0.2 to 3.3)	4.3 (1.7 to 7.3)	-0.1 (-0.6 to 0.7)	0.8 (-9.4 to 7.3)

Table 3. Annual precipitation sensitivity per site, variable (LAI, NEE), and model in CARDAMOM. The values represent the median annual precipitation sensitivity, with the 25th to 75th percentile range in brackets. * LAI sensitivities are scaled by 10^3.

		Annual Precipitation Sensitivity per Site					
		AU-How	FI-Hyy	FR-LBr	FR-Pue	GF-Guy	US-Ha1
S_{LAI}^P *	DALEC$_{CDEA}$	0.0 (0.0 to 0.0)	0.0 (0.0 to 50.3)	58.5 (0.0 to 139.1)	471.9 (309.5 to 656.5)	164.7 (0.0 to 373.3)	0.0 (0.0 to 28.9)
	DALEC$_{Knorr}$	0.2 (0.0 to 1.8)	0.1 (0.0 to 0.8)	0.4 (0.0 to 24.5)	0.1 (0.0 to 0.7)	0.0 (0.0 to 0.3)	0.1 (0.0 to 0.6)
S_{NEE}^P	DALEC$_{CDEA}$	-0.0 (-0.1 to 0)	-1.0 (-5.5 to -0.1)	-30.1 (-53.5 to 0.0)	-96.9 (-146.5 to -68.2)	-20.8 (-36.3 to 0.0)	2.1 (0.4 to 6.3)
	DALEC$_{Knorr}$	-0.5 (-1.8 to -0.2)	-1.5 (-3.6 to -0.6)	-50.5 (-73.5 to -0.6)	-1.1 (-73.1 to 0.5)	0.02 (-0.04 to 0.08)	4.4 (2.0 to 9.8)

The influence of water availability on Knorr LAI has a clear seasonality at some sites with S_{LAI}^P typically being largest during the mid to late growing season (Figs. 4-5). This seasonal dependence is consistent with recent evidence of late growing season water limitation on productivity at many forest sites (Buermann et al., 2018; Zhang et al., 2020). The evident seasonal dependence of water limitation on LAI in the Knorr model gives promise for an MDF approach to exploring the compensatory effects of temperature and water limitation on growing season phenology and productivity.

At most sites the inferred annual median S_{NEE}^P from both DALEC$_{CDEA}$ and Knorr models is negative, indicating an increase in net carbon uptake with increased precipitation. Only US-Ha1 shows a consistently positive annual S_{NEE}^P, across the uncertainty range and for both models. This is due to the strong positive sensitivity of R$_e$ to precipitation and very small positive sensitivity of GPP to precipitation inferred at the US-Ha1 site (Fig. A8). The seasonality in S_{NEE}^P shows considerable differences across sites, but it is evidently influenced by S_{LAI}^P. For example, DALEC$_{Knorr}$ shows a small positive S_{LAI}^P during the peak growing season at AU-How (Austral summer), and this leads to a stronger sensitivity of summer carbon uptake to precipitation. Similarly, the small positive S_{LAI}^P at FI-Hyy during the late growing season produces a slight seasonal shift in S_{NEE}^P to later in the year. In other sites the low S_{LAI}^P from DALEC$_{Knorr}$ is coupled with a very low S_{NEE}^P (FR-Pue, GF-Guy), whereas DALEC$_{CDEA}$ infers large and highly uncertain S_{NEE}^P values at these sites. The addition of the process-based Knorr model seems to help stabilize the sensitivity of NEE to precipitation at these two sites, as the exceptionally large and uncertain S_{LAI}^P from DALEC$_{CDEA}$ maps into S_{NEE}^P.
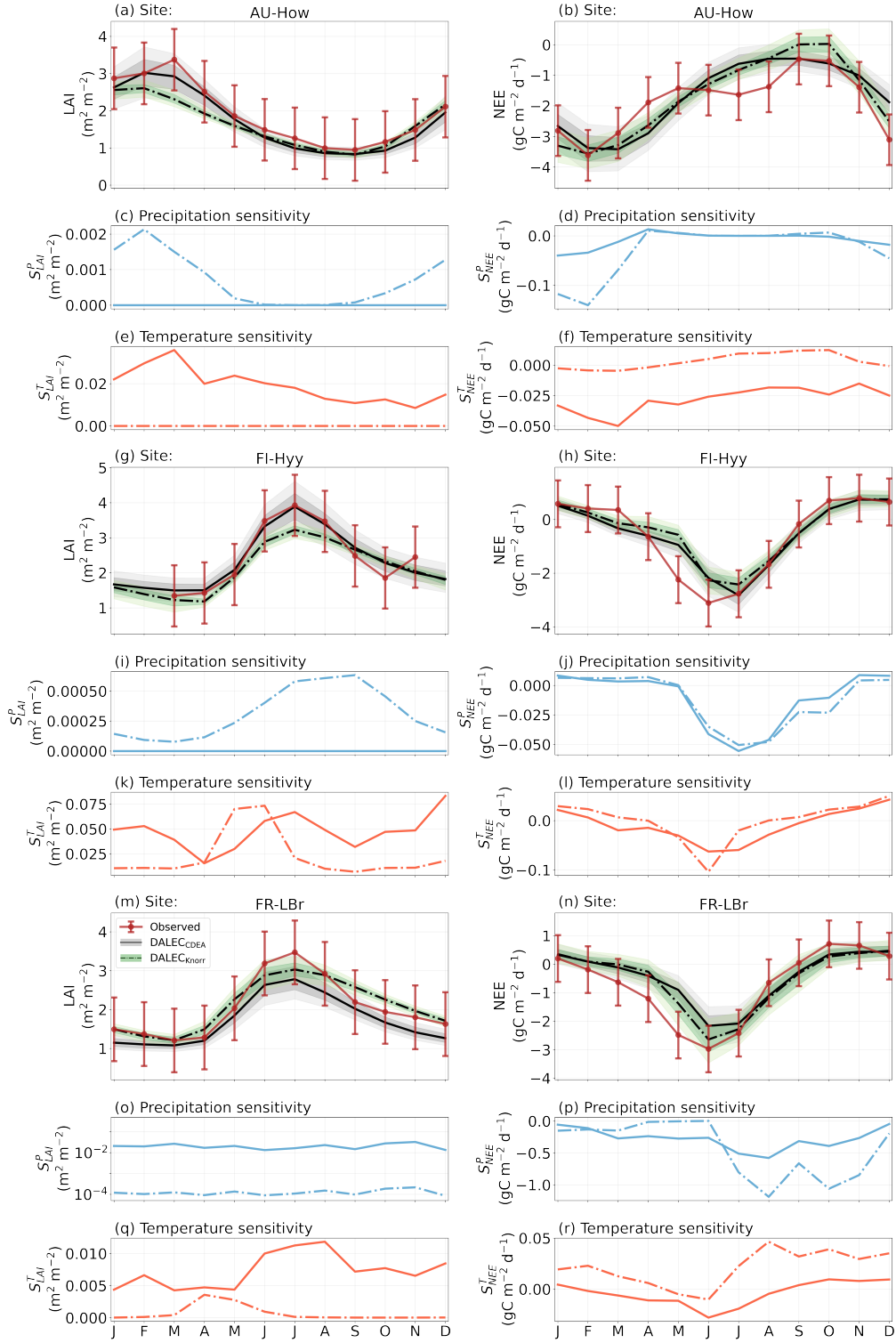
Figure 4. The seasonal pattern of model-data fit to LAI and NEE, and the inferred climate sensitivity of LAI and NEE to interannual variations in precipitation and air temperature.
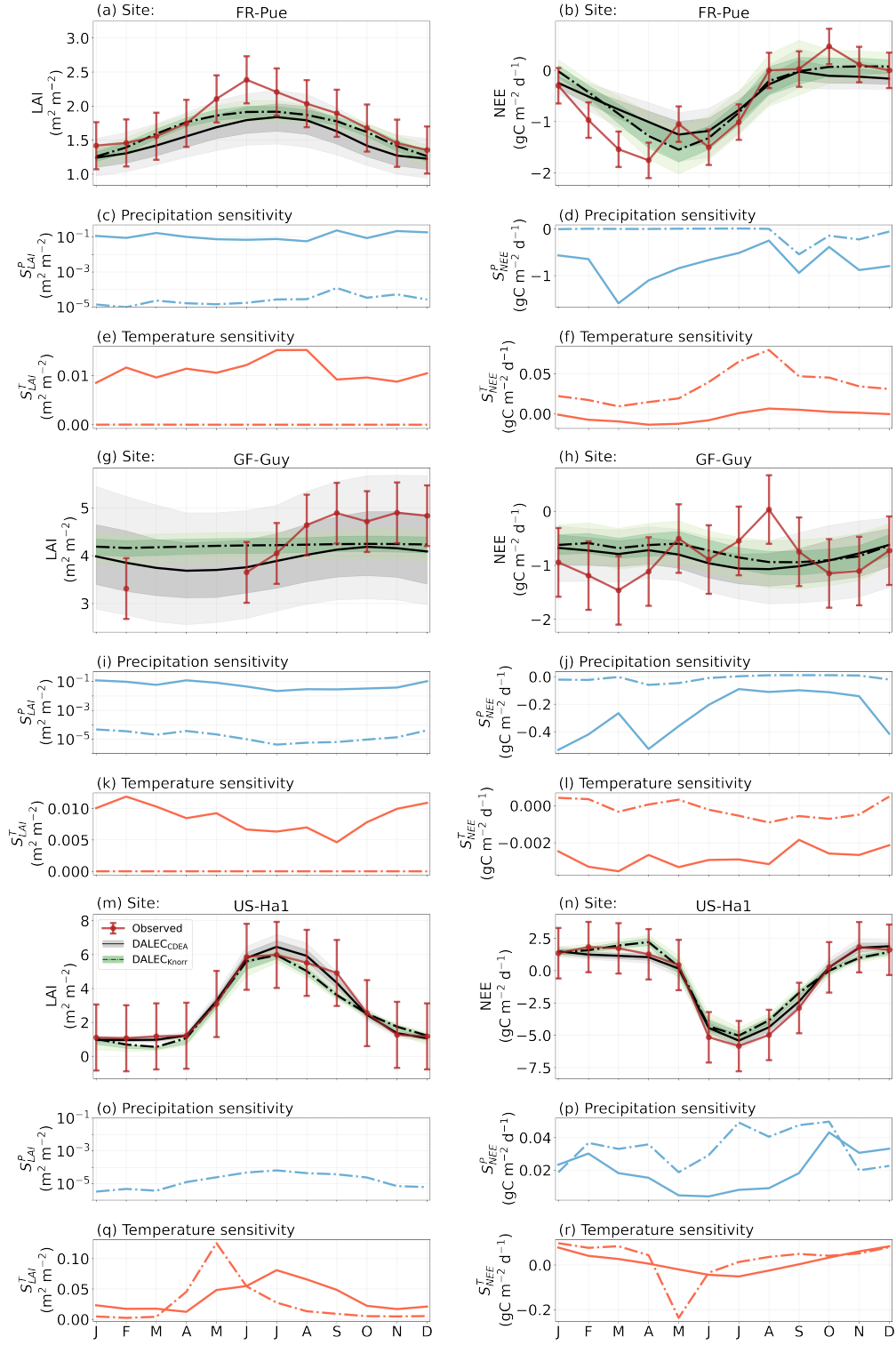
Figure 5. The seasonal pattern of model-data fit to LAI and NEE, and the inferred climate sensitivity of LAI and NEE to interannual variations in precipitation and air temperature.

