Author responses and manuscript changes for "Constraining ecosystem carbon dynamics in a data-limited world: integrating ecological "common sense" in a model-data-fusion framework." by A. A. Bloom and M. Williams

A. A. Bloom & M. Williams

abloom@jpl.nasa.gov

We thank the referees for having provided thorough feedback and for their suggested corrections. Below we have addressed each individual comment from both referees (referee comments are shown in italics). We also denote all manuscript changes: line numbers correspond to the revised marked-up manuscript included below.

Anonymous Referee #1

Bloom and Williams report that incorporating internal 'reality constraints' on model process relations reduces the range of permissible parameter values in a terrestrial ecosystem model. They also report that the use of these reality constraints additionally improves model performance when compared to measured eddy-covariance flux observations out of sample.

The manuscript is very well written, and the approach intuitive and reasonable. The results clearly demonstrate that introducing these additional reality constraints reduces parameter uncertainty. This is a clear result and indeed including such reality constraints in any model endeavor (be it data assimilation or more traditional model assessment) should be standard practice.

My only issue with the results presented is that the model that uses reality constraints does almost too well when compared against eddy-covariance data. In figure 5 we see that it captures the magnitude and seasonal cycle of net ecosystem exchange almost perfectly at two sites, compared to the model that does not use reality constraints. Both model runs use MODIS leaf area index and soil carbon as constraints, but not the eddy-covariance data.

The authors are therefore claiming that with only information on LAI, soil carbon and some general bounds based on how ecosystems are typically structured, we can predict carbon cycling on seasonal and annual timescales. This is quite remarkable given that in a previous study that also included some measure of reality constraints, and a host of other constraints at one of the sites used here (Howland forest; Richardson et al. 2010), the DALEC model

had difficulty in capturing the annual total NEE (i.e. only when annual NEE was used as a constraint, despite being optimized to daily NEE and various other biometric constraints). It is also remarkable in that it suggests that other typically key information such as above ground biomass, photosynthetic potential, soil moisture status, and canopy structure differences between evergreen and deciduous sites (i.e. site specific ACM), are not essential for predicting carbon uptake.

We are confident in the results of our experiments. We link the improved performance particularly to the ecological and dynamic constraints (EDCs) we have introduced our new EDC analyses, as suggested below, help to define the contributions of individual EDCs more clearly.

A lacking component in the manuscript is the identification of which of the reality constraints is responsible for the improved model performance.

We agree with the reviewer's recommendation: to identify which ecological and dynamic constraints (EDCs) have resulted in improved model performance, we have conducted an EDC sensitivity test. We now show which EDCs (a) lead to improved parameter estimates, (b) lead to reduced net ecosystem exchange (NEE) confidence ranges and (c) lead to reduced NEE bias.

In the revised manuscript we have included the above-mentioned sensitivity analysis (the sensitivity analysis is described in lines 320-338; the results are presented in Table 2 and described in lines 351-355, 391-396; the results are discussed in lines 419-429).

It is also not clear why the range of annual model carbon cycling not centered around equilibrium, given the wide range of parameter values used, and information only on soil carbon and leaf area, and a forest typical structure.

For both for synthetic and AmeriFlux experiments, the posterior probability density functions of NEE (e.g. Figures 3, 4 and 5) show that ecosystems could be either net sources or sinks of carbon on annual timescales. Therefore, our results demonstrate that soil carbon and LAI are not sufficient to resolve whether each AmeriFlux site is a net source or sink of carbon on annual timescales.

We now explicitly state this in the discussion section of the revised

manuscript (lines 453-456).

Introduction:

The concept of using internal model constraints, here termed ecological and dynamic constraints, was first introduced by Richardson et al. 2010, there termed a reality constraint. This should be acknowledged in the introduction.

In the revised manuscript we acknowledge that Richardson et al., (2010) introduced internal model constraints in carbon cycle model-data fusion analyses (lines 84-86).

Page 12736, line 25: "therefore. . .". Consider revising this sentence. It does not logically flow from the paragraph.

We agree with the reviewer, and we have revised this sentence (line 67).

Page 12738, line 17: Please do not refer to DALEC2 as a universal ecosystem carbon balance model. It is designed for temperate deciduous and evergreen forests, and will not likely accurately simulate other ecosystem flux dynamics (e.g., tundra, tropical, peatlands, savannah, etc.). Page 12738: Please state the drivers used in the DALEC2 model.

We acknowledge the reviewer's point, and we have re-worded the DALEC2 description (line 112).

Page 12739, line 21: Please clarify that omega here represents a turnover rate. What is OmegaMin?

Equation 5: Clarify what f signifies here.

In the revised manuscript, we have now added an explicit reference to Table 1, where all DALEC2 parameters, notations and ranges are reported (lines 137-140).

Page 12746, line 17-20: Clarify the site selection criteria here. Both Howland and Sylvania have snow cover for far more than two months, which would appear to invalidate the selection criteria based on hydrological concerns.

We agree with the reviewer's remark and acknowledge our oversight. We now clarify that the selected sites exhibit limited water stress and $\leq$ 3 months of below-freezing soil temperatures (lines 279-284).

Page 12747, line 1-10: Please report the values of LAI and soil carbon used for each site.

We now report the 5^{th} and 95^{th} percentile LAI values and the soil carbon value used for each AmeriFlux site experiment (lines 300-302, 308,309).

Page 12748, line 3,5: Please do not confuse error with uncertainty. Parameter vectors have uncertainties, not errors, unless compared against known parameter values. This confusion is apparent throughout the manuscript.

Page 12748, line 14: 'and hence improved estimates of s'. I would argue that what you are really reporting are better constrained estimates of s, though the true values of s are remain unknown.

We agree with the reviewer's two points on error and uncertainty: however, the synthetic datasets are derived from known parameter values s. To better convey this point, we now explicitly state this in the introduction to synthetic experiments (lines 235-241).

Figure 5: I would suggest plotting all three graphs on the same scale to assist between site comparison

In the revised manuscript, we have now plotted all three graphs in Figure 5 on the same scale.

Anonymous Referee #2

The manuscript by Bloom and Williams proposes to include known model parameter relationships in a data assimilation framework in addition to observations. They claim that in a data-poor context these additional constraints will reduce parameter uncertainties. In general I agree with this statement. However, in my opinion the ecological and dynamic constraints (EDCs) that the authors introduce as a novelty are simply part of the prior information we possess for these parameters. I would suggest that the authors highlight this in the manuscript.

The manuscript is well written and presented, but I think some improvements and clarifications are required (see specific comments below).

Specific comments:

In order to obtain a unique solution in an ill-posed problem additional con-

straints are required. This is also known as regularization. Within the Bayesian framework prior parameter information are usually included in form of a covariance matrix, which can include correlations between parameters. The authors mention in the introduction that such correlations limit the possible parameter configuration, but in their example they simply assume no prior knowledge other than the parameter ranges. This seems to be an odd choice, because it means that all values within the given range are equally likely and parameters are independent, which is clearly not the case. The parameter space has not been restricted and it is therefore not surprising that additional information in form of ECDs add large constraints to this problem. I am wondering if this would also be the case if a different prior parameter distribution (i.e. Gaussian) with a defined covariance matrix would have been chosen in the first place. I see the ECDs complementary to the knowledge we include in terms of prior distribution and covariance matrix and not as a replacement.

We agree with the reviewer's point: in contrast to a "flat" parameter prior, a parameter variance-covariance structure would serve as an additional constraint on model parameters, and would reduce the ill-posedness of the problem. However, given that we have poor quantitative knowledge on the realistic values for model parameters and their covariances, constructing a generic, ecologically-appropriate covariance structure is exceedingly difficult. For example, most parameter inter-dependencies presented in our manuscript are dependent on local meteorology: therefore, a meteorology-dependent prior parameter covariance matrix would need to be derived for each AmeriFlux site. By prescribing EDCs, we are instead able to impose ecological knowledge in the form of non-Gaussian state and parameter constraints. We agree with the reviewer that EDCs and a parameter covariance matrix can both be used to resolve ill-posed carbon cycle problems. In the revised manuscript, we now state that prior parameter covariance structures can be used as alternative or complementary constraints to EDCs (lines 446-449).

A number of ECDs are formulated to constrain the parameters and states and it would be interesting to know their individual contribution, i.e. which ECD provides the largest constraint.

We agree with the reviewer's recommendation. We have performed an EDC sensitivity test, whereby we quantify the improvements in model parameter and state estimates associated with each EDC.

In the revised manuscript we have included the above-mentioned sensitivity analysis (the sensitivity analysis is described in lines 320-338; the results are presented in Table 2 and described in lines 351-355, 391-396; the results are discussed in lines 419-429).

Minor comments:

We have implemented all of the following suggested corrections. In particular, the ˜ on P12745, Eq.(16), denotes the median value of $\boldsymbol{E}$(the ˜ was missing from $\boldsymbol{E}$ on P12745 line 18: we have now corrected the text).

P12738,L2 + P12739,L 14: EDC has already been introduced in the abstract and in- troduction (P12737,L18)

(correction on line 99)

P12744, L15 + L20: repetition "We create 40 synthetic experiments ..."

(corrections and re-wording on lines 236,237,241)

P12745, Eq.(16): What is ˜ been used for?

(correction on line 263)

P12759, L1: space between 8 and daily

(correction on line 609)

We have also corrected a minor oversight in the prior parameter ranges shown in Table 1: we used 20-2000 gC m^{-2} for the foliar, labile, fine root and litter carbon pools, and 20-150 day for the leaf-fall period parameter ranges. We have corrected this in the revised manuscript.

(correction in Table 1)

Manuscript prepared for Biogeosciences
with version 5.0 of the LaTeX class copernicus.cls.
Date: 7 January 2015

Constraining ecosystem carbon dynamics in a data-limited world: integrating ecological "common sense" in a model-data-fusion framework.

A. A. Bloom[1,2] and M. Williams[1]

[1]School of GeoSciences, University of Edinburgh, Edinburgh, UK
[2]Now at Jet Propulsion Laboratory, California Institute of Technology, Pasadena, USA

Correspondence to: A. A. Bloom (abloom@jpl.nasa.gov)

Abstract. Many of the key processes represented in global terrestrial carbon models remain largely unconstrained. For instance, plant allocation patterns and residence times of carbon pools are poorly known globally, except perhaps at a few intensively studied sites. As a consequence of data scarcity, carbon models tend to be underdetermined, and so can produce similar net fluxes with very different parameters and internal dynamics. To address these problems, we propose a series of ecological and dynamic constraints (EDCs) on model parameters and initial conditions, as a means to constrain ecosystem variable inter-dependencies in the absence of local data. The EDCs consist of a range of conditions on (a) carbon pool turnover and allocation ratios, (b) steady state proximity, and (c) growth and decay of model carbon pools. We use a simple ecosystem carbon model in a model-data fusion framework to determine the added value of these constraints in a data-poor context. Based only on leaf area index (LAI) time series and soil carbon data, we estimate net ecosystem exchange (NEE) for (a) 40 synthetic experiments and (b) three ~~AMERIFLUX~~ AmeriFlux tower sites. For the synthetic experiments, we show that EDCs lead to an an overall 34 % relative error reduction in model parameters, and a 65 % reduction in the 3 yr NEE 90 % confidence range. In the application at ~~AMERIFLUX~~ AmeriFlux sites all NEE estimates were made independently of NEE measurements. Compared to these observations, EDCs resulted in a 69–93 % reduction in 3 yr cumulative NEE median biases (-0.26 to $+0.08\,\mathrm{kg\,C\,m^{-2}}$), in comparison to standard 3 yr median NEE biases (-1.17 to $-0.84\,\mathrm{kg\,C\,m^{-2}}$). In light of these findings, we advocate the use of EDCs in future model-data fusion analyses of the terrestrial carbon cycle.