3.5 Significance and Limitations

Developing models that balance process realism with reliability and robustness is key to better predictions of carbon-climate feedbacks (Prentice et al., 2015). The climate sensitivity of LAI is an important mechanism of terrestrial carbon-climate feedbacks, as it is closely coupled to both GPP and NEE (Richardson et al., 2013). Here, we implemented a new model for LAI phenology in CARDAMOM that includes processes such as temperature and photoperiod controls on growth and senescence triggers, water limitation on LAI growth rate, and couple LAI dynamics to plant carbon allocation and litterfall. Many previous studies have developed and tested climate-sensitive LAI phenology models against LAI data directly (e.g. Fox et al., 2022; Jolly and Running, 2004; Stöckli et al., 2008; Viskari et al., 2015). However, in the context of carbon-climate feedbacks it is critical that both model formulation and parameterization of LAI considers its close coupling to other processes in the carbon cycle. By confronting multiple processes with observations simultaneously in CARDAMOM we were able to generate a parameterization that provides a more integrated view of the carbon cycle (Kaminski et al., 2013). Furthermore, by extending the validation beyond the assimilated data streams, we were able to rigorously evaluate model skill. This is a key step toward a better model representation of the aggregate behavior of the terrestrial carbon cycle (Fisher and Koven, 2020).

Including the two model formulations allowed us to evaluate how the increase in complexity of LAI phenology influenced the fit to assimilated data streams (LAI, NEE, biomass) and wholly withheld data streams (GPP, R_e). Evidently, the DALEC$_{CDEA}$ model is a parsimonious approach for fitting data and appears to give reliable predictions of the assimilated data streams (Fig. 3) and even outperforms DALEC$_{Knorr}$ at capturing LAI IAV. The close coupling of LAI with GPP and plant allocation in DALEC$_{CDEA}$ may provide flexibility in fitting LAI data. However, at four sites DALEC$_{CDEA}$ predicts significant positive (FI-Hyy) or negative (FR-LBr, FR-Pue, GF-Guy) trends in LAI that are not present in the observations, nor are they predicted by DALEC$_{Knorr}$. At one site (FR-LBr), this leads to a significant trend in DALEC$_{CDEA}$ model NEE that is not supported by the observations, suggesting that DALEC$_{CDEA}$ may lead to large biases in the predicted net carbon balance over longer periods. The tighter coupling between climate and LAI phenology in DALEC$_{Knorr}$ may help moderate LAI and GPP dynamics during the prediction period, preventing unrealistic model trends. This difference in model formulation also leads to a significantly reduced biases in GPP by DALEC$_{Knorr}$, and may imply that the DALEC$_{CDEA}$ model is over-fitting to the LAI data. It is also important to consider that the satellite LAI observations can have systematic biases that we do not consider in the MDF system, as CARDAMOM only considers random errors. Satellite LAI retrievals can be particularly challenging over boreal (e.g. FI-Hyy), tropical (e.g. GF-Guy), and open woody savannas (e.g. AU-How) due to effects such as low solar zenith angle, snow and cloud contamination, visibility of the understory, and a lack of validation data (Fang et al., 2019). For example, the seasonal amplitude of satellite LAI is often overestimated in boreal evergreen systems (Heiskanen et al., 2012). Other potential error sources include spatial sampling differences between the eddy covariance tower footprint (NEE, GPP, R_e), biomass samples, and the footprint of Copernicus satellite LAI data. The limitations of the observations imply caution when over-fitting to any one data stream, and that correcting these sampling biases should allow us to better reconcile model and data. Further evaluation of the two models suggest that DALEC$_{Knorr}$ better captures temporal variability and magnitude of FLUXNET GPP and R_e data, and predicts more realistic annual litterfall rates compared to a global compilation of observations (Zhang et al.,

2014). Overall, despite the skill of DALEC$_{CDEA}$ at simulating LAI and NEE over the validation period, when considering the coupling with the carbon balance and the underlying processes, DALEC$_{Knorr}$ provides more reliable and robust performance.

Future work may explore how additional observations can constrain more uncertain regions of model parameter space. Doing so will help to further constrain ecosystem carbon cycle dynamics and the sensitivity to climate. In particular, constraining processes governing the water cycle and water availability for plant growth are key future directions of research. Observational constraints on the dynamics of W and E would be beneficial considering the large uncertainty of the parameters associated with these processes (Fig. A6), and could lead to differences in the estimates of S^P_{LAI} and S^P_{NEE}. At large scales, observations of terrestrial water storage by the NASA Gravity Recovery and Climate Experiment (GRACE) satellites (Tapley et al., 2004) can help inform hydrologic parameters and state variables in CARDAMOM (e.g. Massoud et al., 2022). At site scale, in-situ soil moisture measurements and eddy covariance measurements of E may be more useful (Smallman and Williams, 2019). Joint constraint of GPP and LAI may also help to constrain the two underlying controls of GPP, that is light interception which is mediated by LAI and light utilization for photosynthesis which is mediated by plant physiology, and may help further resolve the climate sensitivity of these processes. Satellite observations of solar-induced chlorophyll fluorescence show promise in this regard (Frankenberg et al., 2011), as they have been shown to inform on model parameters for both GPP and LAI (Norton et al., 2018) and improve spatiotemporal patterns of GPP (Parazoo et al., 2015; Norton et al., 2019). Even without direct observational constraints on GPP, the MDF setup gave reasonable estimates of the underlying fluxes, GPP and R$_e$, with best performance by DALEC$_{Knorr}$. Separating out NEE into these component fluxes has been a longstanding challenge in carbon cycle science (Schimel and Schneider, 2019). Extending this analysis to more diverse ecosystems is needed to evaluate the robustness of this result. There are opportunities to extend this to larger scales by using available observational data streams, such as atmospheric inversion estimates of NEE, estimates of biomass from passive microwave and radar sensors, and satellite LAI products (Schimel and Schneider, 2019). Overall, our results demonstrate that when assimilating these data streams into a model, the model process representation significantly affects how information from observations is propagated through to parameters and target processes, and that extending model evaluation to withheld data streams provides critical insight into model skill.

Joint constraint of LAI and NEE using observations led to important inferences about the climate sensitivity of the ecosystem carbon balance at the six study sites. From DALEC$_{Knorr}$, most temperate and boreal sites show that S^T_{LAI} occurs almost exclusively during spring onset which leads to a stronger seasonality in S^T_{NEE}. At some sites (AU-How, FR-LBr, FR-Pue) the two models infer S^T_{NEE} of opposite sign demonstrating how process representation of LAI leads to differences in optimized model behavior and response to climate. Furthermore, MDF with DALEC$_{Knorr}$ infers stronger S^T_{LAI} at colder climate forest sites, as expected from ecophysiological understanding and empirical observations (Richardson et al., 2013), highlighting the ability to infer temperature controls on phenology across biomes using this framework. The influence of water availability on LAI at temperate and boreal sites is strongest at the peak or late in the growing season, whereas at the tropical woodland savanna site (AU-How) it impacts LAI from the beginning to end of the growing season, suggesting biome-specific water controls on phenology. Further development of data-constrained S^T_{LAI} and S^P_{LAI} will help to reduce the large spread in earth system model predictions of LAI (Mahowald et al., 2016).

The temporal structure of the inferred climate sensitivities has implications for the response of ecosystem carbon cycling to a changing climate. Variability in climate forcing, due to both natural and anthropogenic factors, is not uniform in time or space, with seasonal dependencies on both the variability and trend (Franzke et al., 2020). The inferred sensitivities to temperature and precipitation (Figs. 4-5) show strong seasonality, so the seasonal structure of future climate change will have a strong impact on the response of the carbon cycle. For example, a change in spring temperature forcing will have markedly different impacts on LAI and NEE than an equivalent change in fall temperature forcing due to seasonal differences in the intrinsic sensitivity of the terrestrial carbon cycle. The intrinsic sensitivity to climate is the combination of a number of distinct yet interrelated processes, each of which can have different effects on the net carbon balance. Here, the inferred temperature sensitivities of LAI, S_{LAI}^T, are positive definite in all cases. However, the inferred temperature sensitivities of NEE, S_{NEE}^T, can be positive or negative depending upon the site and season. The climate sensitivities of NEE integrate over a number of underlying processes, each of which can have differences in sign and magnitude in their sensitivities (e.g. GPP and R_e can respond oppositely to precipitation, Appendix Fig. A7-A8). Constraining these sensitivities using diverse observations provides a robust way toward better representation of carbon-climate feedbacks.

Many phenological processes operate on time scales shorter than the monthly time step used in this analysis. Therefore, accurately resolving specific phenological events such as spring onset or fall senescence events, which can occur within days to weeks, is outside the scope of this study. However, the necessary mechanisms are included in the Knorr LAI model, providing a path toward finer time-scale analyses, which would help to characterize phenological responses to climate and the relationship with NEE. In any case, as outlined by Keenan et al. (2020) data-informed process modeling is key to resolving these processes, as implemented here, so that explicit consideration of processes and observational uncertainties can be mapped onto the inferred climate sensitivities.

4 Conclusions

This study demonstrates a holistic approach to model development for the purpose of improving model representation of the climate response of the terrestrial carbon cycle. We integrated a new formulation for LAI phenology into the DALEC terrestrial biosphere model, outlining the coupling to the carbon and water cycles, and performed MDF using CARDAMOM to calibrate the model against diverse Earth observations. Relative to the previous LAI phenology model in DALEC, the new DALEC model showed improved representation of the underlying processes governing the net ecosystem carbon balance, GPP and R_e, yet similar performance against the assimilated observations, LAI, NEE, and biomass. This analysis was carried forward to evaluate the data-informed climate sensitivity of the new model structure and parameterization, which showed large changes in the seasonality, sign, and magnitude of LAI and NEE sensitivities to temperature and precipitation. The added process realism of LAI phenology in DALEC/CARDAMOM provided more realistic and robust predictions of the terrestrial carbon cycle and its response to climate, highlighting the important role that LAI phenology plays in representing the terrestrial carbon cycle especially when considered in a MDF system.

Code and data availability. The model code, including DALEC and CARDAMOM, used in this paper is available at https://github.com/

635   CARDAMOM-framework/CARDAMOM_v2.3 (last access: May 2023). The input data (meteorological forcing, observations, calibration

configuration for CARDAMOM), output data (calibrated model states, fluxes and parameters) and post-processing code used for the statistical

analyses and to generate figures are available at Norton et al. (2023, DOI: 10.5281/zenodo.7793974).

Appendix A: DALEC Model Parameters

Table A1. DALEC model parameters that are shared, with the same physical meaning in both DALEC$_{CDEA}$ and DALEC$_{Knorr}$. This includes process parameters and initial conditions along with their prior range. Note: W = plant-available soil water pool, W$_u$ = plant-unavailable soil water pool. The allocation fraction to wood f_{wood} is computed as 1 - f_{auto} - f_{lab} f_{fol} - f_{root} in DALEC$_{CDEA}$, and 1 - f_{auto} - f_{lab} - f_{root} in DALEC$_{Knorr}$.

Class	#	Description	Symbol	Prior Range
Vegetation carbon	1	Canopy GPP efficiency	c_{eff}	5-50
	2	Wilting point	ω	1-10000
	3	NPP fraction to C_{lab} [a]	f_{lab}	0.01-0.5
	4	Leaf carbon mass per area	LMA	5-200
	5	Fraction of GPP respired	f_{auto}	0.2-0.8
	6	NPP fraction to C_{root} [a]	f_{root}	0.01-1
	7	Wood carbon turnover rate	θ_{wood}	0.000025-0.001
	8	Root carbon turnover rate	θ_{root}	0.0001-0.01
Litter & soil carbon	9	Litter carbon turnover rate	θ_{lit}	0.0001-0.01
	10	Soil carbon turnover rate	θ_{soil}	0.0000001-0.001
	11	Turnover rate for litter-soil transfer	$\theta_{lit2soil}$	0.0001-0.01
	12	Decomposition temp. rate	θ	0.018-0.08
	13	Moisture factor for decomposition	s_p	0.01-1
Water cycle	14	Underlying water-use efficiency	uWUE	0.5-30
	15	W runoff focal point	W^{Qmax}	1-100000
	16	W$_u$ runoff focal point	W_u^{Qmax}	1-100000
	17	W to W$_u$ runoff fraction	$h2o_{xfer}$	0.01-1
Initial conditions	18	Initial C_{lab}	C_{lab}^i	1-2000
	19	Initial C_{fol}	C_{fol}^i	1-2000
	20	Initial C_{root}	C_{root}^i	1-2000
	21	Initial C_{wood}	C_{wood}^i	1-100000
	22	Initial C_{lit}	C_{lit}^i	1-2000
	23	Initial C_{soil}	C_{soil}^i	1-200000
	24	Initial W pool	W^i	1-10000
	25	Initial W$_u$ pool	W_u^i	1-10000

Appendix B: CDEA Model Description and Parameters

The leaf onset factor (ϕ_{onset}), which is used to compute the carbon pool transfer from C_{lab} to C_{fol}, was originally defined in Bloom and Williams (Eq. A7, 2015) but was subsequently updated to include a variable residence time for C_{lab} (Bloom et al., 2020), with a formulation of

$$\phi_{onset}(t) = \frac{2}{\sqrt{\pi}} \cdot \left(\frac{ln(\theta_{lab}) - ln(\theta_{lab} - 1)}{c_{onset}\sqrt{2}} \right) \cdot e^{-(sin(\frac{t - d_{onset} + osl}{s}) \cdot \frac{2s}{c_{ronset}\sqrt{2}})^2} \tag{B1}$$

where θ_{lab}, c_{onset}, and d_{onset} are time-invariant parameters (Table B1), and osl is calculated by

$$osl = offset(\theta_{lab}, \frac{c_{ronset}\sqrt{2}}{2}) \tag{B2}$$

and $s = \frac{365.25}{\pi}$.

The leaf fall factor (ϕ_{fall}), which is used to compute the carbon pool transfer from C_{fol} to C_{lit} (i.e. litter production), was originally defined in Bloom and Williams (Eq. A8, 2015) but was subsequently updated to include a variable residence time for C_{fol} (Bloom et al., 2020), with a formulation of:

$$\phi_{fall}(t) = \frac{2}{\sqrt{\pi}} \cdot \left(\frac{ln(\theta_{fol}) - ln(\theta_{fol} - 1)}{c_{rfall}\sqrt{2}} \right) \cdot e^{-(sin(\frac{t - d_{rfall} + osf}{s}) \cdot \frac{2s}{c_{rfall}\sqrt{2}})^2} \tag{B3}$$

where θ_{fol}, c_{rfall}, and d_{rfall} are time-invariant parameters (Table B1), and where osl is calculated by

$$osf = offset(\theta_{fol}, \frac{c_{rfall}\sqrt{2}}{2}) \tag{B4}$$

The C_{fol} is updated at each time step by:

$$C_{fol}(t+1) = f_{fol}NPP(t) + \phi_{onset}(t)C_{lab}(t) + (1 - \phi_{fall}(t))C_{fol}(t) \tag{B5}$$

And C_{lab} is updated at each time step by:

$$C_{lab}(t+1) = f_{lab}(NPP(t) - f_{fol}NPP(t)) + (1 - \phi_{onset}(t))C_{lab}(t) \tag{B6}$$

Table B1. DALEC$_{CDEA}$ LAI phenology parameters including process parameters and initial conditions, along with their prior range. ^{a} This parameter is used in DALEC$_{CDEA}$ to allow for direct carbon allocation to C_{fol}, whereas DALEC$_{Knorr}$ does not, as C_{fol} is only supplied with carbon from C_{lab}. ^{b} Uses a circular prior range that extends beyonda 365.25 to prevent edge jumping during optimization (e.g. Dec 31 to Jan 1), as this parameter is used in a sine function with an annual period, so the actual day-of-year value can be computed as modulo 365.25.

Class	#	Description	Symbol	Prior Range	Units
LAI phenology	1	NPP fraction to C_{fol} ^{a}	f_{foliar}	0.01-0.5	-
	2	C_{lab} (leaf) lifespan	θ_{fol}	1.01-8	-
	3	C_{lab} lifespan	θ_{lab}	1.01-8	-
	4	Peak day of year for C_{lab} turnover	d_{onset}	365.25-1461^b	days
	5	Peak day of year for C_{fol} turnover	d_{rfall}	365.25-1461^b	days
	6	C_{lab} release period	c_{ronset}	30.4375-100	days
	7	Leaf fall period	c_{rfall}	30.4375-150	days
	8	Leaf carbon mass per area	LMA	5-200	g C m^{-2}
	9	Initial C_{lab}	C_{lab}^i	1-2000	g C m^{-2}
	10	Initial C_{fol}	C_{fol}^i	1-2000	g C m^{-2}

Appendix C: Knorr Model Description

The labile production flux is computed by:

$$F_{labprod}(t) = (GPP - R_a)f_{lab} = NPPf_{lab} \tag{C1}$$

where GPP is the gross primary productivity, R_a is the autotrophic respiration, NPP is the net primary productivity, and f_{lab} is a parameter representing the fraction NPP allocated to the labile pool.

The quadratic smoothing function is given in Knorr et al. (Eq. 25, 2010), we use η=0.99.

$$\upsilon(x,y) = \frac{x + y - \sqrt{(x+y)^2 - 4\eta xy}}{2\eta} \tag{C2}$$

Table C1. DALEC$_{Knorr}$ LAI phenology parameters including process parameters and initial conditions, along with their prior range. [a] Only during canopy senescence. [b] Defined as the fraction of LAI_{max}.