1 Introduction

Terrestrial ecosystem carbon exchange is a fundamental part of the global carbon cycle link to biosphere processes. Atmospheric CO_2 measurements indicate the presence of a global land C sink, i.e. uptake by the terrestrial biosphere exceeds losses. However, relative to all major terms in the global carbon budget, the global land sink exhibits both the largest inter-annual variability and the largest uncertainty (Le Quéré et al., 2013). The terrestrial carbon budget uncertainty stems largely from unknowns in the size, spatial distribution and temporal dynamics of the major terrestrial carbon pools. As a result, there is little agreement among modelled land sink projections for the 21 st century (Todd-Brown et al., 2013; Friend et al., 2013), reflecting uncertainty in knowledge on the current state of the terrestrial C cycle and its dynamics.

In recent years a growing volume of data from flux towers, satellites and plant trait databases has been used to constrain some of the key components of the terrestrial carbon cycle (e.g. Baldocchi et al., 2001; Simard et al., 2011; Kattge et al., 2011). In particular, a range of ecosystem carbon models and datasets have been brought together in model-data fusion (MDF) frameworks to produce an enhanced analysis of ecosystem carbon cycling (e.g. Williams et al., 2005; Fox et al., 2009; Carvalhais et al., 2010; Luo et al., 2011; Ziehn et al., 2012; Smith et al., 2013). Where mutliple data streams are available, MDF approaches can provide an extensive insight into carbon pool dynamics, turnover rates, and carbon allocation fractions (Richardson et al., 2010; Keenan et al., 2013). However, even at research intensive sites, MDF studies can produce a wide range of acceptable model parameter sets, due to under-determination of the carbon budget with available data. Some of these optimized parameter sets, even though they generate realistic fluxes over short timescales, are associated with major changes to larger carbon pools (soil, wood) that are nonsensical (Fox et al., 2009). For regional and global scale model implementation, the lack of in-situ measurements amplifies this problem, sometimes referred to as equifinality (Beven and Freer, 2001). Ultimately, we need to overcome data limitations and under-determination by integrating models and ecosystem knowledge in a common framework. This framework must ensure ecologically realistic outcomes, while still encompassing (i.e. effectively quantifying) the uncertainty associated with parameter estimation given observation errors (Hill et al., 2012).

Although a range of process-based models have been used to represent the dynamics of the terrestrial carbon cycle and land–atmosphere CO_2 exchange (e.g. Sitch et al., 2008; Schwalm et al., 2010), there are advantages in using simpler models to estimate ecosystem carbon state variables. Firstly, there is a trade-off between model complexity, such as the number of model parameters, and a model's ability to reproduce observations (e.g. Akaike, 1974): therefore a low-complexity model is preferable when it can reproduce ecosystem observations with comparable skill. Secondly, complex models are often computationally expensive, and this is an inhibiting factor when using iterative methods (such as Monte Carlo approaches) to estimate model parameters and their uncertainty. Ideally, the key terms of ecosystem carbon dynamics can be constrained by combining ecosystem

observations with a model of appropriate complexity in a computationally efficient MDF framework.

Previous MDF studies have invariably relied on net ecosystem exchange (NEE) measurements (real and synthetic), along with other site-level observations (Williams et al., 2009). In a global context, the FLUXNET flux-tower network (Baldocchi et al., 2001) consists of hundreds of flux tower sites where hectare-scale NEE measurements have been made over the past two decades. In addition to NEE, complimentary site-level biometric data can help resolve model parameters and state variables in an MDF context (Richardson et al., 2010; Hill et al., 2012; Keenan et al., 2013), alleviating the problem of under-determination. However, the terrestrial biosphere will inevitably remain poorly sampled by FLUXNET. Alternative estimates of NEE from atmospheric CO_2 measurements (e.g. Peters et al., 2010; Feng et al., 2011) are only produced at continental-scale resolutions. ~~Therefore,~~ Therefore, given the limited span of the FLUXNET flux-tower network, are spatially resolved global carbon cycle analyses limited by the sparsity of eddy flux and biometric data?

NEE, the difference between photosynthesis and ecosystem respiration, is a function of the dynamics of all carbon pools over a range of timescales. In the absence of NEE observations, model NEE estimates depend on a knowledge of carbon pool sizes and model parameter values. In reality, carbon pools and model parameters (especially those related to plant allocation fractions and pool turnover rates) are poorly constrained, and therefore NEE estimates are subject to a comparably large uncertainty. Nonetheless, fundamental knowledge on ecosystem behaviour can potentially be used to overcome the lack of location specific data or parameter values. For example, while parameters related to phenology, C allocation and turnover may vary across multiple orders of magnitude (Kattge et al., 2011; Fox et al., 2009), these parameters are strongly correlated (e.g. Sloan et al., 2013), and the range of possible parameter configurations is therefore limited. Such examples include correlations between leaf lifespan and leaf mass per area (Wright et al., 2004), leaf area index and total foliar N (Williams and Rastetter, 1999), and between foliar and root biomass (Sloan et al., 2013). These correlations can confine parameter searches to a smaller hyper-volume. Equally, while ecosystems exhibit a large range of non-steady state dynamic behaviours, strong inter-relationships are expected between inputs, outputs, carbon pool magnitudes and turnover rates (Luo and Weng, 2011)~~, and model~~ . Richardson et al. (2010) introduced the concept of reality constraints (or internal model constraints) on carbon pool dynamics within a carbon cycle MDF analysis: such constraints on the model state can potentially be used to improve estimates of model parameters. Here we propose that a broad range of model parameter combinations can be discarded when ~~the~~ phenology, carbon allocation, turnover rates and pool dynamics are considered ecologically "nonsensical". Here we seek to address the following question: can we improve ecosystem model parameter and NEE estimates by incorporating ecological "common sense" into carbon cycle MDF analyses?

In this paper we propose a series of Ecological and Dynamic Constraints (EDCs) on model parameters: these include turnover and allocation parameter inter-relations, carbon pool dynamics and steady state proximity conditions (Sect. 2). We quantify the added value of imposing EDCs in syn-

thetic and real data MDF contexts using a simple ecosystem carbon model, by measuring bias and confidence interval reductions of carbon cycle analyses relative to independent data (Sect. 3). Finally we discuss the prospects and limitations of our approach, as well as the implications of a wider EDC implementation in terrestrial carbon cycle MDF methods (Sect. 4).

2 Methods

Here we present a series of ~~ecological and dynamic constraints (EDCs) for~~ EDCs for a daily box budget terrestrial C cycle model, the Data Assimilation Linked Ecosystem Carbon model version two (DALEC2). Within an MDF context, we test the added value of implementing EDCs. Our aims are (1) to quantify our ability to estimate DALEC2 parameters and NEE within a synthetic framework, and (2) to validate our ability to estimate NEE at three temperate forest AmeriFlux sites. We use simulated and real observations of (a) satellite-derived leaf area index (LAI) and (b) soil organic carbon from the Harmonized World Soil Database (HWSD, Hiederer and Köchy, 2012) in our MDF analyses. The choice of these two data sets serves as an analogue for the limited ecosystem carbon datasets available on a global scale.

2.1 DALEC2

DALEC has been extensively used in MDF frameworks (e.g. Williams et al., 2005; Quaife et al., 2008; Richardson et al., 2010, amongst others). In particular, a range of MDF approaches were used in the REFLEX project, where ecosystem observations were assimilated into DALEC to produce carbon state analyses (Fox et al., 2009). Here we use ~~a universal~~ the DALEC2 ecosystem carbon balance model ~~– DALEC2 –~~, which combines components of DALEC evergreen and DALEC deciduous (Williams et al., 2005; Fox et al., 2009) into a single model. Gross primary production (GPP) in DALEC2 is determined from the aggregated canopy model (Williams et al., 1997), and is allocated to the biomass pools (foliar, labile, wood, and fine roots) and to autotrophic respiration (R_a); degraded carbon from biomass pools goes to two dead organic matter pools with temperature dependent losses (heterotrophic respiration, R_h). The net ecosystem exchange is summarised as $\mathrm{NEE} = R_\mathrm{a} + R_\mathrm{h} - \mathrm{GPP}$. C flow in DALEC2 is determined as a function of 23 parameters (including six initial carbon pool states, Table 1). We henceforth refer to the 23 parameters required to initiate DALEC2 as a parameter vector $\boldsymbol{x}$. DALEC2 C pools and fluxes are iteratively calculated at a daily time-step: the DALEC2 model equations are fully described in Appendix A. We henceforth refer to the ensemble of all model state variables (such as daily NEE, GPP, respiration terms and carbon pool trajectories) as $\mathrm{DALEC2}(\boldsymbol{x})$.

2.2 Ecological and Dynamic Constraints

In previous work, DALEC MDF approaches (Williams et al., 2005; Fox et al., 2009; Richardson et al., 2010; Hill et al., 2012) did not explicitly impose any conditions on the inter-relationships between model parameters, therefore parameter prior information had only consisted of prescribed parameter ranges. In reality, broader ecological knowledge can be informative in terms of the inter-relationships between parameter values. For example, long-term leaf turnover rate must be faster than woody biomass turnover (e.g. Norby et al., 2002): such a relationship can provide a relative constraint on model parameter values, without imposing any further constraints to the prior parameter ranges (Table 1).

Here we propose a sequence of ecological and dynamic constraints (EDCs) on DALEC2 parameters and pool dynamics. For any given DALEC2 parameter vector $\boldsymbol{x}$, all EDCs presented in this section (henceforth EDC 1, EDC 2, etc.) are implemented. The probability of parameters (henceforth $P_{\text{EDC}}(\text{DALEC2}(\boldsymbol{x}))$ is 1 if all EDCs are met, otherwise $P_{\text{EDC}}(\text{DALEC2}(\boldsymbol{x})) = 0$. All DALEC2 parameters (allocation fractions f_{auto}, f_{lab},f_{fol}, f_{roo}, f_{woo}; turnover rate parameters θ_{woo}, θ_{roo}, θ_{lit}, θ_{som}, θ_{min}, Θ; canopy parameters d_{onset},d_{fall},c_{eff},c_{lma},c_{lf},c_{ronset},c_{rfall}; carbon pools at time t C^t_{lab}, C^t_{fol},C^t_{woo},C^t_{som},C^t_{lit},C^t_{som}) are described in Table 1.

2.2.1 Turnover Constraints

We impose the following constraints on the relative sizes of turnover rates:

$$\text{EDC 1: } \theta_{\text{som}} < \theta_{\text{lit}}, \tag{1}$$

$$\text{EDC 2: } \theta_{\text{som}} < \theta_{\text{min}}, \tag{2}$$

$$\text{EDC 3: } c_{\text{lf}} > 1 - (1 - \theta_{\text{woo}})^{365.25}, \tag{3}$$

$$\text{EDC 4: } (1 - \theta_{\text{roo}})^N > \Pi_{i=1}^{N}(1 - \theta_{\text{som}} e^{\Theta T_i}), \tag{4}$$

where T_i are daily temperature values during an N-day time window (e.g. three years). These constraints ensure the turnover rate ratios are consistent with knowledge of the carbon pool relative residence times (e.g. Gaudinski et al., 2000; Norby et al., 2002; Trumbore, 2006). In particular, we expect a faster litter turnover in contrast to soil organic matter (SOM) turnover (EDC 1), a faster conversion rate of litter to SOM relative to SOM turnover (EDC 2), the annual leaf loss fraction is greater than the annual woody biomass loss fraction (EDC 3), and a faster fine root turnover in contrast to SOM turnover (EDC 4).