Class	#	Description	Symbol	Prior Range	Units
LAI phenology	1	Mean temperature at leaf onset	T_ϕ	268.15-323.15	°K
	2	Spatial range of T_ϕ	T_r	0.1-10	°K
	3	Linear growth constant	ξ	0.001-0.5	days^{-1}
	4	Inverse of leaf longevity[a]	k_{leaf}	0.001-0.5	days^{-1}
	5	Max intrinsic LAI	LAI_{max}	0.1-10	m^2 m^{-2}
	6	Length of dry spell before leaf shedding	τ_W	0.1-300	days
	7	Mean day length at leaf shedding	t_c	2-22	hrs
	8	Spatial range of t_c	t_r	0.1-6	hrs
	9	Background leaf turnover rate	θ_{foliar}	0.001-0.1	-
	10	Leaf carbon mass per area	LMA	5-200	g C m^{-2}
	11	Initial C_{lab}	C_{lab}^i	1-2000	g C m^{-2}
	12	Initial C_{fol}	C_{fol}^i	1-2000	g C m^{-2}
	13	Initial air temp. memory	T^i	268.15-323-15	°K
	14	Initial LAI_W [b]	LAI_W^i	0.01-1	-
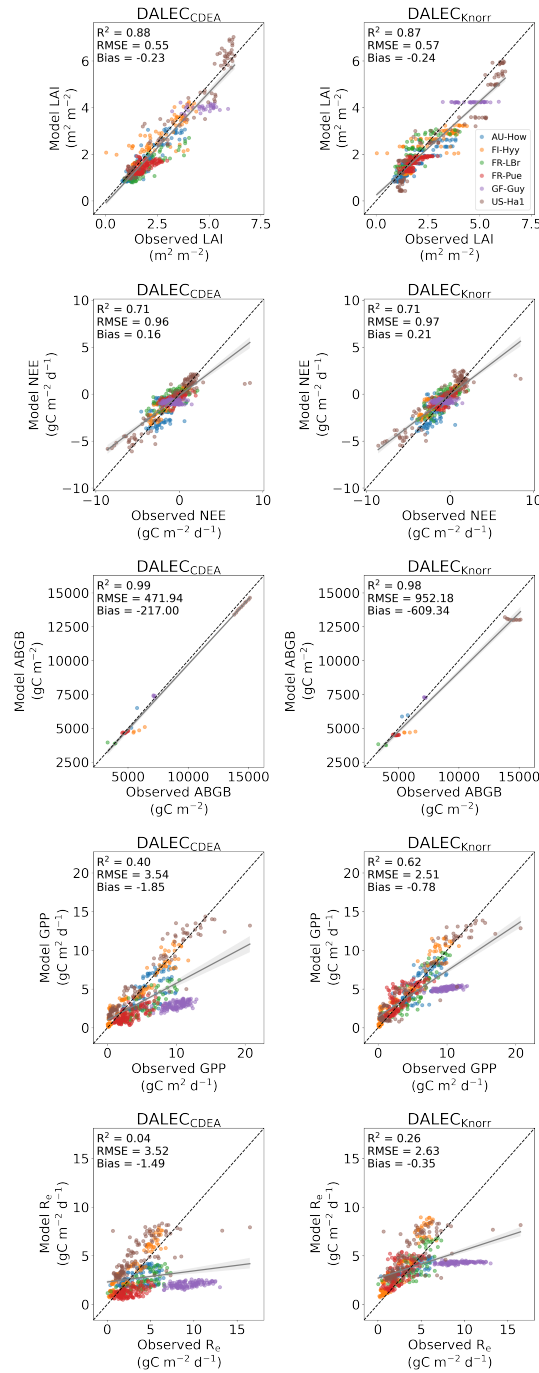
Figure A1. Model-data fit to assimilated observations (NEE, LAI, ABGB) and wholly withheld observations (GPP, R_e) over the validation period, for the DALEC$_{CDEA}$ model (left) and DALEC$_{Knorr}$ model (right). Each color represents a different site.
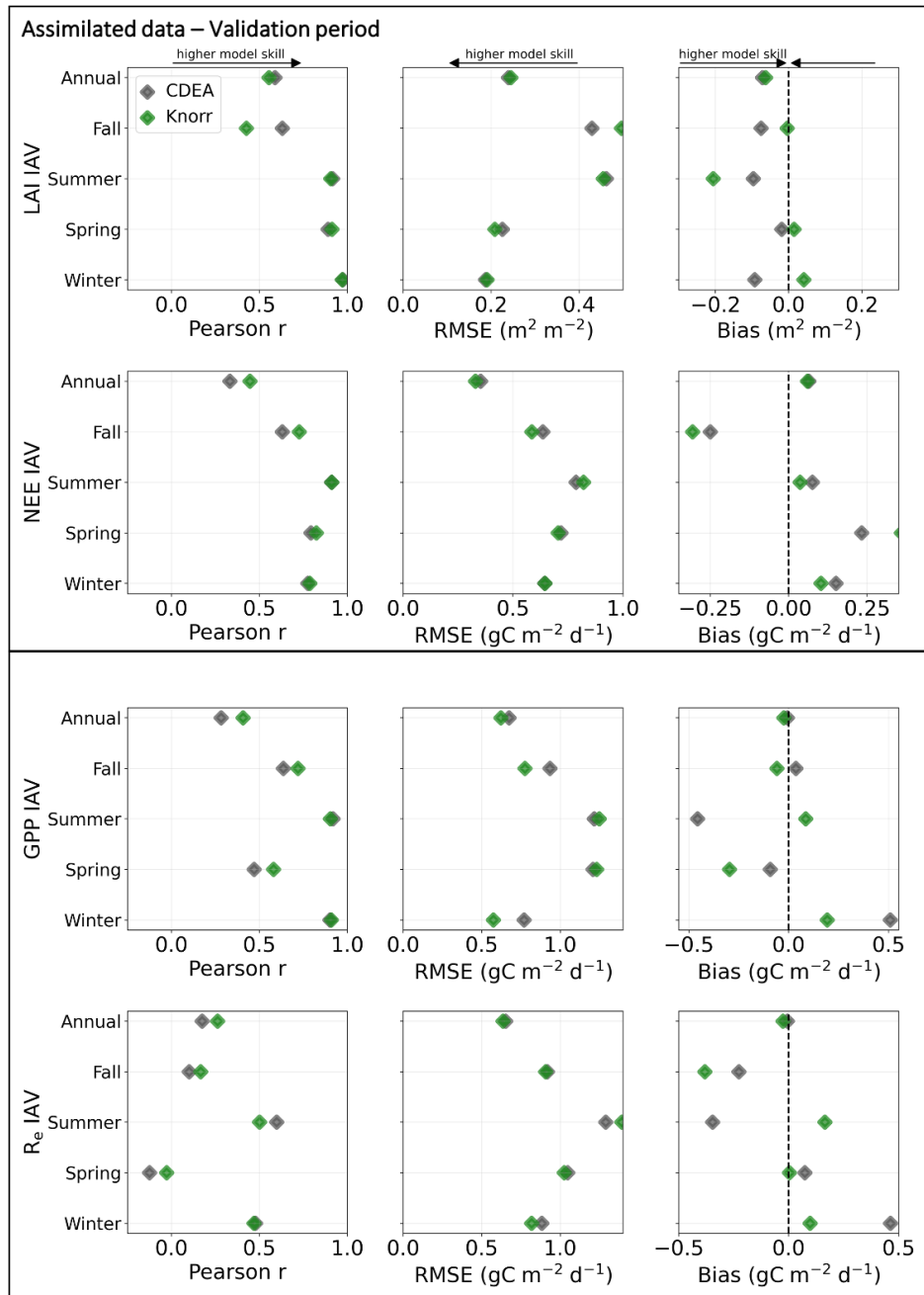
Figure A2. Model-data fit statistics for the interannual variability on an annual and seasonal basis against the assimilated observations (LAI, NEE) and wholly withheld data (GPP, R_e) over the validation period, with all sites combined. Markers show the Pearson's correlation coefficient (r), RMSE, and bias, per study site and for all site data combined.
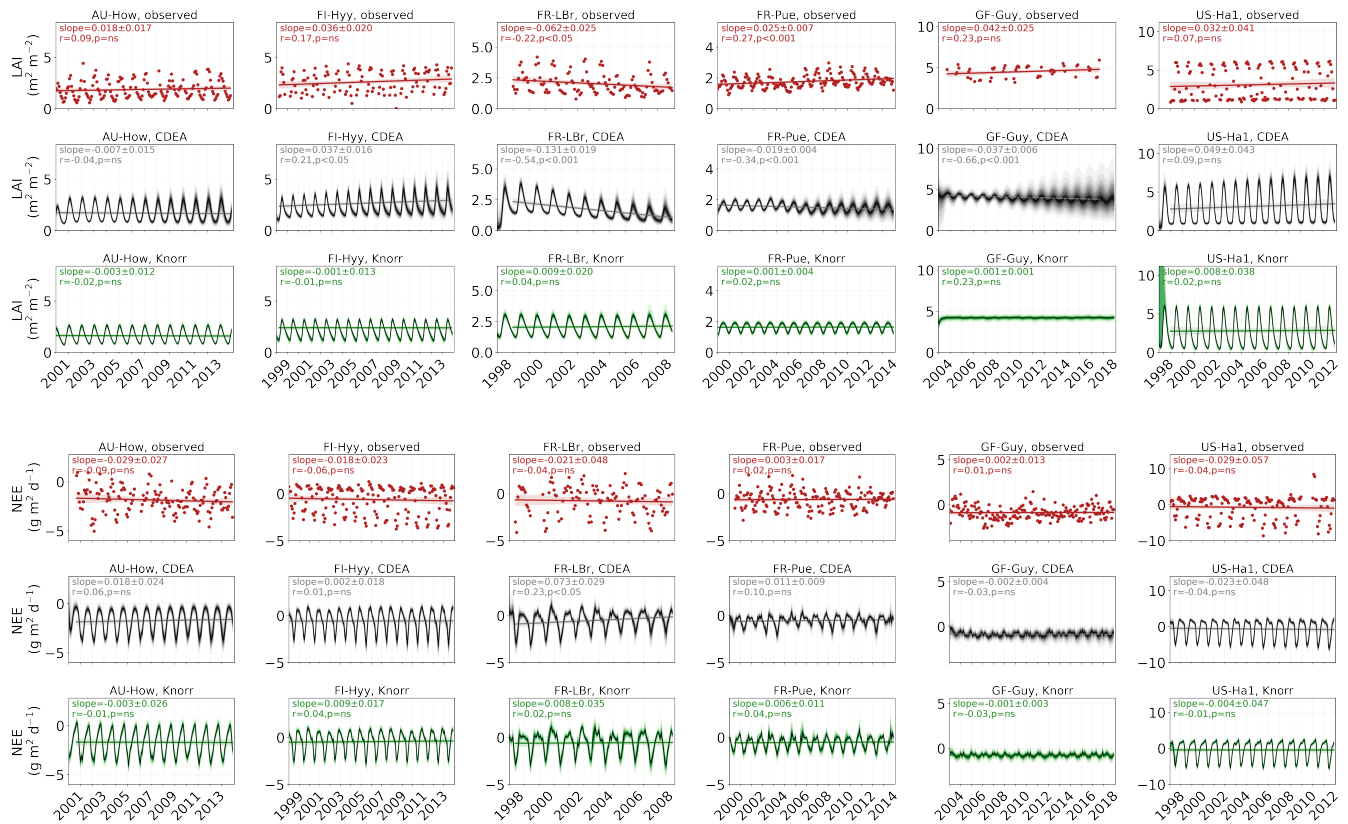
Figure A3. Observed and modeled trends over the calibration and validation period for LAI and NEE, including the observations (red dots), DALEC_CDEA model (grey line and shading), and DALEC_Knorr model (green line and shading). The slope, Pearson's r, and significance level for the linear regression is shown in each panel of the figure.

Author contributions. AJN and AAB designed the research. AJN conducted the modeling and analysis, with support from AAB, PAL, and SM. LTS processed the observational data. AJN prepared the manuscript and handled revisions. All authors contributed to interpretation and manuscript revisions.

Competing interests. TEXT

Disclaimer. TEXT
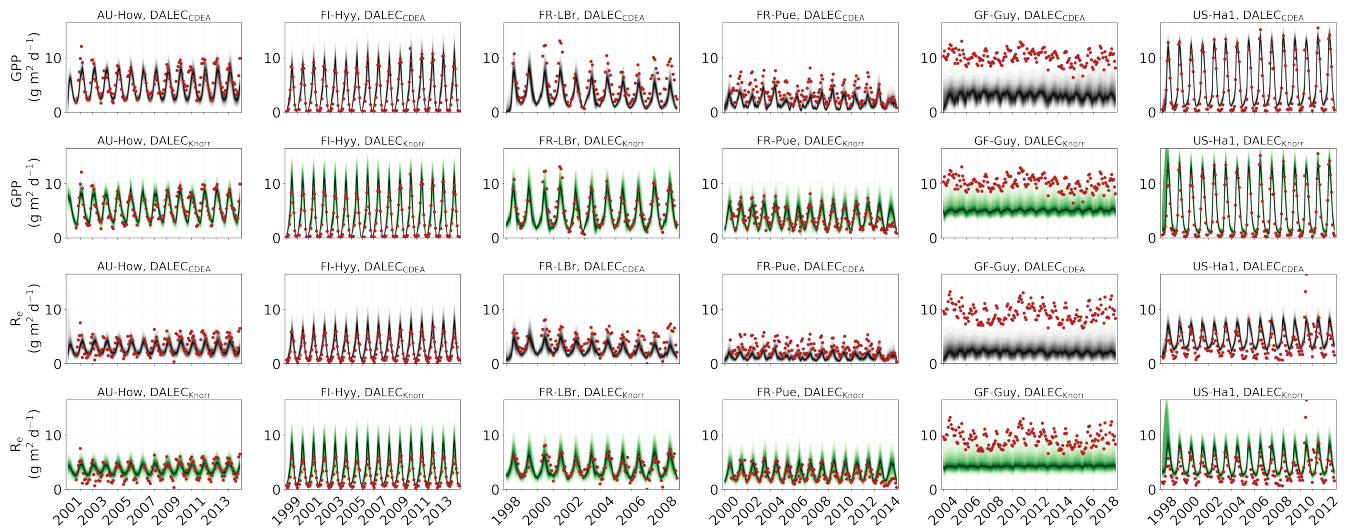
Figure A4. Model-data fit shown as time-series at each site and for each CARDAMOM LAI model formulation, for GPP (top panel) and R_e (bottom panel). Both GPP and R_e data were withheld from the MDF for validation. Observations are shown in red markers (validation data). The gray shading shows the DALEC$_{CDEA}$ model and the green shading shows the DALEC$_{Knorr}$ model.

Acknowledgements. A portion of this research was supported by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. Part of the funding for this study was provided through a NASA Carbon Cycle Science grant (#NNH20ZDA001N-CARBON). Alexander Norton, Anthony Bloom, Nicholas Parazoo, Paul Levine, Shuang Ma, and Renato Braghiere were all supported by the Jet Propulsion Laboratory, California Institute of Technology. NCP acknowledges funds and support provided under the NASA Earth Science Division (ESD) MEASURES and Arctic Boreal Vulnerability Experiment (ABoVE) programs, and the NSF Arctic Natural Sciences program. TLS was supported by the UK's National Centre for Earth Observation.