2.2.2 Root-Foliar C allocation constraints

Strong correlations are expected between foliar and fine root carbon pools (e.g. Mokany et al., 2006; Sloan et al., 2013). We constrain the C allocation and dynamics of the root and foliar pools:

$$\text{EDC 5: } 0.2 f_{\text{roo}} < f_{\text{fol}} + f_{\text{lab}} < 5 f_{\text{roo}}, \tag{5}$$

$$\text{EDC 6: } 0.2\overline{C_{\text{fol}}} < \overline{C_{\text{roo}}} < 5\overline{C_{\text{fol}}}, \tag{6}$$

where $\overline{C_{\text{fol}}}$ and $\overline{C_{\text{roo}}}$ are the mean foliar and fine root carbon pool sizes over the model run period. EDC 5 ensures that the GPP allocated fraction to C_{roo} and C_{fol} (directly or via the labile C pool) are within a factor of 5 of each other. EDC 6 ensures that the mean fine root and foliar pool sizes are within a factor of 5 of each other.

2.2.3 Carbon Pool Growth

While we expect pools to potentially grow through time, we assume no recent disturbance and therefore limit the relative growth rate of pools. We constrain pool growth as follows:

$$\text{EDC 7: } \frac{\overline{C_{\text{pool}}^{\text{year}=n}}}{\overline{C_{\text{pool}}^{\text{year}=1}}} < 1 + G_{\text{max}} \frac{n-1}{10}, \tag{7}$$

where $\overline{C_{\text{pool}}^{\text{year}=1}}$ is the mean carbon pool size in year 1, and $\overline{C_{\text{pool}}^{\text{year}=n}}$ is the mean carbon pool size after $n-1$ years. We choose a value of $G_{\text{max}} = 0.1$, which is equivalent to a 10 % yearly growth rate (or doubling of carbon over 10 yr) as the maximum growth rate for each pool in EDC 7. This assumption is conservative, given data on global forest biomass growth rates (Baker et al., 2004; Luyssaert et al., 2008).

2.2.4 Carbon pool exponential decay trajectories

While carbon pools are expected to grow and contract through time, in the absence of major and recent disturbance events carbon pool trajectories are expected to exhibit gradual changes on inter-annual timescales (e.g. Bellamy et al., 2005). Under these circumstances, rapid exponential decay in modelled DALEC2 carbon pools can only occur as a result of an ecologically inconsistent $\boldsymbol{x}$. We examine the system response within a three-year period by imposing a constraint on exponential pool trajectories (Fig. 1): we numerically fit an exponential decay curve $a + be^{ct}$ to all carbon pools, where t is time in days, and a, b and c are the fitted exponential decay parameters.

DALEC2 pool trajectories are rejected if the half-life of carbon pool changes is less than three years~~(~~, i.e. ~~EDC 8: $c < -\frac{365.25 \times 3}{\log(2)}$; we~~

$$\text{EDC 8: } c < -\frac{365.25 \times 3}{\log(2)} \tag{8}$$

We fully describe the numerical derivation of c in Appendix B.

2.2.5 Steady State Proximity

For ecosystems with no recent disturbance events, we propose that each pool is within an order of magnitude of its steady state attractor. We use mean gross primary production ($\overline{F_{\text{gpp}}}$) as a proxy for

long-term GPP to estimate the steady state attractors, C^{∞}_{pool}, of four carbon pools (SOM, litter, wood and root). The steady state attractors for C_{som}, C_{lit}, C_{woo} and C_{roo} are analytically derived as follows:

$$C^{\infty}_{\text{som}} = \frac{(f_{\text{woo}} + (f_{\text{fol}} + f_{\text{roo}} + f_{\text{lab}})\theta_{\text{min}})\overline{F_{\text{gpp}}}}{(\theta_{\text{min}} + \theta_{\text{lit}})\theta_{\text{som}} e^{\Theta\overline{T}}}, \tag{9}$$

$$C^{\infty}_{\text{lit}} = \frac{(f_{\text{fol}} + f_{\text{roo}} + f_{\text{lab}})\overline{F_{\text{gpp}}}}{\theta_{\text{lit}} e^{\Theta\overline{T}}}, \tag{10}$$

$$C^{\infty}_{\text{woo}} = \frac{f_{\text{woo}}\overline{F_{\text{gpp}}}}{\theta_{\text{woo}}}, \tag{11}$$

$$C^{\infty}_{\text{roo}} = \frac{f_{\text{roo}}\overline{F_{\text{gpp}}}}{\theta_{\text{woo}}}, \tag{12}$$

where $\overline{T}$ is the mean annual temperature ($^{\circ}C$). For each pool, we impose an order-of-magnitude constraint on the proximity of C^{∞}_{pool} from the initial C_{pool} value:

$$\text{EDCs 9–12: } \frac{C^{0}_{\text{pool}}}{10} < C^{\infty}_{\text{pool}} < 10C^{0}_{\text{pool}} \tag{13}$$

where C^{0}_{pool} is the initial C_{som}, C_{lit}, C_{woo} and C_{roo} value for EDCs 9, 10, 11 and 12 respectively.

The twelve presented EDCs are what we believe to be the most ecologically suitable constraints on DALEC2 parameters and state variables, and are based on broader ecological knowledge of carbon dynamics. We discuss the advantages and the limitations of the proposed EDCs in Sect. 4 of this paper.

2.3 Model-Data Fusion

Given LAI observations, soil organic carbon estimates, prior parameter ranges (Table 1) and EDCs (Sect. 2.2), our aim for each experiment is to estimate the ~~likelihood~~ probability distribution of parameters $\boldsymbol{x}$. We assume no prior knowledge, other than the parameter ranges shown in Table 1: we therefore prescribe a uniform (i.e. non-informative) prior probability distribution onto all parameters. Within a Bayesian framework (e.g. Hill et al., 2012; Ziehn et al., 2012), we combine the above-mentioned information to derive the posterior probability density function of $\boldsymbol{x}$, $P(\boldsymbol{x}|\boldsymbol{O})$, where

$$P(\boldsymbol{x}|\boldsymbol{O}) \propto P(\boldsymbol{O}|\boldsymbol{x}) \cdot P_{\text{range}}(\boldsymbol{x}) \cdot P_{\text{EDC}}(\text{DALEC2}(\boldsymbol{x})) \tag{14}$$

$P(\boldsymbol{O}|\boldsymbol{x})$ is the probability of the observations given $\boldsymbol{x}$, $P_{\text{range}}(\boldsymbol{x}) = 1$ if all parameters are within the ranges prescribed in Table 1 (otherwise $P_{\text{range}}(\boldsymbol{x}) = 0$), and $P_{\text{EDC}}(\text{DALEC2}(\boldsymbol{x})) = 1$ if all EDCs are met (otherwise $P_{\text{EDC}}(\text{DALEC2}(\boldsymbol{x})) = 0$). For N observations, we derive the observation probability given $\boldsymbol{x}$, $P(\boldsymbol{O}|\boldsymbol{x})$, as follows:

$$P(\boldsymbol{O}|\boldsymbol{x}) = e^{-\frac{1}{2}\sum_{n=1}^{N}(M_n - O_n)^2/\sigma_n^2}, \tag{15}$$

where O_n is the nth observation, M_n is the corresponding state variable, and σ_n^2 is the nth error variance for each observation (e.g. Xu et al., 2006): here we assume no error covariance between observation errors.

We employ an adaptive Metropolis Hastings Markov Chain Monte Carlo (MHMCMC) approach to draw 5×10^6 samples from $P(\boldsymbol{x}|\boldsymbol{O})$. This approach has been widely used to estimate the probability density function of ecosystem model parameters (Xu et al., 2006; Hill et al., 2012; Ziehn et al., 2012; Caldararu et al., 2012; Smith et al., 2013; Keenan et al., 2013, amongst others) and is ideal to explore parameter space without a need to define normal prior distributions for each parameter (e.g. Richardson et al., 2010). We repeat the MHMCMC algorithm four times (i.e. four chains), to ensure convergence between $P(\boldsymbol{x}|\boldsymbol{O})$ distributions from each chain. To minimise sample correlations we use 500 $\boldsymbol{x}$ samples from the latter half of the accepted parameter vectors. We describe the details of our MHMCMC approach in Appendix C.

2.4 Synthetic truth – DALEC2 analyses

To quantify our ability to estimate synthetic DALEC2 ecosystem states, we perform the MDF approach over a three year period using LAI and SOM observations created from a synthetic DALEC2 truth ~~– our~~ , based on known DALEC2 parameters. Our choice of synthetic ~~data~~ DALEC2 states represents globally spanning datasets of satellite LAI retrievals and soil carbon map data. ~~We~~ Based on 40 DALEC2 parameter combinations, we create 40 synthetic ~~experiments~~ datasets representing typical temperate forest carbon dynamics, with three years of semi-continuous LAI data and one simulated soil organic carbon estimate. We use the three-year meteorology drivers (temperate climate) from the REFLEX synthetic experiments (Fox et al., 2009).

We ~~create~~ select 40 synthetic ~~experiments~~ parameter combinations by randomly sampling parameter vectors $\boldsymbol{x}$ within the DALEC2 parameter space (Table 1), where (i) $P_{\text{EDC}}(\text{DALEC2}(\boldsymbol{x})) = 1$, and (ii) $\boldsymbol{x}$ values are relevant to temperate forest ecosystems (see Appendix D). We remove approximately 95 % of daily LAI points to create an 8 day resolution semi-continuous LAI time-series. We add noise to the remaining 3 yr synthetic DALEC2 LAI: each LAI value is multiplied by a random error factor of $2^{N(0,1)}$, where $N(0,1)$ is a random number derived from a normal distribution with a mean of zero and a standard deviation of 1. For each synthetic soil carbon observation, we multiply C^0_{som} at $t = 0$ by a random error factor of $2^{N(0,1)}$. We fully explain the derivation of the synthetic experiment parameter vectors, (henceforth $\boldsymbol{s}$) in Appendix D.