670
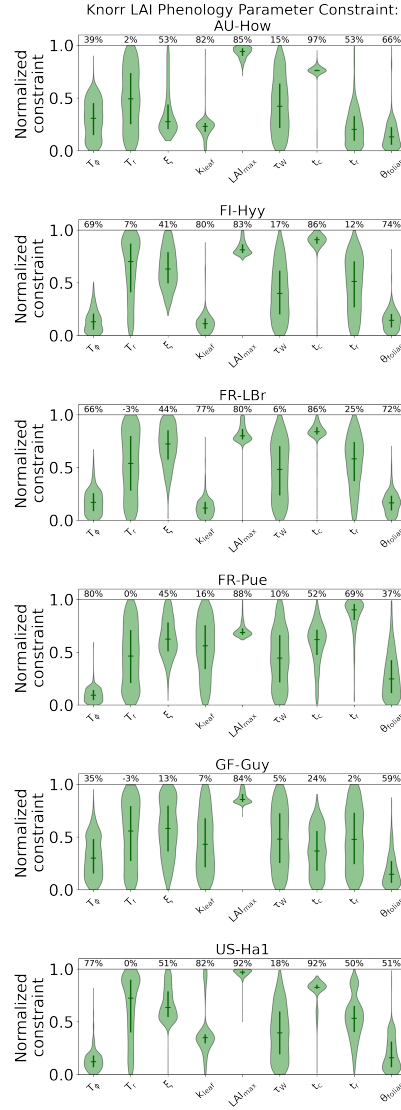
Figure A5. LAI phenology posterior parameter PDFs for the Knorr model, shown as the log-normalized PDFs between the minimum (normalized constraint=0) and maximum (normalized constraint=1) parameter bounds. The violin plot shows the full posterior PDF, the solid vertical bars indicate the 25th to 75th percentile range (IQR), and the horizontal solid lines indicate the median. The percentages at the top of the figure indicate the parameter uncertainty reduction from prior to posterior, reported as the reduction in log-normalized IQR (see methods).
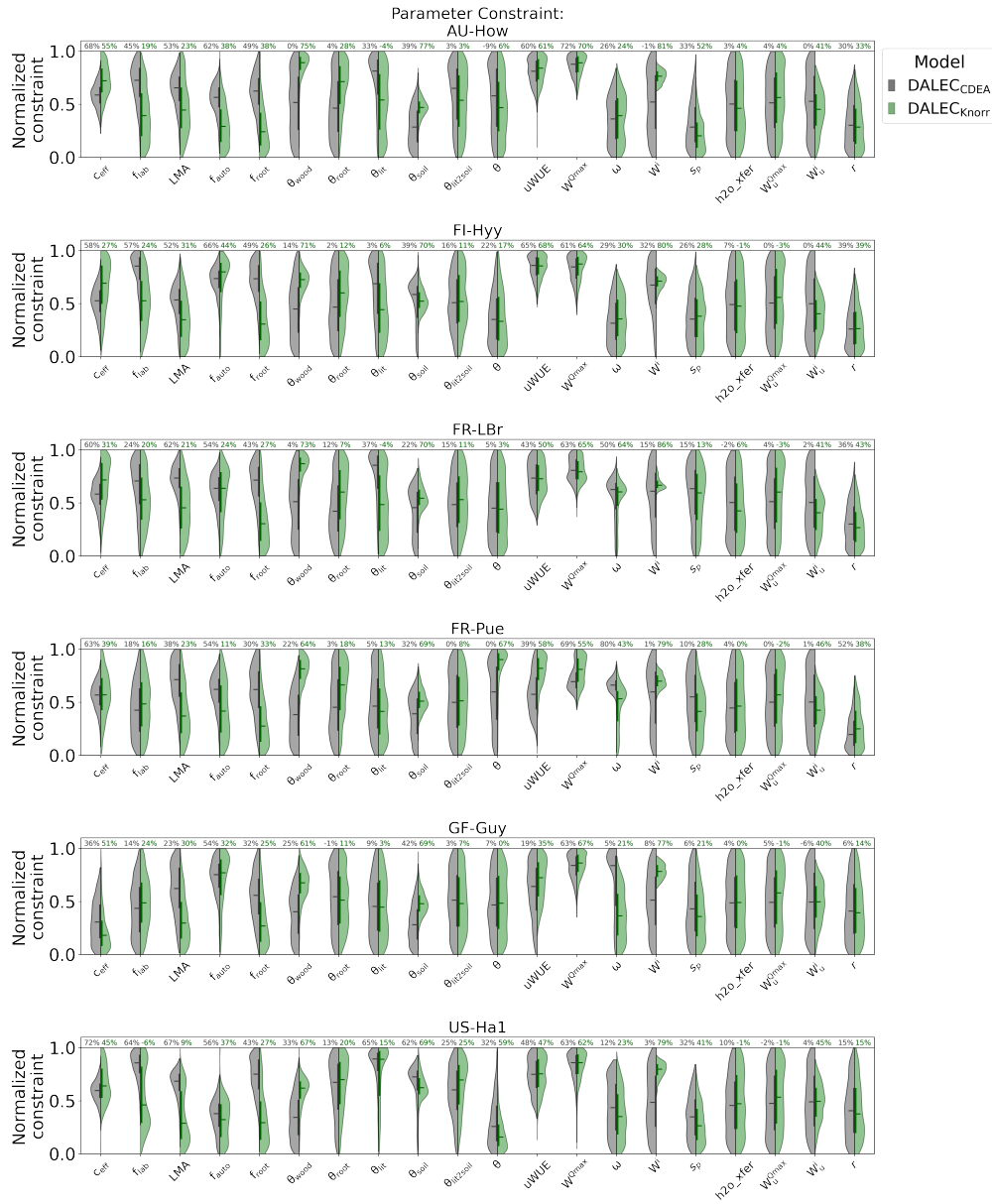
Figure A6. Posterior parameter PDFs for the common process parameters of the DALEC_CDEA model (gray) and DALEC_Knorr model (green), excluding initial conditions. Violin plots of 'normalized contraint' show the log-normalized PDFs between the minimum (normalized constraint=0) and maximum (normalized constraint=1) parameter bounds. The violin plot shows the full posterior PDF, the solid vertical bars indicate the 25th to 75th percentile range (IQR), and the horizontal solid lines indicate the median. The percentages at the top of the figure indicate the parameter uncertainty reduction from prior to posterior, reported as the reduction in log-normalized IQR (see methods).
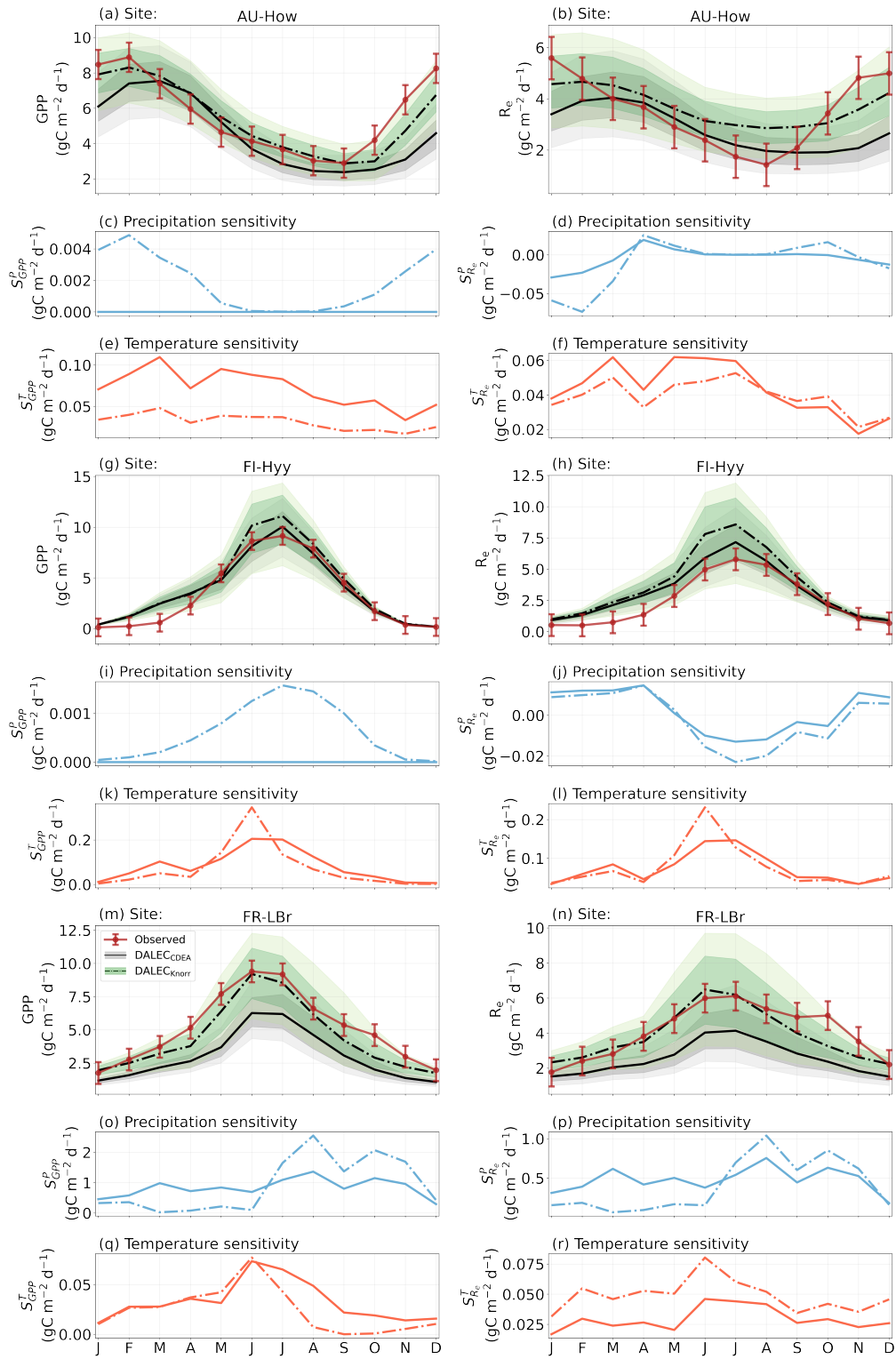
Figure A7. The seasonal pattern of model simulated GPP and R_e, and the inferred climate sensitivity of GPP and R_e to interannual variations in precipitation and air temperature.
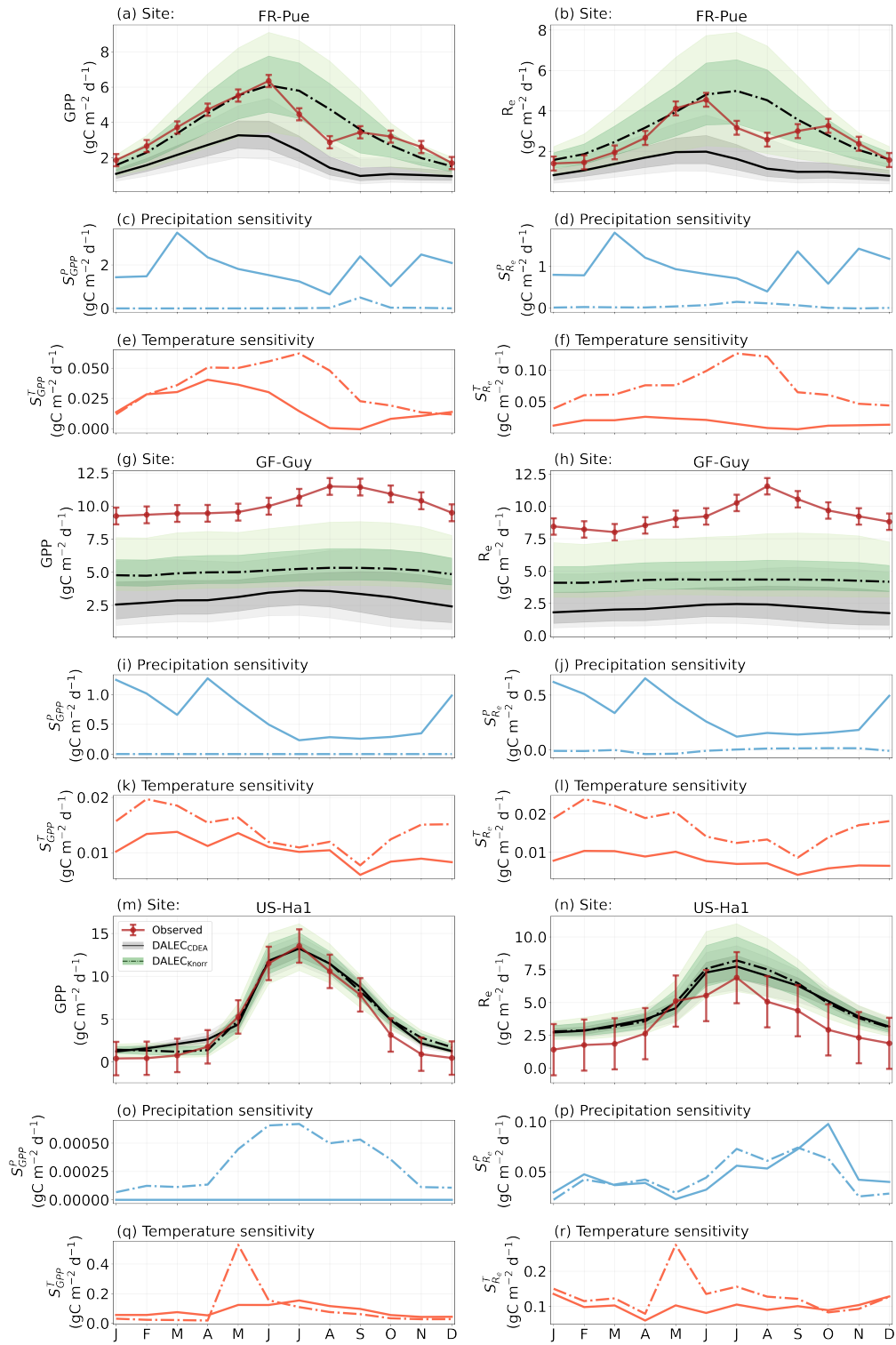
Figure A8. The seasonal pattern of model simulated GPP and R_e, and the inferred climate sensitivity of GPP and R_e to interannual variations in precipitation and air temperature.

References

Albert, L. P., Restrepo-Coupe, N., Smith, M. N., Wu, J., Chavana-Bryant, C., Prohaska, N., Taylor, T. C., Martins, G. A., Ciais, P., Mao, J., Arain, M. A., Li, W., Shi, X., Ricciuto, D. M., Huxman, T. E., McMahon, S. M., and Saleska, S. R.: Cryptic phenology in plants: Case studies, implications, and recommendations, Global Change Biology, 25, 3591–3608, https://doi.org/10.1111/gcb.14759, 2019.

Baldocchi, D.: TURNER REVIEW No. 15. 'Breathing' of the terrestrial biosphere: Lessons learned from a global network of carbon dioxide flux measurement systems, Australian Journal of Botany, 56, 1–26, https://doi.org/10.1071/BT07151, 2008.

Bloom, A. A. and Williams, M.: Constraining ecosystem carbon dynamics in a data-limited world: Integrating ecological "common sense" in a model-data fusion framework, Biogeosciences, 12, 1299–1315, https://doi.org/10.5194/bg-12-1299-2015, 2015.

Bloom, A. A., Exbrayat, J. F., Van Der Velde, I. R., Feng, L., and Williams, M.: 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 of the United States of America, 113, 1285–1290, https://doi.org/10.1073/pnas.1515160113, 2016.

Bloom, A. A., Bowman, K. W., Liu, J., Konings, A. G., Worden, J. R., Parazoo, N. C., Meyer, V., Reager, J. T., Worden, H. M., Jiang, Z., Quetin, G. R., Luke Smallman, T., Exbrayat, J. F., Yin, Y., Saatchi, S. S., Williams, M., and Schimel, D. S.: Lagged effects regulate the inter-annual variability of the tropical carbon balance, Biogeosciences, 17, 6393–6422, https://doi.org/10.5194/bg-17-6393-2020, 2020.

Booth, B. B. B., Jones, C. D., Collins, M., Totterdell, I. J., Cox, P. M., Sitch, S., Huntingford, C., Betts, R. A., Harris, G. R., and Lloyd, J.: High sensitivity of future global warming to land carbon cycle processes, Environmental Research Letters, 7, 024 002, https://doi.org/10.1088/1748-9326/7/2/024002, 2012.

Buermann, W., Forkel, M., O'Sullivan, M., Sitch, S., Friedlingstein, P., Haverd, V., Jain, A. K., Kato, E., Kautz, M., Lienert, S., Lombardozzi, D., Nabel, J. E. M. S., Tian, H., Wiltshire, A. J., Zhu, D., Smith, W. K., and Richardson, A. D.: Widespread seasonal compensation effects of spring warming on northern plant productivity, Nature, 562, 110–114, https://doi.org/10.1038/s41586-018-0555-7, 2018.

Caldararu, S., Purves, D. W., and Palmer, P. I.: Phenology as a strategy for carbon optimality: A global model, Biogeosciences, 11, 763–778, https://doi.org/10.5194/bg-11-763-2014, 2014.

Clelend, E., Chuine, I., Menzel, A., Mooney, H., and Schwartz, M.: Shifting plant phenology in response to global change, Trends in Ecology & Evolution, 22, 357–365, https://doi.org/10.1016/j.tree.2007.04.003, 2007.

Cole, E. F. and Sheldon, B. C.: The shifting phenological landscape: Within- and between-species variation in leaf emergence in a mixed-deciduous woodland, Ecology and Evolution, 7, 1135–1147, https://doi.org/10.1002/ece3.2718, 2017.

Cooke, J. E. K., Eriksson, M. E., and Junttila, O.: The dynamic nature of bud dormancy in trees: environmental control and molecular mechanisms, Plant, Cell & Environment, 35, 1707–1728, https://doi.org/10.1111/j.1365-3040.2012.02552.x, 2012.

Delpierre, N., Vitasse, Y., Chuine, I., Guillemot, J., Bazot, S., Rutishauser, T., and Rathgeber, C. B.: Temperate and boreal forest tree phenology: from organ-scale processes to terrestrial ecosystem models, Annals of Forest Science, 73, 5–25, https://doi.org/10.1007/s13595-015-0477-6, 2016.

Famiglietti, C. A., Smallman, T. L., Levine, P. A., Flack-Prain, S., Quetin, G. R., Meyer, V., Parazoo, N. C., Stettz, S. G., Yang, Y., Bonal, D., Bloom, A. A., Williams, M., and Konings, A. G.: Optimal model complexity for terrestrial carbon cycle prediction, Biogeosciences, 18, 2727–2754, https://doi.org/10.5194/bg-18-2727-2021, 2021.

Fang, H., Baret, F., Plummer, S., and Schaepman-Strub, G.: An Overview of Global Leaf Area Index (LAI): Methods, Products, Validation, and Applications, Reviews of Geophysics, 57, 739–799, https://doi.org/10.1029/2018RG000608, 2019.

Fisher, R. A. and Koven, C. D.: Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2018MS001453, 2020.

Fox, A. M., Huo, X., Hoar, T. J., Dashti, H., Smith, W. K., MacBean, N., Anderson, J. L., Roby, M., and Moore, D. J. P.: Assimilation of Global Satellite Leaf Area Estimates Reduces Modeled Global Carbon Uptake and Energy Loss by Terrestrial Ecosystems, Journal of Geophysical Research: Biogeosciences, 127, https://doi.org/10.1029/2022JG006830, 2022.

Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J.-E., Toon, G. C., Butz, A., Jung, M., Kuze, A., and Yokota, T.: New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity, Geophysical Research Letters, 38, n/a–n/a, https://doi.org/10.1029/2011GL048738, 2011.

Franzke, C. L. E., Barbosa, S., Blender, R., Fredriksen, H., Laepple, T., Lambert, F., Nilsen, T., Rypdal, K., Rypdal, M., Scotto, M. G., Vannitsem, S., Watkins, N. W., Yang, L., and Yuan, N.: The Structure of Climate Variability Across Scales, Reviews of Geophysics, 58, https://doi.org/10.1029/2019RG000657, 2020.

Friedl, M. and Sulla-Menashe, D.: MCD12C1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 0.05Deg CMG V006 [Data set], 2015.

Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., and Knutti, R.: Uncertainties in CMIP5 Climate Projections due to Carbon Cycle Feedbacks, Journal of Climate, 27, 511–526, https://doi.org/10.1175/JCLI-D-12-00579.1, 2014.

Fuster, B., Sánchez-Zapero, J., Camacho, F., García-Santos, V., Verger, A., Lacaze, R., Weiss, M., Baret, F., and Smets, B.: Quality Assessment of PROBA-V LAI, fAPAR and fCOVER Collection 300 m Products of Copernicus Global Land Service, Remote Sensing, 12, 1017, https://doi.org/10.3390/rs12061017, 2020.

Ge, R., He, H., Zhang, L., Ren, X., Williams, M., Yu, G., Luke Smallman, T., Zhou, T., Li, P., Xie, Z., Wang, S., Wang, H., Zhou, G., Zhang, Q., Wang, A., Fan, Z., Zhang, Y., Shen, W., Yin, H., and Lin, L.: Climate Sensitivies of Carbon Turnover Times in Soil and Vegetation: Understanding Their Effects on Forest Carbon Sequestration, Journal of Geophysical Research: Biogeosciences, 127, https://doi.org/10.1029/2020JG005880, 2022.

Haario, H., Saksman, E., and Tamminen, J.: An Adaptive Metropolis Algorithm, Bernoulli, 7, 223, https://doi.org/10.2307/3318737, 2001.

Heiskanen, J., Rautiainen, M., Stenberg, P., Mõttus, M., Vesanto, V.-H., Korhonen, L., and Majasalmi, T.: Seasonal variation in MODIS LAI for a boreal forest area in Finland, Remote Sensing of Environment, 126, 104–115, https://doi.org/10.1016/j.rse.2012.08.001, 2012.

Hill, T. C., Ryan, E., and Williams, M.: The use of CO2 flux time series for parameter and carbon stock estimation in carbon cycle research, Global Change Biology, 18, 179–193, https://doi.org/10.1111/j.1365-2486.2011.02511.x, 2012.

Huntzinger, D. N., Michalak, A. M., Schwalm, C., Ciais, P., King, A. W., Fang, Y., Schaefer, K., Wei, Y., Cook, R. B., Fisher, J. B., Hayes, D., Huang, M., Ito, A., Jain, A. K., Lei, H., Lu, C., Maignan, F., Mao, J., Parazoo, N., Peng, S., Poulter, B., Ricciuto, D., Shi, X., Tian, H., Wang, W., Zeng, N., and Zhao, F.: Uncertainty in the response of terrestrial carbon sink to environmental drivers undermines carbon-climate feedback predictions, Scientific Reports, 7, 4765, https://doi.org/10.1038/s41598-017-03818-2, 2017.

Hutley, L. B. and Beringer, J.: Disturbance and climatic drivers of carbon dynamics of a North Australian tropical Savanna, in: Ecosystem Function in Savannas: Measurement and Modeling at Landscape to Global Scales, edited by Hill, M. J. and Hanan, N. P., pp. 57–75, CRC Press, 1 edn., https://doi.org/10.1201/b10275, 2010.

Iio, A., Hikosaka, K., Anten, N. P., Nakagawa, Y., and Ito, A.: Global dependence of field-observed leaf area index in woody species on climate: A systematic review, Global Ecology and Biogeography, 23, 274–285, https://doi.org/10.1111/geb.12133, 2014.

Jolly, W. M. and Running, S. W.: Effects of precipitation and soil water potential on drought deciduous phenology in the Kalahari, Global Change Biology, 10, 303–308, https://doi.org/10.1046/j.1365-2486.2003.00701.x, 2004.