We perform the MHMCMC and label the posterior parameter ensemble ($4 \times 500 \times 40$ $\boldsymbol{x}$ samples) as $\boldsymbol{x}\boldsymbol{s}_{\text{STA}}$ (standard synthetic MDF) and $\boldsymbol{x}\boldsymbol{s}_{\text{EDC}}$ (synthetic MDF with EDCs). We assign an uncertainty factor of 2 to all synthetic observations, hence O_n and M_n are log-transformed observations and $\sigma_n = \log(2)$. For each posterior DALEC2 $\boldsymbol{x}$, we determine the log-normalised parameter-space error $\epsilon(\boldsymbol{x})$ by comparing $\boldsymbol{x}$ with its corresponding synthetic truth vector $\boldsymbol{s}$:

$$\epsilon(\boldsymbol{x}) = \frac{\sqrt{\sum_{n=1}^{N} \left(\frac{\log(x(n)) - \log(s(n))}{\log(x(n)_{\max}) - \log(x(n)_{\min})} \right)^2}}{\sqrt{N}} \tag{16}$$

where $x(n)$ and $s(n)$ represent the nth parameters of $\boldsymbol{x}$ and $\boldsymbol{s}$, N is the number of parameters in $\boldsymbol{x}$,

and $x(n)_{\text{min}}$, $x(n)_{\text{max}}$ are the minimum and maximum parameter values (see Table 1). To assess the parameter estimation capability for each experiment, we derive the ϵ ($\boldsymbol{x}$) for each parameter vector in (a) $\boldsymbol{xs}_{\text{STA}}$ (b) $\boldsymbol{xs}_{\text{EDC}}$ and (c) for uniformly random samples where $P_{\text{range}}(\boldsymbol{x}) = 1$ (henceforth $\boldsymbol{xs}_{\text{RAN}}$). We refer to the ensemble of $\epsilon(\boldsymbol{x})$ values for $\boldsymbol{xs}_{\text{STA}}$, $\boldsymbol{xs}_{\text{EDC}}$ and $\boldsymbol{xs}_{\text{RAN}}$ as $\boldsymbol{E}(\boldsymbol{xs}_{\text{STA}})$, $\boldsymbol{E}(\boldsymbol{xs}_{\text{EDC}})$ and $\boldsymbol{E}(\boldsymbol{xs}_{\text{RAN}})$. We quantify the overall EDC associated error reduction (I_{EDC}) as follows:

$$I_{\text{EDC}} = \left(\frac{\tilde{\boldsymbol{E}}(\boldsymbol{xs}_{\text{RAN}}) - \tilde{\boldsymbol{E}}(\boldsymbol{xs}_{\text{EDC}})}{\tilde{\boldsymbol{E}}(\boldsymbol{xs}_{\text{RAN}}) - \tilde{\boldsymbol{E}}(\boldsymbol{xs}_{\text{STA}})} - 1 \right) \times 100\,\% \tag{17}$$

where ~~$\boldsymbol{E}$~~ $\tilde{\boldsymbol{E}}$ represents the median of $\boldsymbol{E}$ for each posterior parameter ensemble. This allows us to assess the relative improvement of $\boldsymbol{xs}_{\text{EDC}}$ over $\boldsymbol{xs}_{\text{STA}}$ parameter estimates against the $\boldsymbol{xs}_{\text{RAN}}$ "zero-knowledge" case. In addition, we determine the I_{EDC} for two parameter subgroups: (a) directly constrained parameters, and (b) indirectly constrained parameters. We assign c_{lf}, c_{ronset}, c_{rfall}, d_{onset}, d_{fall} and C^0_{som} to parameter group A: these parameters can be directly inferred from the LAI and soil organic carbon observations. We assign the remaining parameters to parameter group B: these can only be inferred from the DALEC2 model structure and – potentially – EDCs. Finally we compare NEE from DALEC2($\boldsymbol{xs}_{\text{EDC}}$) and DALEC2($\boldsymbol{xs}_{\text{STA}}$) against the NEE synthetic "truths" – DALEC2($\boldsymbol{s}$).

2.5 AmeriFlux – DALEC2 analyses

For the flux-tower experiments, we constrain DALEC2 parameters using (a) MODIS derived Leaf Area Index (LAI), and (b) total soil carbon from the harmonised world soil database (HWSD Hiederer and Köchy, 2011). We perform daily resolution three year DALEC2 analyses for three forest categories: evergreen needle-leaf (ENF), deciduous broad-leaf (DBF), and mixed forest (MF). We chose one AmeriFlux site from each forest type. To establish a suitable site for our method we chose sites with NEE data spanning across three years between 2001 and 2010. ~~We narrowed our selection by choosing~~ Our selected sites for each forest type are Howland Forest (US-Ho1, evergreen needleleaf forest – 45.2041° N, 68.7402° W – Hollinger et al., 1999) , Morgan Monroe State Forest (US-MMS, deciduous broadleaf forest – 39.3231° N, 86.4131° W – Schmid et al., 2000) and Sylvania Wilderness (US-Syv, mixed forest – 46.2420° N, 89.3476° W – Desai et al., 2005) . We chose temperate sites with ~~less than 2~~ little expected water-stress, and with a ≤ 3 months of recorded ~~freezing temperaturesand little expected water-stress: these~~ below-freezing soil temperatures. These criteria reflect the current capabilities of DALEC2, as hydrological processes are not explicitly portrayed in the model. ~~Our selected sites for each forest type are Howland Forest (US-Ho1, evergreen needleleaf forest – 45.2041° N, 68.7402° W – Hollinger et al., 1999) , Morgan Monroe State Forest (US-MMS, deciduous broadleaf forest – 39.3231° N, 86.4131° W – Schmid et al., 2000) and Sylvania Wilderness (US-Syv, mixed forest – 46.2420° N, 89.3476° W – Desai et al., 2005) .~~

For each AmeriFlux site, we extract the corresponding MODIS LAI retrievals from the MOD15A2 LAI 8 day version 005 1 km resolution product (downloaded from the Land Processes Distributed

Active Archive Centre http://lpdaac.usgs.gov/): we only keep maximum quality flag data. Standard deviations are provided for 1 km MODIS LAI retrievals, however these (a) do not reflect the magnitude variability in uncertainty, (b) often imply the existence of negative LAI observations ($\sigma_{\mathrm{LAI}} > \mathrm{LAI}$) and (c) are occasionally missing. While various MODIS LAI evaluations have been performed (e.g. Sea et al., 2011; Serbin et al., 2013), large-scale spatiotemporal LAI retrieval errors remain poorly quantified. For the sake of simplicity, we assign a factor of 2 uncertainty (i.e. $\log(\mathrm{LAI}) \pm \log(2)$) for each MODIS LAI observation. To minimise spatial discrepancies between MODIS and AmeriFlux sites, each LAI observation is the arithmetic mean of all available LAI retrievals within a 9 pixel $3\,\mathrm{km} \times 3\,\mathrm{km}$ area (centred on each AmeriFlux site). Overall, we use 95, 120 and 119 LAI values at US-Syv, US-Ho1 and US-MMS (5th–95th percentile ranges for LAI values are 0.4-5.8, 1.0-5.6 and 0.4-5.5 respectively).

For each site we extract total soil carbon density from the nearest Harmonised World Soil Database 30 arc seconds resolution total soil carbon content (approx. 1 km at equator Hiederer and Köchy, 2011): the authors have performed multiple comparisons of the global HWSD against other products, however no pixel-scale uncertainties are provided. We chose to assign an uncertainty factor of 2 on each site-scale HWSD SOC estimate. The HWSD SOC values are 2.3×10^4 $\mathrm{g\,C\,m^{-2}}$, 2.3×10^4 $\mathrm{g\,C\,m^{-2}}$ and 5.2×10^3 $\mathrm{g\,C\,m^{-2}}$.

To limit our study to the use of globally spanning datasets, we extract DALEC2 drivers from $0.125\,^\circ \times 0.125\,^\circ$ ERA interim meteorology (see Appendix A for details). The DALEC2 analyses for each site are therefore completely independent from all site-level measurements (we note, however, that extensive meteorological and biometric data are meticulously recorded across the AmeriFlux site network). Therefore, we produce a fully independent ecosystem carbon cycle analysis, which can be evaluated against measured NEE at each flux-tower site.

As done for the synthetic experiments, we perform the MHMCMC approach at each site – with and without EDCs – and label the posterior parameter ensembles (4 chains $\times$ 500 $\boldsymbol{x}$ samples) as $\boldsymbol{xa}_{\mathrm{STA}}$ (standard AmeriFlux MDF) and $\boldsymbol{xa}_{\mathrm{EDC}}$ (AmeriFlux MDF + EDCs). We compare the DALEC2 NEE analyses, DALEC2($\boldsymbol{xa}_{\mathrm{EDC}}$) and DALEC2($\boldsymbol{xa}_{\mathrm{STA}}$) against NEE measurements at each AmeriFlux site.

2.6 EDC sensitivity test

To determine the sensitivity of our results to EDCs 1-12, we repeat MDF estimates of $\boldsymbol{xs}_{\mathrm{EDC}}$ and $\boldsymbol{xa}_{\mathrm{EDC}}$ by imposing only one EDC at a time (henceforth $\boldsymbol{xs}_{\mathrm{EDC}(n)}$ and $\boldsymbol{xa}_{\mathrm{EDC}(n)}$, where n is the nth EDC). For the synthetic experiments, we determine the relative contribution of the nth EDC by quantifying the overall EDC associated error reduction ($I_{\mathrm{EDC}(n)}$, see equation 17) for each estimate of $\boldsymbol{xs}_{\mathrm{EDC}(n)}$. Given the large computational cost of estimating $\boldsymbol{xs}_{\mathrm{EDC}(n)}$ for each EDC (40 synthetic experiments $\times$ 12 EDCs $\times$ 4 chains), we limit our sensitivity analysis to I_{EDC} estimates based on 4 (out of 40) synthetic experiments.

We compare 3 yr integrated DALEC2 NEE estimates and AmeriFlux NEE measurements at all three sites (AmeriFlux NEE measurement temporal gaps have been consistently excluded from DALEC2 3 yr NEE estimates). We determine the DALEC2 3 yr NEE 50% confidence range (50% CR: 25th–75th percentile interval) reduction as follows:

$$(1 - \frac{R_{\text{NEE,EDC(n)}}}{R_{\text{NEE,STA}}}) \times 100\% \tag{18}$$

where $R_{\text{NEE,EDC(n)}}$ and $R_{\text{NEE,STA}}$ are the 50% CR of DALEC2($\boldsymbol{xa}_{\text{EDC(n)}}$) and DALEC2($\boldsymbol{xa}_{\text{STA}}$) 3 yr NEE estimates. Similarly, we calculate the 3 yr NEE bias reduction (relative to AmeriFlux NEE measurements) as follows:

$$(1 - \frac{|B_{\text{NEE,EDC(n)}}|}{|B_{\text{NEE,STA}}|}) \times 100\% \tag{19}$$

where $B_{\text{NEE,EDC(n)}}$ and $B_{\text{NEE,STA}}$ are the median biases of DALEC2($\boldsymbol{xa}_{\text{EDC(n)}}$) and DALEC2($\boldsymbol{xa}_{\text{STA}}$) 3 yr NEE estimates.