Jolly, W. M., Nemani, R., and Running, S. W.: A generalized, bioclimatic index to predict foliar phenology in response to climate, Global Change Biology, 11, 619–632, https://doi.org/10.1111/j.1365-2486.2005.00930.x, 2005.

Kaminski, T., Knorr, W., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P. P.: Consistent assimilation of MERIS FAPAR and atmospheric CO2 into a terrestrial vegetation model and interactive mission benefit analysis, Biogeosciences, 9, 3173–3184, https://doi.org/10.5194/bg-9-3173-2012, 2012.

Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P. J., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R., Gobron, N., Grant, J. P., Heimann, M., Hooker-Stroud, A., Houweling, S., Kato, T., Kattge, J., Kelley, D., Kemp, S., Koffi, E. N., Köstler, C., Mathieu, P.-P., Pinty, B., Reick, C. H., Rödenbeck, C., Schnur, R., Scipal, K., Sebald, C., Stacke, T., van Scheltinga, A. T., Vossbeck, M., Widmann, H., and Ziehn, T.: The BETHY/JSBACH Carbon Cycle Data Assimilation System: experiences and challenges, Journal of Geophysical Research: Biogeosciences, 118, 1414–1426, https://doi.org/10.1002/jgrg.20118, 2013.

Kato, T., Knorr, W., Scholze, M., Veenendaal, E., Kaminski, T., Kattge, J., and Gobron, N.: Simultaneous assimilation of satellite and eddy covariance data for improving terrestrial water and carbon simulations at a semi-arid woodland site in Botswana, Biogeosciences, 10, 789–802, https://doi.org/10.5194/bg-10-789-2013, 2013.

Keenan, T. F., Richardson, A. D., and Hufkens, K.: On quantifying the apparent temperature sensitivity of plant phenology, New Phytologist, 225, 1033–1040, https://doi.org/10.1111/nph.16114, 2020.

Knorr, W., Kaminski, T., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P. P.: Carbon cycle data assimilation with a generic phenology model, Journal of Geophysical Research: Biogeosciences, 115, 1–16, https://doi.org/10.1029/2009JG001119, 2010.

Levine, P., Bilir, E., Bloom, A. A., Braghiere, R., Famiglietti, C., Konings, A., Longo, M., Ma, S., Massoud, E., Meyer, V., Norton, A. J., Parazoo, N., Quetin, G., Smallman, L., Williams, M., Worden, J., Worden, M., Worden, S., and Yang, Y.: Constraining carbon, water, and energy cycling using diverse Earth observations across scales: the CARDAMOM 3.0 approach, in: EGU General Assembly 2023, EGU23-10918, Vienna, Austria, https://doi.org/10.5194/egusphere-egu23-10918, 2023.

MacBean, N., Peylin, P., Chevallier, F., Scholze, M., and Schürmann, G.: Consistent assimilation of multiple data streams in a carbon cycle data assimilation system, Geoscientific Model Development, 9, 3569–3588, https://doi.org/10.5194/gmd-9-3569-2016, 2016.

Mahowald, N., Lo, F., Zheng, Y., Harrison, L., Funk, C., Lombardozzi, D., and Goodale, C.: Projections of leaf area index in earth system models, Earth System Dynamics, 7, 211–229, https://doi.org/10.5194/esd-7-211-2016, 2016.

Manzoni, S., Vico, G., Thompson, S., Beyer, F., and Weih, M.: Contrasting leaf phenological strategies optimize carbon gain under droughts of different duration, Advances in Water Resources, 84, 37–51, https://doi.org/10.1016/j.advwatres.2015.08.001, 2015.

Marchand, L. J., Dox, I., Gričar, J., Prislan, P., Leys, S., Van den Bulcke, J., Fonti, P., Lange, H., Matthysen, E., Peñuelas, J., Zuccarini, P., and Campioli, M.: Inter-individual variability in spring phenology of temperate deciduous trees depends on species, tree size and previous year autumn phenology, Agricultural and Forest Meteorology, 290, 108 031, https://doi.org/10.1016/j.agrformet.2020.108031, 2020.

Martínez-Vilalta, J., Sala, A., Asensio, D., Galiano, L., Hoch, G., Palacio, S., Piper, F. I., and Lloret, F.: Dynamics of non-structural carbohydrates in terrestrial plants: a global synthesis, Ecological Monographs, 86, 495–516, https://doi.org/10.1002/ecm.1231, 2016.

Massoud, E. C., Bloom, A. A., Longo, M., Reager, J. T., Levine, P. A., and Worden, J. R.: Information content of soil hydrology in a west Amazon watershed as informed by GRACE, Hydrology and Earth System Sciences, 26, 1407–1423, https://doi.org/10.5194/hess-26-1407-2022, 2022.

Migliavacca, M., Sonnentag, O., Keenan, T. F., Cescatti, A., O'Keefe, J., and Richardson, A. D.: On the uncertainty of phenological responses to climate change, and implications for a terrestrial biosphere model, Biogeosciences, 9, 2063–2083, https://doi.org/10.5194/bg-9-2063-2012, 2012.

Norton, A., Bloom, A. A., Parazoo, N. C., Levine, P. A., Ma, S., Braghiere, R. K., and Smallman, L. T.: CARDAMOM Phenology Study: Dataset and Analysis Code, https://doi.org/10.5281/zenodo.7793974, 2023.

Norton, A. J., Rayner, P. J., Koffi, E. N., and Scholze, M.: Assimilating solar-induced chlorophyll fluorescence into the terrestrial biosphere model BETHY-SCOPE v1.0: Model description and information content, Geoscientific Model Development, 11, 1517–1536, https://doi.org/10.5194/gmd-11-1517-2018, 2018.

Norton, A. J., Rayner, P. J., Koffi, E. N., Scholze, M., Silver, J. D., and Wang, Y. P.: Estimating global gross primary productivity using chlorophyll fluorescence and a data assimilation system with the BETHY-SCOPE model, Biogeosciences, 16, 3069–3093, https://doi.org/10.5194/bg-16-3069-2019, 2019.

Parazoo, N. C., Barnes, E., Worden, J., Harper, A. B., Bowman, K. B., Frankenberg, C., Wolf, S., Litvak, M., and Keenan, T. F.: Influence of ENSO and the NAO on terrestrial carbon uptake in the Texas-northern Mexico region, Global Biogeochemical Cycles, 29, 1247–1265, https://doi.org/10.1002/2015GB005125, 2015.

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., Isaac, P., Polidori, D., Reichstein, M., Ribeca, A., van Ingen, C., Vuichard, N., Zhang, L., Amiro, B., Ammann, C., Arain, M. A., Ardö, J., Arkebauer, T., Arndt, S. K., Arriga, N., Aubinet, M., Aurela, M., Baldocchi, D., Barr, A., Beamesderfer, E., Marchesini, L. B., Bergeron, O., Beringer, J., Bernhofer, C., Berveiller, D., Billesbach, D., Black, T. A., Blanken, P. D., Bohrer, G., Boike, J., Bolstad, P. V., Bonal, D., Bonnefond, J.-M., Bowling, D. R., Bracho, R., Brodeur, J., Brümmer, C., Buchmann, N., Burban, B., Burns, S. P., Buysse, P., Cale, P., Cavagna, M., Cellier, P., Chen, S., Chini, I., Christensen, T. R., Cleverly, J., Collalti, A., Consalvo, C., Cook, B. D., Cook, D., Coursolle, C., Cremonese, E., Curtis, P. S., D'Andrea, E., da Rocha, H., Dai, X., Davis, K. J., Cinti, B. D., de Grandcourt, A., Ligne, A. D., De Oliveira, R. C., Delpierre, N., Desai, A. R., Di Bella, C. M., di Tommasi, P., Dolman, H., Domingo, F., Dong, G., Dore, S., Duce, P., Dufrêne, E., Dunn, A., Dušek, J., Eamus, D., Eichelmann, U., ElKhidir, H. A. M., Eugster, W., Ewenz, C. M., Ewers, B., Famulari, D., Fares, S., Feigenwinter, I., Feitz, A., Fensholt, R., Filippa, G., Fischer, M., Frank, J., Galvagno, M., Gharun, M., Gianelle, D., Gielen, B., Gioli, B., Gitelson, A., Goded, I., Goeckede, M., Goldstein, A. H., Gough, C. M., Goulden, M. L., Graf, A., Griebel, A., Gruening, C., Grünwald, T., Hammerle, A., Han, S., Han, X., Hansen, B. U., Hanson, C., Hatakka, J., He, Y., Hehn, M., Heinesch, B., Hinko-Najera, N., Hörtnagl, L., Hutley, L., Ibrom, A., Ikawa, H., Jackowicz-Korczynski, M., Janouš, D., Jans, W., Jassal, R., Jiang, S., Kato, T., Khomik, M., Klatt, J., Knohl, A., Knox, S., Kobayashi, H., Koerber, G., Kolle, O., Kosugi, Y., Kotani, A., Kowalski, A., Kruijt, B., Kurbatova, J., Kutsch, W. L., Kwon, H., Launiainen, S., Laurila, T., Law, B., Leuning, R., Li, Y., Liddell, M., Limousin, J.-M., Lion, M., Liska, A. J., Lohila, A., López-Ballesteros, A., López-Blanco, E., Loubet, B., Loustau, D., Lucas-Moffat, A., Lüers, J., Ma, S., Macfarlane, C., Magliulo, V., Maier, R., Mammarella, I., Manca, G., Marcolla, B., Margolis, H. A., Marras, S., Massman, W., Mastepanov, M., Matamala, R., Matthes, J. H., Mazzenga, F., McCaughey, H., McHugh, I., McMillan, A. M. S., Merbold, L., Meyer, W., Meyers, T., Miller, S. D., Minerbi, S., Moderow, U., Monson, R. K., Montagnani, L., Moore, C. E., Moors, E., Moreaux, V., Moureaux, C., Munger, J. W., Nakai, T., Neirynck, J., Nesic, Z., Nicolini, G., Noormets, A., Northwood, M., Nosetto, M., Nouvellon, Y., Novick, K., Oechel, W., Olesen, J. E., Ourcival, J.-M., Papuga, S. A., Parmentier, F.-J., Paul-Limoges, E., Pavelka, M., Peichl, M., Pendall, E., Phillips, R. P., Pilegaard, K., Pirk, N., Posse, G., Powell, T., Prasse, H., Prober, S. M., Rambal, S., Rannik, Ü., Raz-Yaseef, N., Rebmann, C., Reed, D., de Dios, V. R., Restrepo-Coupe, N., Reverter, B. R., Roland, M., Sabbatini, S., Sachs, T., Saleska, S. R., Sánchez-Cañete, E. P., Sanchez-Mejia, Z. M., Schmid, H. P., Schmidt, M., Schneider, K., Schrader, F., Schroder, I., Scott, R. L., Sedlák, P., Serrano-Ortíz, P., Shao, C., Shi, P., Shironya, I., Siebicke, L., Šigut, L., Silberstein, R., Sirca, C., Spano, D., Steinbrecher, R., Stevens, R. M., Sturtevant, C., Suyker, A., Tagesson, T., Takanashi, S., Tang, Y., Tapper, N., Thom, J., Tomassucci, M., Tuovinen, J.-P., Urbanski, S., Valentini, R., van der Molen, M., van Gorsel, E., van Huissteden, K., Varlagin, A., Verfaillie, J., Vesala, T., Vincke, C., Vitale, D., Vygodskaya, N., Walker, J. P., Walter-Shea, E., Wang,

825   H., Weber, R., Westermann, S., Wille, C., Wofsy, S., Wohlfahrt, G., Wolf, S., Woodgate, W., Li, Y., Zampedri, R., Zhang, J., Zhou, G., Zona, D., Agarwal, D., Biraud, S., Torn, M., and Papale, D.: The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data, Scientific Data, 7, 225, https://doi.org/10.1038/s41597-020-0534-3, 2020.