3 Results

3.1 Synthetic Experiments

The inclusion of EDCs resulted in substantial error reductions in posterior DALEC2 parameter and state variable estimates. We found an overall reduction in the posterior MHMCMC EDC parameter vector errors $\boldsymbol{E}(\boldsymbol{xs}_{\text{EDC}})$, relative to both the standard MHMCMC errors $\boldsymbol{E}(\boldsymbol{xs}_{\text{STA}})$ and the randomly sampled parameter vector errors $\boldsymbol{E}(\boldsymbol{xs}_{\text{RAN}})$: we found an improvement of $I_{\text{EDC}} = 34\,\%$ associated with using EDCs (Fig. 2c). For the directly constrained parameters (parameter group A) we found similar ~~likelihood functions~~ distributions for both $\boldsymbol{E}(\boldsymbol{xs}_{\text{STA}})$ and $\boldsymbol{E}(\boldsymbol{xs}_{\text{EDC}})$ errors relative to $\boldsymbol{E}(\boldsymbol{xs}_{\text{RAN}})$ errors (Fig. 2a), and similarly lower $\boldsymbol{xs}_{\text{STA}}$ and $\boldsymbol{xs}_{\text{EDC}}$ errors values relative to $\boldsymbol{xs}_{\text{RAN}}$ errors ($\tilde{\boldsymbol{E}}(\boldsymbol{xs}_{\text{STA}}) = 0.19$, $\tilde{\boldsymbol{E}}(\boldsymbol{xs}_{\text{EDC}}) = 0.21$, $\tilde{\boldsymbol{E}}(\boldsymbol{xs}_{\text{RAN}}) = 0.42$, group A: $I_{\text{EDC}} = -6\,\%$). For the indirectly constrained parameters (group B), we found significantly smaller $\boldsymbol{xs}_{\text{EDC}}$ errors relative to $\boldsymbol{xs}_{\text{STA}}$ and $\boldsymbol{xs}_{\text{RAN}}$ ($\tilde{\boldsymbol{E}}(\boldsymbol{xs}_{\text{EDC}}) = 0.29$, $\tilde{\boldsymbol{E}}(\boldsymbol{xs}_{\text{STA}}) = 0.34$, and $\tilde{\boldsymbol{E}}(\boldsymbol{xs}_{\text{RAN}})$ = 0.38), and hence improved estimates of $\boldsymbol{s}$ when we implemented EDCs (group B: $I_{\text{EDC}} = 88\,\%$, Fig. 2b). We found that EDCs 5 and 8 accounted for the largest error reduction in DALEC2 parameter estimates ($I_{\text{EDC(5,8)}} \geq 3\,\%$, Table 2), followed by EDCs 6, 10 and 12 ($I_{\text{EDC(6,10,12)}} = 2\,\%$). EDC 7 led to an overall parameter error increase ($I_{\text{EDC(7)}} = -13\,\%$). The remaining EDCs accounted for small or negative error reductions.

We compared EDC total $\boldsymbol{xs}_{\text{EDC}}$, $\boldsymbol{xs}_{\text{STA}}$ and $\boldsymbol{xs}_{\text{RAN}}$ live biomass ($C_{\text{roo}} + C_{\text{fol}} + C_{\text{lab}} + C_{\text{woo}}$) and dead biomass ($C_{\text{som}} + C_{\text{lit}}$) pool biases relative to their corresponding synthetic truths (Fig. 2d–e).

For dead biomass, both $\boldsymbol{xs}_{\mathrm{EDC}}$ and $\boldsymbol{xs}_{\mathrm{STA}}$ perform comparably better than $\boldsymbol{xs}_{\mathrm{RAN}}$ (Fig. 2e), as dead biomass is mostly accounted for by the synthetic C_{som} observations: the $\boldsymbol{xs}_{\mathrm{EDC}}$ and $\boldsymbol{xs}_{\mathrm{STA}}$ median bias factors (1.1, 0.91) are close to 1 (i.e. a bias of zero) relative to $\boldsymbol{xs}_{\mathrm{RAN}}$ median bias factor (0.04). For live biomass pools, $\boldsymbol{xs}_{\mathrm{EDC}}$ live biomass bias estimates are smaller than $\boldsymbol{xs}_{\mathrm{STA}}$ (Fig. 2d): the $\boldsymbol{xs}_{\mathrm{EDC}}$ bias distribution (median = 1.20) is closer to 1 relative to the $\boldsymbol{xs}_{\mathrm{STA}}$ bias distribution (0.48), with respect to $\boldsymbol{xs}_{\mathrm{RAN}}$ median bias (0.20). For total biomass estimates, we found similar bias distributions relative to $\boldsymbol{xs}_{\mathrm{RAN}}$: ($\boldsymbol{xs}_{\mathrm{EDC}}$ median bias factor = 1.22, $\boldsymbol{xs}_{\mathrm{STA}}$ bias factor = 0.98): both bias factors are closer to 1 relative to $\boldsymbol{xs}_{\mathrm{RAN}}$ (bias factor = 0.16).

We found that incorporating EDCs resulted in a reduced mode and 90 % confidence range (90% CR: 95th–5th percentile interval) for three year NEE biases (Fig. 3). We found a 65 % reduction in the DALEC2($\boldsymbol{xs}_{\mathrm{EDC}}$) three year NEE bias 90 % CR ($9.0\,\mathrm{g\,C\,m^{-2}\,d^{-1}}$), relative to the DALEC2($\boldsymbol{xs}_{\mathrm{STA}}$) three year NEE bias 90 % CR ($26.9\,\mathrm{g\,C\,m^{-2}\,d^{-1}}$). The three year NEE bias modes for DALEC2($\boldsymbol{xs}_{\mathrm{EDC}}$) and DALEC2($\boldsymbol{xs}_{\mathrm{STA}}$) are $0.0\,\mathrm{g\,C\,m^{-2}\,d^{-1}}$ and -0.5 $\mathrm{g\,C\,m^{-2}\,d^{-1}}$ (at $0.5\,\mathrm{g\,C\,m^{-2}\,d^{-1}}$ intervals).

3.2 AmeriFlux results

The DALEC2($\boldsymbol{xa}_{\mathrm{EDC}}$) analyses outperformed the standard DALEC2($\boldsymbol{xa}_{\mathrm{STA}}$) analyses at the AmeriFlux tower sites. The inclusion of EDCs in DALEC2 analyses amounted to overall NEE bias reductions at all sites (US-Syv, US-Ho1, US-MMS, we henceforth present all site results in this order). The aggregated DALEC2($\boldsymbol{xa}_{\mathrm{EDC}}$) median daily NEE biases (-0.02, 0.13, $-0.03\,\mathrm{g\,C\,m^{-2}\,d^{-1}}$) are closer to the AmeriFlux measured NEE by roughly one order of magnitude in contrast to DALEC2($\boldsymbol{xa}_{\mathrm{STA}}$) median NEE biases ($-0.52$, -0.86, $-1.15\,\mathrm{g\,C\,m^{-2}\,d^{-1}}$). The aggregated daily DALEC2($\boldsymbol{xa}_{\mathrm{EDC}}$) NEE 90 % confidence ranges at each site (10.9, 10.1, $8.3\,\mathrm{g\,C\,m^{-2}\,d^{-1}}$) were all smaller (53–87 %) than the corresponding DALEC2($\boldsymbol{xa}_{\mathrm{STA}}$) NEE bias 90 % CR (20.3, 18.3, $9.5\,\mathrm{g\,C\,m^{-2}\,d^{-1}}$). The reductions in bias are consistent across the three year comparison period at each site (Fig. 4).

Cumulative AmeriFlux NEE observations are compared against corresponding DALEC2($\boldsymbol{xa}_{\mathrm{STA}}$) and DALEC2($\boldsymbol{xa}_{\mathrm{EDC}}$) NEE estimates (Fig. 5); AmeriFlux NEE temporal gaps have been omitted from both DALEC2 and AmeriFlux derived cumulative NEE time series. DALEC2($\boldsymbol{xa}_{\mathrm{EDC}}$) integrated NEE estimates outperformed DALEC2($\boldsymbol{xa}_{\mathrm{STA}}$) NEE estimates at all three sites. DALEC2($\boldsymbol{xa}_{\mathrm{EDC}}$) median NEE biases over the 3 yr period (-0.26, 0.07, $0.08\,\mathrm{kg\,C\,m^{-2}}$) are smaller than the equivalent DALEC2($\boldsymbol{xa}_{\mathrm{STA}}$) biases ($-0.84$, -1.09, $-1.18\,\mathrm{kg\,C\,m^{-2}}$), with relative EDC bias reductions of 69 %, 93 % and 93 %. The inclusion of EDCs also resulted in a reduction in NEE confidence intervals: DALEC2($\boldsymbol{xa}_{\mathrm{EDC}}$) 50 % CR (1.17, 1.57, 1.16 $\mathrm{kg\,C\,m^{-2}}$) are 32–48 % smaller than the corresponding DALEC2($\boldsymbol{xa}_{\mathrm{STA}}$) 50 % CR (2.04, 3.00, 1.70 $\mathrm{kg\,C\,m^{-2}}$). Based on DALEC2($\boldsymbol{xa}_{\mathrm{EDC}(u)}$) 3 yr NEE estimates, EDC 10 resulted in a $\geq 18\%$ bias reduction and a $\geq 5\%$ 50 % CR reduction at all three sites, relative to DALEC2($\boldsymbol{xa}_{\mathrm{STA}}$) (Table 2). EDCs 2 and 8 resulted in

a $> 10\%$ 3 yr NEE 50 % CR reduction and an increase in 3 yr NEE bias at all three sites (NEE bias reduction $\leq -22\%$). EDCs 7 and 9 resulted in a $\geq 50\%$ 3 yr NEE bias reduction and an increase in 3 yr NEE 50 % CR at all three sites (NEE 50 % CR reduction $\leq -15\%$).

4 Discussion

With the use of a simple model and globally available data, i.e. leaf area dynamics and soil carbon observations, we have demonstrated that the EDC approach provides an improved ability to infer the magnitude of carbon fluxes, live carbon pools and model parameters, in comparison to a standard parameter optimisation approach (STA).

For ecologically relevant synthetic truths, EDCs provide improved estimates of the DALEC2 parameters and state variables. The EDC approach resulted in (a) parameter estimation error reductions, (b) NEE bias and confidence range reductions, and (c) improved estimates of the live biomass C pools, in contrast to the STA parameter and flux and C pool estimates. While there is little difference between directly inferable (Group A) estimated parameter errors between the EDC and STA approach, using EDCs led to a marked reduction in estimated parameter error for indirectly inferable (Group B) parameters. The indirectly inferred parameters include allocation fractions, subsurface pools and turnover rates, which are typically difficult to observe at field sites and virtually impossible to observe remotely (i.e. at regional scales).

By comparing DALEC2 analyses against independent AmeriFlux NEE measurements over real ecosystems, we further validated the advantages of using EDCs. At each AmeriFlux site, we found that EDCs led to an increased confidence and a largely reduced NEE bias; our DALEC2 model analyses suggests that the use of EDCs regionally and globally could significantly enhance our ability to estimate ecosystem state variables in the absence of direct observational constraints. In light of the large differences between earth system models (Todd-Brown et al., 2013; Friend et al., 2013), we anticipate that EDCs may help constrain ecosystem carbon terms on global scales, where carbon pools and their residence times are typically difficult or impossible to measure.

Together, EDCs 1-12 lead to overall improvements in parameter estimates and AmeriFlux site NEE confidence range/bias (Table 2): however, with the exception of EDC 10, when EDCs were tested individually, they did not lead to comprehensive improvements. For example, EDC 8 alone (no rapid exponential pool decay) resulted in large AmeriFlux site NEE confidence range reductions, as well as improved synthetic parameter estimates; however, EDC 8 resulted in higher AmeriFlux site NEE biases. Conversely, EDC 9 (steady state proximity of the soil carbon pool) resulted in the largest AmeriFlux site bias reductions, while NEE confidence was lower. EDC 5 (comparable fine root and foliar/labile allocation) led to the largest parameter improvements; however, the associated changes in AmeriFlux site NEE estimates were relatively small. Our findings demonstrate that robust improvements in carbon cycling parameter and state variable estimates only arise when EDCs are

used collectively.