Piao, S., Sitch, S., Ciais, P., Friedlingstein, P., Peylin, P., Wang, X., Ahlström, A., Anav, A., Canadell, J. G., Cong, N., Huntingford, C., Jung, M., Levis, S., Levy, P. E., Li, J., Lin, X., Lomas, M. R., Lu, M., Luo, Y., Ma, Y., Myneni, R. B., Poulter, B., Sun, Z., Wang, T., Viovy, N.,
830   Zaehle, S., and Zeng, N.: Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO 2 trends, Global Change Biology, 19, 2117–2132, https://doi.org/10.1111/gcb.12187, 2013.

Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., and Zhu, X.: Plant phenology and global climate change: Current progresses and challenges, Global Change Biology, 25, 1922–1940, https://doi.org/10.1111/gcb.14619, 2019.

Piao, S., Wang, X., Wang, K., Li, X., Bastos, A., Canadell, J. G., Ciais, P., Friedlingstein, P., and Sitch, S.: Interannual variation of terrestrial
835   carbon cycle: Issues and perspectives, Global Change Biology, 26, 300–318, https://doi.org/10.1111/gcb.14884, 2020.

Prentice, I. C., Liang, X., Medlyn, B. E., and Wang, Y.-P.: Reliable, robust and realistic: the three R's of next-generation land-surface modelling, Atmospheric Chemistry and Physics, 15, 5987–6005, https://doi.org/10.5194/acp-15-5987-2015, 2015.

Quetin, G. R., Bloom, A. A., Bowman, K. W., and Konings, A. G.: 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,
840   https://doi.org/10.1029/2019MS001889, 2020.

Rayner, P. J., Michalak, A. M., and Chevallier, F.: Fundamentals of data assimilation applied to biogeochemistry, Atmospheric Chemistry and Physics, 19, 13 911–13 932, https://doi.org/10.5194/acp-19-13911-2019, 2019.

Richardson, A. D., Anderson, R. S., Arain, M. A., Barr, A. G., Bohrer, G., Chen, G., Chen, J. M., Ciais, P., Davis, K. J., Desai, A. R., Dietze, M. C., Dragoni, D., Garrity, S. R., Gough, C. M., Grant, R., Hollinger, D. Y., Margolis, H. A., Mccaughey, H., Migliavacca, M., Monson,
845   R. K., Munger, J. W., Poulter, B., Raczka, B. M., Ricciuto, D. M., Sahoo, A. K., Schaefer, K., Tian, H., Vargas, R., Verbeeck, H., Xiao, J., and Xue, Y.: Terrestrial biosphere models need better representation of vegetation phenology: Results from the North American Carbon Program Site Synthesis, Global Change Biology, 18, 566–584, https://doi.org/10.1111/j.1365-2486.2011.02562.x, 2012.

Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O., and Toomey, M.: Climate change, phenology, and phenological control of vegetation feedbacks to the climate system, Agricultural and Forest Meteorology, 169, 156–173,
850   https://doi.org/10.1016/j.agrformet.2012.09.012, 2013.

Schiestl-Aalto, P., Kulmala, L., Mäkinen, H., Nikinmaa, E., and Mäkelä, A.: CASSIA - a dynamic model for predicting intra-annual sink demand and interannual growth variation in Scots pine, New Phytologist, 206, 647–659, https://doi.org/10.1111/nph.13275, 2015.

Schimel, D. and Schneider, F. D.: Flux towers in the sky: global ecology from space, New Phytologist, 224, 570–584, https://doi.org/10.1111/nph.15934, 2019.

855 Schwalm, C. R., Schaefer, K., Fisher, J. B., Huntzinger, D., Elshorbany, Y., Fang, Y., Hayes, D., Jafarov, E., Michalak, A. M., Piper, M., Stofferahn, E., Wang, K., and Wei, Y.: Divergence in land surface modeling: linking spread to structure, Environmental Research Communications, 1, 111 004, https://doi.org/10.1088/2515-7620/ab4a8a, 2019.

Seiler, C., Melton, J. R., Arora, V. K., Sitch, S., Friedlingstein, P., Anthoni, P., Goll, D., Jain, A. K., Joetzjer, E., Lienert, S., Lombardozzi, D., Luyssaert, S., Nabel, J. E. M. S., Tian, H., Vuichard, N., Walker, A. P., Yuan, W., and Zaehle, S.: Are Terrestrial Biosphere Models Fit for
860   Simulating the Global Land Carbon Sink?, Journal of Advances in Modeling Earth Systems, 14, https://doi.org/10.1029/2021MS002946, 2022.

Sellers, P. J., Dickinson, R. E., Randall, D. A., Betts, A. K., Hall, F. G., Berry, J. A., Collatz, G. J., Denning, A. S., Mooney, H. A., Nobre, C. A., Sato, N., Field, C. B., and Henderson-Sellers, A.: Modeling the Exchanges of Energy, Water, and Carbon Between Continents and the Atmosphere, Science, 275, 502–509, https://doi.org/10.1126/science.275.5299.502, 1997.

865 Smallman, T. L. and Williams, M.: Description and validation of an intermediate complexity model for ecosystem photosynthesis and evapotranspiration: ACM-GPP-ETv1, Geoscientific Model Development, 12, 2227–2253, https://doi.org/10.5194/gmd-12-2227-2019, 2019.

Smallman, T. L., Milodowski, D. T., Neto, E. S., Koren, G., Ometto, J., and Williams, M.: Parameter uncertainty dominates C-cycle forecast errors over most of Brazil for the 21st century, Earth System Dynamics, 12, 1191–1237, https://doi.org/10.5194/esd-12-1191-2021, 2021.

Stöckli, R., Rutishauser, T., Dragoni, D., O'Keefe, J., Thornton, P. E., Jolly, M., Lu, L., and Denning, A. S.: Remote sensing data assimilation
870 for a prognostic phenology model, Journal of Geophysical Research: Biogeosciences, 113, 1–19, https://doi.org/10.1029/2008JG000781, 2008.

Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F., and Watkins, M. M.: GRACE Measurements of Mass Variability in the Earth System, Science, 305, 503–505, https://doi.org/10.1126/science.1099192, 2004.

Ter Braak, C. J. F.: A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real
875 parameter spaces, Statistics and Computing, 16, 239–249, https://doi.org/10.1007/s11222-006-8769-1, 2006.

Trugman, A. T., Detto, M., Bartlett, M. K., Medvigy, D., Anderegg, W. R. L., Schwalm, C., Schaffer, B., and Pacala, S. W.: Tree carbon allocation explains forest drought-kill and recovery patterns, Ecology Letters, 21, 1552–1560, https://doi.org/10.1111/ele.13136, 2018.

Van Bodegom, P. M., Douma, J. C., Witte, J. P. M., Ordoñez, J. C., Bartholomeus, R. P., and Aerts, R.: Going beyond limitations of plant functional types when predicting global ecosystem-atmosphere fluxes: exploring the merits of traits-based approaches, Global Ecology
880 and Biogeography, 21, 625–636, https://doi.org/10.1111/j.1466-8238.2011.00717.x, 2012.

Verger, A., Baret, F., and Weiss, M.: Near Real-Time Vegetation Monitoring at Global Scale, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7, 3473–3481, https://doi.org/10.1109/JSTARS.2014.2328632, 2014.

Viskari, T., Hardiman, B., Desai, A. R., and Dietze, M. C.: Model-data assimilation of multiple phenological observations to constrain and predict leaf area index, Ecological Applications, 25, 546–558, https://doi.org/10.1890/14-0497.1, 2015.

885 Wheeler, K. I. and Dietze, M. C.: Improving the monitoring of deciduous broadleaf phenology using the Geostationary Operational Environmental Satellite (GOES) 16 and 17, Biogeosciences, 18, 1971–1985, https://doi.org/10.5194/bg-18-1971-2021, 2021.

Williams, M., Rastetter, E. B., Fernandes, D. N., Goulden, M. L., Shaver, G. R., and Johnson, L. C.: Predicting gross primary productivity in terrestrial ecosystems, Ecological Applications, 7, 882–894, https://doi.org/10.1890/1051-0761(1997)007[0882:PGPPIT]2.0.CO;2, 1997.

Williams, M., Schwarz, P. A., Law, B. E., Irvine, J., and Kurpius, M. R.: An improved analysis of forest carbon dynamics using data
890 assimilation, Global Change Biology, 11, 89–105, https://doi.org/10.1111/j.1365-2486.2004.00891.x, 2005.

Wu, J., Albert, L. P., Lopes, A. P., Restrepo-Coupe, N., Hayek, M., Wiedemann, K. T., Guan, K., Stark, S. C., Christoffersen, B., Prohaska, N., Tavares, J. V., Marostica, S., Kobayashi, H., Ferreira, M. L., Campos, K. S., da Silva, R., Brando, P. M., Dye, D. G., Huxman, T. E., Huete, A. R., Nelson, B. W., and Saleska, S. R.: Leaf development and demography explain photosynthetic seasonality in Amazon evergreen forests, Science, 351, 972–976, https://doi.org/10.1126/science.aad5068, 2016.

895 Xin, Q., Zhou, X., Wei, N., Yuan, H., Ao, Z., and Dai, Y.: A Semiprognostic Phenology Model for Simulating Multidecadal Dynamics of Global Vegetation Leaf Area Index, Journal of Advances in Modeling Earth Systems, 12, 1–17, https://doi.org/10.1029/2019MS001935, 2020.

Yang, J., Medlyn, B. E., De Kauwe, M. G., and Duursma, R. A.: Applying the Concept of Ecohydrological Equilibrium to Predict Steady State Leaf Area Index, Journal of Advances in Modeling Earth Systems, 10, 1740–1758, https://doi.org/10.1029/2017MS001169, 2018.

Yang, Y., Bloom, A. A., Ma, S., Levine, P., Norton, A., Parazoo, N. C., Reager, J. T., Worden, J., Quetin, G. R., Smallman, T. L., Williams, M., Xu, L., and Saatchi, S.: CARDAMOM-FluxVal version 1.0: a FLUXNET-based validation system for CARDAMOM carbon and water flux estimates, Geoscientific Model Development, 15, 1789–1802, https://doi.org/10.5194/gmd-15-1789-2022, 2022.

Yin, Y., Bloom, A. A., Worden, J., Saatchi, S., Yang, Y., Williams, M., Liu, J., Jiang, Z., Worden, H., Bowman, K., Frankenberg, C., and Schimel, D.: Fire decline in dry tropical ecosystems enhances decadal land carbon sink, Nature Communications, 11, 1900, https://doi.org/10.1038/s41467-020-15852-2, 2020.

Zhang, H., Yuan, W., Dong, W., and Liu, S.: Seasonal patterns of litterfall in forest ecosystem worldwide, Ecological Complexity, 20, 240–247, https://doi.org/10.1016/j.ecocom.2014.01.003, 2014.

Zhang, Y., Parazoo, N. C., Williams, A. P., Zhou, S., and Gentine, P.: Large and projected strengthening moisture limitation on end-of-season photosynthesis, Proceedings of the National Academy of Sciences, 117, 9216–9222, https://doi.org/10.1073/pnas.1914436117, 2020.