Here we developed a group of EDCs suitable to ecosystems with no recent major disturbance. However, we note that our EDCs can be adapted for a wider range of ecosystem dynamics. For example, recently disturbed ecosystems may be (a) rapidly recovering and (b) growing towards a steady state where carbon pools are greater than one order of magnitude from the initial carbon pools. Therefore a subset of our EDCs (EDCs 7–12) can be adapted to better represent ecological "common sense" in recovering ecosystems.

Ultimately, EDCs can be adapted to best represent ecological knowledge in a variety of ecosystem carbon model MDF applications, where the ecosystem observations are insufficient to constrain all model state variables (e.g. Fox et al., 2009). For example, on regional and global spatial scales, there is often no explicit knowledge on various model parameter values and their associated uncertainty. In such cases, our EDC approach imposes inter-parameter constraints while simultaneously allowing a global parameter exploration across several orders of magnitude (see Table 1). Hence EDCs allow us to incorporate ecologically consistent relationships between parameters (i.e. allocation ratios, turnover ratios), without the need to constrain otherwise unknown parameter and state variables. Moreover, as an alternative to imposing plant-functional-type priors, which risk being subjective and over-rigid, ecosystem trait inter-relationships derived from plant trait data (e.g. Wright et al., 2004; Kattge et al., 2011) could be incorporated as additional EDCs. Given a quantitative knowledge of parameter inter-relationships, we also note that a prior parameter variance-covariance structure – in addition to EDCs – can also be used as an alternative or complementary constraint on the model state and parameters. Finally, we note that our choice of EDCs is open to adaptation and adjustment: we maintained relatively broad constraints (e.g. EDC 6 permissible root:foliar C range $>$ one order of magnitude), which can likely be refined through further study.

In this study we limited our observational constraints to globally spanning MODIS LAI retrievals and the HWSD soil map. Given these two datasets, we have demonstrated that EDCs lead to improved model parameter estimates and reduced NEE bias and confidence ranges. Nonetheless, based on the posterior NEE probability density function, we are unable to determine whether sites are net carbon sinks or sources on annual timescales. However, an increasing number of continental and global scale biospheric datasets are becoming available: these include a global canopy height map by Simard et al. (2011), pan-tropical biomass maps by Saatchi et al. (2011); Baccini et al. (2012) and a pan-boreal carbon density map by Thurner et al. (2013). These products can potentially be used in conjuncture with MODIS LAI, HWSD data and our EDC approach in a MDF framework to better constrain terrestrial carbon cycle dynamics.

5 Concluding Remarks

We have addressed the under-determined nature of the carbon cycle problem by applying a group of widely applicable ecological and dynamic constraints (EDCs) on an ecosystem carbon model in a model-data fusion (MDF) framework. Particularly where extensive in-situ measurements are not available, EDCs can be used to incorporate ecological knowledge, such as parameter inter-relationships and pool dynamics constraints, into ecosystem carbon model analyses. In a synthetic data experiment, we found improved estimates of DALEC2 model parameters, live carbon pools and net ecosystem exchange (NEE) when using EDCs in DALEC2 MDF analyses. By validating our DALEC2 MDF analyses against independent AmeriFlux NEE measurements, we found that EDCs led to a 69–93 % reduction in three year NEE biases. We incorporated twelve EDCs in DALEC2 analyses of temperate forest ecosystem carbon cycling: these EDCs can potentially be adapted for a range of models and biomes. Moreover, additional EDCs can be derived to incorporate parameter inter-relationships derived from regional or global plant trait datasets into ecosystem carbon model analyses. Here we have shown that EDCs can be used to constrain the poorly resolved components of the carbon cycle: we therefore advocate the use of EDCs in future MDF analyses of the terrestrial carbon cycle.

Appendix A

DALEC2 model

The full DALEC2 model dynamics can be expressed as six equations:

$$C_{\text{lab}}^{t+1} = (1 - \Phi_{\text{onset}}(t, d_{\text{onset}}, c_{\text{ronset}}))C_{\text{lab}}^{t} + f_{\text{lab}}F_{\text{gpp}}^{t} \tag{A1}$$

$$C_{\text{fol}}^{t+1} = (1 - \Phi_{\text{fall}}(t, d_{\text{fall}}, c_{\text{rfall}}, c_{\text{lspan}}))C_{\text{fol}}^{t} + \Phi_{\text{onset}}(t, d_{\text{onset}}, c_{\text{ronset}})C_{\text{lab}}^{t} + f_{\text{fol}}F_{\text{gpp}}^{t} \tag{A2}$$

$$C_{\text{roo}}^{t+1} = (1 - \theta_{\text{roo}})C_{\text{roo}}^{t} + f_{\text{roo}}F_{\text{gpp}}^{t} \tag{A3}$$

$$C_{\text{woo}}^{t+1} = (1 - \theta_{\text{woo}})C_{\text{woo}}^{t} + f_{\text{woo}}F_{\text{gpp}}^{t} \tag{A4}$$

$$C_{\text{lit}}^{t+1} = (1 - (\theta_{\text{lit}} + \theta_{\text{min}})e^{\Theta T_t})C_{\text{lit}}^{t} + \theta_{\text{roo}}C_{\text{roo}}^{t} + \Phi_{\text{fall}}(t, d_{\text{fall}}, c_{\text{rfall}}, c_{\text{lspan}})C_{\text{fol}}^{t} \tag{A5}$$

$$C_{\text{som}}^{t+1} = (1 - \theta_{\text{som}}e^{\Theta T_t})C_{\text{som}}^{t} + \theta_{\text{woo}}C_{\text{woo}}^{t} + (\theta_{\text{min}})e^{\Theta T_t}C_{\text{lit}}^{t} \tag{A6}$$

The 23 free parameters and carbon pool symbols are summarised in Table 1. The daily gross primary production F_{gpp}^{t}, is derived from the aggregated canopy model (ACM, Williams et al., 1997), and is a function of daily driver data $\boldsymbol{M}$ (day of year, atmospheric CO_2, minimum and maximum temperature, and global radiation), and parameters C_{fol}, c_{lma} and c_{eff} (parameter c_{eff}, the canopy efficiency, is a replacement for the nitrogen $\times$ nitrogen use efficiency product in ACM).

The model is initiated with six initial carbon pool values (C^0_{lab}, C^0_{fol}, C^0_{roo}, C^0_{woo}, C^0_{lit}, C^0_{som}) and these are iteratively updated at a daily timestep. The leaf onset (labile to foliar pool C transfer) and leaf fall (foliar to litter pool C transfer) functions, Φ_{onset} and Φ_{fall} are defined below:

$$\Phi_{\text{onset}}(t, d_{\text{onset}}, c_{\text{ronset}}) = \frac{\sqrt{2}}{\sqrt{\pi}} \cdot \left(\frac{6.9088}{c_{\text{ronset}}}\right) \cdot e^{-\left(\sin\left(\frac{t - d_{\text{onset}} - 0.6245 c_{\text{ronset}}}{s}\right) \cdot \frac{\sqrt{2}s}{c_{\text{ronset}}}\right)^2} \tag{A7}$$

$$\Phi_{\text{fall}}(t, d_{\text{fall}}, c_{\text{rfall}}, c_{\text{lspan}}) = \frac{\sqrt{2}}{\sqrt{\pi}} \cdot \left(\frac{\log(c_{\text{lspan}}) - \log(c_{\text{lspan}} - 1)}{c_{\text{rfall}}}\right) \cdot e^{-\left(\sin\left(\frac{t - c_{\text{rfall}} + \psi_f}{s}\right) \cdot \frac{\sqrt{2}s}{c_{\text{rfall}}}\right)^2} \tag{A8}$$

The Φ_{fall} continuous cyclical step function derivatives were derived such that (a) $\prod_{t=1}^{t=365}(1 - \Phi_{\text{onset}}(t, d_{\text{onset}}, c_{\text{ronset}})) = 1 - \frac{1}{c_{\text{lf}}}$ (b) the maximum leaf loss rate occurs annually at $t\,\%\,365.25 = d_{\text{fall}}$. (c) 68 % of leaf loss occurs within c_{ronset} days and 95 % of leaf loss within $2c_{\text{rfall}}$ days.

ψ_f is an offset term included to ensure that the maximum leaf loss rate, .i.e. $\frac{\mathrm{d}^2 C_{\text{fol}}}{\mathrm{d}t^2} = 0$, occurs at $t\,\%\,365.25 = d_{\text{fall}}$ – it is a numerical solution to the following equation:

$$2\sqrt{\pi} \cdot \log\left(\frac{c_{\text{lspan}}}{c_{\text{lspan}} - 1}\right) \cdot \psi + e^{-\psi^2} = 0. \tag{A9}$$

where $\psi = \frac{\sqrt{2}\psi_f}{c_{\text{rfall}}}$ (we note that Eq. (A9) can be solved using a Lambert W function, where $\psi = W(f(c_{\text{lf}}))$ - however, Lambert W functions cannot be solved analytically). We created a look up function for ψ by fitting a 6th order polynomial between ψ and $log(c_{\text{lf}} - 1)$ – the full polynomial is included in the downloadable DALEC2 code. Φ_{onset} is a special case of the Φ_{fall} formula: it was derived such that 99.99 % of C_{lab} is transferred to C_{fol} annually at $t\%365.25 = d_{\text{onset}}$ and 68 % of leaf onset occurs within c_{ronset} day. The Φ functions are advantageous in that (a) the daily turnover rates result in a continuous and specified loss of carbon throughout a known time period, and (b) the functions are cyclical and hence do not need to be reset, "switched on" or "switched off" throughout the model run period. We also note that while we treated d_{onset} and d_{fall} as constant parameters, the Φ functions can easily accommodate temporally variable definitions for leaf onset and leaf fall. Total ecosystem respiration F^t_{rec} and the net ecosystem exchange F^t_{nee} fluxes are derived at each timestep and are shown below.

$$F^t_{\text{rec}} = f_{\text{auto}} F^t_{\text{gpp}} + (\theta_{\text{lit}} C^t_{\text{lit}} + \theta_{\text{som}} C^t_{\text{som}}) e^{\Theta T_t} \tag{A10}$$

$$F^t_{\text{nee}} = F^t_{\text{rec}} - F^t_{\text{gpp}} \tag{A11}$$

At time t leaf area index (LAI) is defined as:

$$\text{LAI}^t = \frac{C^t_{\text{fol}}}{c_{\text{lma}}} \tag{A12}$$

A schematic of the of the carbon fluxes in DALEC2 is shown in Fig. 6.

For ~~AMERIFLUX~~ AmeriFlux DALEC2 analyses we used daily meteorological drivers for DALEC2 from 0.125 ° × 0.125 ° ERA-interim re-analyses. For each site we obtained coordinates from ameriflux.ornl.gov. We downloaded 6 h temperature and 12 h downward surface solar radiation data for all site locations and years from apps.ecmwf.int/datasets. We averaged temperature and

radiation from the four nearest $0.125\,^\circ \times 0.125\,^\circ$ ERA-interim grid-points. We obtained minimum and maximum temperatures from the 6 h ERA-interim temperature range. For M daily radiation values we used the sum of the two 12 h radiation re-analyses.

Appendix B

Exponential C_{pool} decay

Exponentially decaying C^t_{pool} trajecories can be approximated as $C_{\text{exp}} = a + be^{ct}$, where a, b and c are constants, t is time in days. To implement EDC 8, we numerically estimate parameter c: to derive c for each carbon pool trajectory (C_{pool}) we derive (i) the gradient between yearly means for years 1 and 2:

$$\Delta C_0 = \frac{[\sum_{t=365+1}^{365\times 2} + C^t_{\text{pool}} - \sum_{t=1}^{365} C^t_{\text{pool}}]}{365} \tag{B1}$$

and (ii) the gradient between the yearly means with a 1 day offset:

$$\Delta C_1 = \frac{[\sum_{t=365+1+1}^{365\times 2+1} C^t_{\text{pool}} - \sum_{t=1+1}^{365+1} C^t_{\text{pool}}]}{365} \tag{B2}$$

Parameter c can be expressed as:

$$c = \log(\frac{\Delta C_1}{\Delta C_0}) \tag{B3}$$

In the case of a true exponential curve with a known value of c, the numeric derivation of c shown in equations B1–B3 is exact (i.e. within numerical precision of the true c). In cases where there is no exponential decay, c is either positive or complex. While C_{exp} is an approximation of C^t_{pool}, in practice this approach is both computationally fast and effectively able to identify rapid exponential decay ($c < \frac{\log(2)}{365.25\times 3}$) trajectories.

Appendix C

Adaptive MHMCMC algorithm

For the standard MDF parameter estimates the normalised parameter probability is:

$$P(\boldsymbol{x}|\boldsymbol{O}) = \alpha \cdot P(\boldsymbol{O}|\boldsymbol{x}) \cdot P_{\text{range}}(\boldsymbol{x}), \tag{C1}$$

and for EDC MDF parameter estimates the normalised parameter probability is:

$$P(\boldsymbol{x}|\boldsymbol{O}) = \alpha \cdot P(\boldsymbol{O}|\boldsymbol{x}) \cdot P_{\text{range}}(\boldsymbol{x}) \cdot P_{\text{EDC}}(\text{DALEC2}(\boldsymbol{x})), \tag{C2}$$

where α is a scaling constant ensuring $\int(P(\boldsymbol{x}|\boldsymbol{O}))\mathrm{d}\boldsymbol{x} = 1$. For each chain, we search for a random $\boldsymbol{x}_0$ starting point where $P_{\text{range}}(\boldsymbol{x}_0) \cdot P_{\text{EDC}}(\text{DALEC2}(\boldsymbol{x}_0)) = 1$ (for standard runs we randomly sample $\boldsymbol{x}_0$ from $P_{\text{range}}(\boldsymbol{x}_0) = 1$).

Based on the Ziehn et al. (2012) algorithm, we then iterate through the following steps:

1. $\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \boldsymbol{d}$.
2. Run DALEC2($\boldsymbol{x}_{i+1}$).
3. If $\frac{P(\boldsymbol{x}_{i+1}|\boldsymbol{O})}{P(\boldsymbol{x}_i|\boldsymbol{O})} > U(0,1)$, accept $\boldsymbol{x}_{i+1}$, and $i = i + 1$.

where $\boldsymbol{d}$ is the stepsize. The ratio $\frac{P(\boldsymbol{x}_{i+1}|\boldsymbol{O})}{P(\boldsymbol{x}_i|\boldsymbol{O})}$ is derived from equations C1 and C2 for standard and EDC MHMCMC iterations, respectively (therefore knowledge of α is not required). At each iteration, for each parameter dimension n, $\boldsymbol{d}(n) = \boldsymbol{s}(n)N(0,1)$, where $\boldsymbol{s}$ is the proposal distribution and $N(0,1)$ is a random number sampled from a normal distribution with mean = 0 and variance = 1. This sequence repeated until 10^7 samples of $\boldsymbol{x}$ have been accepted. While any proposal distribution $\boldsymbol{s}$ can be used, adapting the proposal distribution can reduce the number of steps required to reach the maximum probability parameter space. For the first 5×10^6 samples, we adapt the proposal distribution $\boldsymbol{s}$ every 100 iterations by (i) scaling $\boldsymbol{s}$ to ensure an acceptance rate of 23–44 % (Ziehn et al., 2012), and (ii) scale individual dimensions of $\boldsymbol{s}$ to ensure that $2\boldsymbol{s}_n > \sigma_{\boldsymbol{x}}(n)$ where $\sigma_{\boldsymbol{x}}(n)$ is the nth parameter standard deviation over 100 iterations. The $P(\boldsymbol{x}|\boldsymbol{O})$ distribution is then derived from the second 5×10^6 samples.

The MHMCMC parameter sampling approach is then repeated four times (four chains): to determine whether all four chains have converged to the same parameter distributions, we use the Gelman-Rubin convergence criterion R, where for each parameter $R < 1.1$ indicates an acceptable chain convergence (Gelman and Rubin, 1992; Xu et al., 2006). If the chains have not converged for all parameters, we sequentially test all N chain combinations (where $N \geq 2$) to (a) repeat the GR criterion, and (b) determine the combination with the maximum number of converged chains.

Appendix D

Temperate forest synthetic truths

The 40 synthetic experiments were created by searching for parameter vectors $\boldsymbol{s}$ where $P_{\text{EDC}}(\text{DALEC2}(\boldsymbol{s})) = 1$. To create synthetic experiments parameter vectors $\boldsymbol{s}$ relevant to temperate forest ecosystems - henceforth $P_{\text{TF}}(\text{DALEC2}(\boldsymbol{s})) = 1$ - we imposed the following parameter and state variable conditions:

1. $60 < d_{\text{onset}} < 150$
2. $242 < d_{\text{fall}} < 332$

3. $c_{\text{ronset}} > 20$

4. $c_{\text{rfall}} > 30$

5. $c_{\text{lf}} > 0.25$

6. $1 < \overline{\text{LAI}^t} < 8$

7. $3\text{kg}\,\text{C}\,\text{m}^{-2} < C^0_{\text{woo}} < 30\text{kg}\,\text{C}\,\text{m}^{-2}$

8. $1\text{kg}\,\text{C}\,\text{m}^{-2} < C^0_{\text{som}} < 100\text{kg}\,\text{C}\,\text{m}^{-2}$

9. $\overline{F^t_{\text{gpp}}} > 2\,\text{g}\,\text{C}\,\text{m}^{-2}\,\text{d}^{-1}$

10. $\frac{5}{6} < \frac{\overline{\text{LAI}_{\text{year}=3}}}{\overline{\text{LAI}_{\text{year}=1}}} < \frac{6}{5}$

For a given vector $\boldsymbol{s}$, all conditions must be met when $P_{\text{TF}}(\text{DALEC2}(\boldsymbol{s})) = 1$. The above-listed conditions ensure that the selected $\boldsymbol{s}$ vectors broadly reflect canopy dynamics (1–4), carbon pool sizes (5–6) mean photosynthetic uptake (7) and limited year-to-year canopy changes (8) associated with temperate forest ecosystems (e.g. Fox et al., 2009). We derive each parameter vector $\boldsymbol{s}$ by selecting a random parameter vector $\boldsymbol{s}_0$ and incrementally adjusting it until $P_{\text{EDC}}(\text{DALEC2}(\boldsymbol{s}))$ $P_{\text{TF}}(\text{DALEC2}(\boldsymbol{s})) = 1$. To represent a range of canopy dynamics, we also imposed either (a) a deciduous condition (where $c_{\text{rfall}} > \frac{10}{11}$) or (b) a mixed forest or evergreen condition (where $c_{\text{rfall}} < \frac{10}{11}$), with a 50 %-50 % propability for either constraint.

We simplistically simulate the 8-daily MODIS LAI data and soil carbon map HWSD products from DALEC2($\boldsymbol{s}$) LAI and C_{som}: we multiplied each soil organic carbon "truth" at $t = 0$ (C^0_{som}) by $2^{N(0,1)}$, where $N(0,1)$ is a random number sampled from normally distribution with mean = 0 and variance = 1.

For LAI synthetic observations, we only kept one in eight LAI values, and created correlated gaps in the remaining LAI data of random lengths until at least 50 % of the ~~8daily~~ 8 daily data is removed. Overall, between 65 and 68 LAI observations are kept for each 3 yr synthetic experiment. Twenty-two parameter vectors are categorised as deciduous, and eighteen as evergreen. Mean 3 yr F^t_{gpp} ranges from 2.04–8.79 $\text{g}\,\text{C}\,\text{m}^{-2}\,\text{d}^{-1}$ (median $= 4.75\text{g}\,\text{C}\,\text{m}^{-2}\,\text{d}^{-1}$) and mean 3 yr F^t_{nee} ranges from -3.71 to $2.87\text{g}\,\text{C}\,\text{m}^{-2}\,\text{d}^{-1}$ (median $= -0.72\,\text{g}\,\text{C}\,\text{m}^{-2}\,\text{d}^{-1}$).

Acknowledgements. This project was funded by the NERC National Centre for Earth Observation, U.K. This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF, http://www.ecdf.ed.ac.uk/). The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 283080, project GEOCARBON. This work used eddy covariance data acquired by the FLUXNET community and in particular by the AmeriFlux (US Department of Energy, Biological and Environmental Research, Terrestrial Carbon Program (DEFG0204ER63917 and DEFG0204ER63911)). We acknowledge the financial support to the eddy covariance data harmonization

provided by CarboEuropeIP, FAOGTOSTCO, iLEAPS, Max Planck Institute for Biogeochemistry, National Science Foundation, University of Tuscia, Université Laval and Environment Canada and US Department of Energy and the database development and technical support from Berkeley Water Center, Lawrence Berkeley National Laboratory, Microsoft Research eScience, Oak Ridge National Laboratory, University of California Berkeley, University of Virginia. AmeriFlux is funded by the the United States Department of Energy (DOE – TES), Department of Commerce (DOC – NOAA), the Department of Agriculture (USDA – Forest Service), the National Aeronautics and Space Administration (NASA), the National Science Foundation (NSF). US-Syv received funding support from Department of Energy, Ameriflux Network Management Project Support for UW ChEAS Cluster and the Department of Energy, Terrestrial Carbon Processes. The writing of this paper was partially carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. We are grateful for feedback from L. Smallman and discussions with J. Exbrayat and T. Hill.

References

Akaike, H.: New look at statistical-model identification, IEEE Transactions on Automatic Control, Ac., 19, 716–723., 1974.

Baccini, A., Goetz,S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D.,Hackler, J., Beck, P. S. A., Dubayah, R., Friedl, M. A., Samanta, S., and Houghton, R. A.: Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps, Nature Climate Change, 2, 182–185, 2012.

Baker, T. R., Phillips, O. L., Malhi, Y., Almeida, S., Arroyo,L., Di Fiore, A., Erwin, T., Higuchi, N., Killeen, T. J., Laurance, S. G. , Laurance, W. F., Lewis, S. L., Monteagudo, A., Neill, D. A., Núñez Vargas, P., Pitman, N. C. A. , Silva, J. N. M., and Vásquez Martínez, R.: Increasing biomass in amazonian forest plots, P. T. Roy. Soc. Lond., 359, 353–365, 2004.

Baldocchi, D., Falge, E. , Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S. Fluxnet: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, B. Am. Meteorol. Soc., 82, 2415–2434, 2001.

Bellamy, P. H., Loveland, P. J., Bradley, R. I., Lark, R. M., and Kirk, G. J.: Carbon losses from all soils across england and wales 1978–2003, Nature, 437, 245–248, 2005.

Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the glue methodology, J. Hydrol., 249, 11–29, 2001.

Caldararu, S., Palmer, P. I., and Purves, D. W.: Inferring Amazon leaf demography from satellite observations of leaf area index, Biogeosciences, 9, 1389–1404, doi:10.5194/bg-9-1389-2012, 2012.

Carvalhais, N., Reichstein, M., Ciais, P., Collatz, G. J., Mahecha, M. D., Montagnani, L., Papale, D., Rambal, S., and Seixas, J.: Identification of vegetation and soil carbon pools out of equilibrium in a process model via eddy covariance and biometric constraints, Glob. Change. Biol., 16, 2813–2829,2010.

Desai, A. R., Bolstad, P. V., Cook, B. D., Davis, K. J., and Carey, E. V.: Comparing net ecosystem exchange of carbon dioxide between an old-growth and mature forest in the upper midwest, usa, Agr. Forest. Meteorol., 128, 33–55, 2005.

Feng, L., Palmer, P. I., Yang, Y., Yantosca, R. M., Kawa, S. R., Paris, J.-D., Matsueda, H., and Machida, T.: Evaluating a 3-D transport model of atmospheric CO_2 using ground-based, aircraft, and space-borne data, Atmos. Chem. Phys., 11, 2789–2803, doi:10.5194/acp-11-2789-2011, 2011.

Fox, A., Williams, M., Richardson, A. D., Cameron, D., Gove, J. H., Quaife, T., Ricciuto, D., Re- ichstein, M., Tomelleri, E., Trudinger, C. M., and van Wijk, M. T.: The reflex project: comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data, Agr. Forest. Meteorol., 149, 1597–1615, 2009.

Friend, A. D., Lucht, W., Rademacher, T. T., Keribin, R., Betts, R., Cadule, P., Ciais, P., Clark, D. B., Dankers, R., Falloon, P. D., Ito, A., Kahana, R., Kleidon, A., Lomas, M. R., Nishina, K., Ostberg, S., Pavlick, R., Peylin, P., Schaphoff, S., Vuichard, N., Warszawski, L., Wiltshire, A., and Woodward, F. I.: Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO2, Proc. Natl. Acad. Sci., 111, 3280–3285, 201222477, 2013.

Gaudinski, J. B., Trumbore, S. E., Davidson, E. A., and Zheng, S.: Soil carbon cycling in a temperate forest:

radiocarbon-based estimates of residence times, sequestration rates and partitioning of fluxes, Biogeochemistry, 51, 33–69, 2000.

Gelman, A. and Rubin, D. B.: Inference from iterative simulation using multiple sequences, Statistical science, 7, 457–472, 1992.

Hiederer, R. and M. Köchy: Global soil organic carbon estimates and the harmonized world soil database, EUR, 79, 25225, doi:10.2788/13267, 2011.

Hill, T. C., Ryan, E., and Williams, M.: The use of CO_2 flux time series for parameter and carbon stock estimation in carbon cycle research, Glob. Change Biol., 18, 179–193, 2012.

Hollinger, D., Goltz, S., Davidson, E., Lee, J., Tu, K., and Valentine, H.: Seasonal patterns and environmental control of carbon dioxide and water vapour exchange in an ecotonal boreal forest, Glob. Change Biol., 5, 891–902, 1999.

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

Keenan, T. F., Davidson, E. A., Munger, J. W., and Richardson, A. D.: Rate my data: quantifying the value of ecological data for the development of models of the terrestrial carbon cycle, Ecol. Appl., 23, 273–286, 2013.

Le Quéré, C., Andres, R. J., Boden, T., Conway, T., Houghton, R. A., House, J. I., Marland, G., Peters, G. P., van der Werf, G., Ahlström, A., Andrew, R. M., Bopp, L., Canadell, J. G., Ciais, P., Doney, S. C., Enright, C., Friedlingstein, P., Huntingford, C., Jain, A. K., Jourdain, C., Kato, E., Keeling, R. F., Klein Goldewijk, K., Levis, S., Levy, P., Lomas, M., Poulter, B., Raupach, M. R., Schwinger, J., Sitch, S., Stocker, B. D., Viovy, N., Zaehle, S., and Zeng, N.: The global carbon budget 1959–2011, Earth Syst. Sci. Data, 5, 165–185, doi:10.5194/essd-5-165-2013, 2013.

Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., Clark, J. S., and Schimel, D. S.: Ecological forecasting and data assimilation in a data-rich era, Ecol. Appl., 21, 1429–1442, 2011.

Luo, Y. and Weng, E.: Dynamic disequilibrium of the terrestrial carbon cycle under global change, Trends Ecol. Evol., 26, 96–104, 2011.

Luyssaert, S., Schulze, E.-D., Börner, A., Knohl, A., Hessenmöller, D., Law, B. E., Ciais, P., and Grace, J.: Old-growth forests as global carbon sinks, Nature, 455, 213–215, 2008.

Mokany, K., Raison, R., and Prokushkin, A. S.: Critical analysis of root: shoot ratios in terrestrial biomes, Glob. Change Biol., 12, 84–96, 2006.

Norby, R. J., Hanson, P. J., O'Neill, E. G., Tschaplinski, T. J., Weltzin, J. F., Hansen, R. A., Cheng, W., Wullschleger, S. D., Gunderson, C. A., Edwards, N. T., and Johnson, D. W.: Net primary productivity of a CO_2-enriched deciduous forest and the implications for carbon storage, Ecol. Appl., 12, 1261–1266, 2002.

Peters, W., Krol, M. C., Van Der Werf, G. R., Houweling, S., Jones, C. D., Hughes, J., Schaefer, K., Masarie, K. A., Jacobson, A. R., Miller, J. B., Cho, C. H., Ramonet, M., Schmidt, M., Ciattaglia, L., Apadula, F., Heltai, D., Meinhardt, F., Di Sarra, A. G., Piacentino, S., Sferlazzo, D., Aalto, T., Hatakka, J., Ström, J., Haszpra, L., Meijer, H. A. J., Van Der Laan, S., Neubert, R. E. M., Jordan, A., Rodó, X., Morguí, J.-A., Vermeulen, A. T., Popa, E., Rozanski, K., Zimnoch, M., Manning, A. C., Leuenberger, M., Uglietti, C., Dolman, A. J., Ciais, P., Heimann, M., and Tans, P. P. Seven years of recent european net terrestrial carbon dioxide exchange constrained by atmospheric observations, Glob. Change Biol., 16, 1317–1337, 2010.

Quaife, T., Lewis, P., De Kauwe, M., Williams, M., Law, B. E., Disney, M., and Bowyer, P.: Assimilating canopy reflectance data into an ecosystem model with an ensemble kalman filter, Remote Sens. Environ., 112, 1347–1364, 2008.

Richardson, A. D., Williams, M., Hollinger, D. Y., Moore, D. J., Dail, D. B., Davidson, E. A., Scott, N. A., Evans, R. S., Hughes, H., Lee, J. T., Rodrigues, C., and Savage, K.: Estimating parameters of a forest ecosystem c model with measurements of stocks and fluxes as joint constraints, Oecologia, 164, 25–40, 2010.

Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S., Petrova, S., White, L., Silman, M., and Morel, A.: Benchmark map of forest carbon stocks in tropical regions across three continents, P. Natl. Acad. Sci., 108, 9899–9904, 2011.

Schmid, H. P., Grimmond, C., S. B., Cropley, F., Offerle, B., and Su, H.-B.: Measurements of CO2 and energy fluxes over a mixed hardwood forest in the mid-western united states, Agr. Forest. Meteorol., 103, 357–374, 2000.

Schwalm, C. R., Williams, C. A., Schaefer, K., Anderson, R., Arain, M. A., Baker, I., Barr, A., Black, T. A., Chen, G., Chen, J. M., Ciais, P., Davis, K. J., Desai, A., Dietze, M., Dragoni, D., Fischer, M. L., Flanagan, L. B., Grant, R., Gu, L. Hollinger, D. Izaurralde, R. C., Kucharik, C. Lafleur, P., Law, B. E., Li, L., Li, Z. Liu, S., Lokupitiya, E., Luo, Y., Ma, S., Margolis, H., Matamala, R., McCaughey, H., Monson, R. K., Oechel, W. C., Peng, C. Poulter, B. Price, D. T. Riciutto, D. M., Riley, W., Sahoo, A. K., Sprintsin, M., Sun, J., Tian, H., Tonitto, C., Verbeeck, H., and Verma, S. B.: A model-data intercomparison of CO2 exchange across north america: Results from the north american carbon program site synthesis, J. Geophys. Res.-Biogeo., 115, G00H05, doi:10.1029/2009JG001229, 2010.

Sea, W. B., Choler, P., Beringer, J., Weinmann, R. A., Hutley, L. B., and Leuning, R.: Documenting improvement in leaf area index estimates from modis using hemispherical photos for australian savannas, Agr. Forest. Meteorol., 151, 1453–1461, 2011.

Serbin, S. P., Ahl, D. E., and Gower, S. T.: Spatial and temporal validation of the modis lai and fpar products across a boreal forest wildfire chronosequence, Remote Sens. Environ., 133, 71–84, 2013.

Simard, M., Pinto, N., Fisher, J. B., and Baccini, A.: Mapping forest canopy height globally with spaceborne lidar, J. Geophys. Res.-Biogeo., 116, G04021, doi:10.1029/2011JG001708, 2011.

Sitch, S., Huntingford, C., Gedney, N., Levy, P., Lomas, M., Piao, S., Betts, R., Ciais, P., Cox, P., Friedlingstein, P., Jones, C. D., Prentice, I. C., and Woodward, F. I.: Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five dynamic global vegetation models (dgvms), Glob. Change Biol., 14, 2015–2039, 2008.

Sloan, V. L., Fletcher, B. J., Press, M. C., Williams, M., and Phoenix, G. K.: Leaf and fine root carbon stocks and turnover are coupled across arctic ecosystems, Glob. Change Biol., 19, 3668–3676, 2013.

Smith, M. J., Purves, D., Vanderwel, M., Lyutsarev, V., and Emmott, S.: The climate dependence of the terrestrial carbon cycle, including parameter and structural uncertainties, Biogeosciences, 10, 583–606, doi:10.5194/bg-10-583-2013, 2013.

Thurner, M., Beer, C., Santoro, M., Carvalhais, N., Wutzler, T., Schepaschenko, D., Shvi-denko, A., Kompter, E., Ahrens, B., Levick, S. R., and Schmullius, C.: Carbon stock and density of northern boreal and temperate forests, Global Ecol. Biogeogr., 23, 297–310, 2013.

Todd-Brown, K., Randerson, J., Post, W., Hoffman, F., Tarnocai, C., Schuur, E., and Allison, S.: Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations, Biogeosciences, 10, 1717–1736, doi:10.5194/bg-10-1717-2013, 2013.

Trumbore, S.: Carbon respired by terrestrial ecosystems–recent progress and challenges, Glob. Change Biol., 12, 141–153, 2006.

Williams, M. and Rastetter, E. B.: Vegetation characteristics and primary productivity along an arctic transect: implications for scaling-up, J. Ecol., 87, 885–898, 1999.

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, Ecol. Appl., 7, 882–894, 1997.

Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., Carvalhais, N., Jung, M., Hollinger, D. Y., Kattge, J., Leuning, R., Luo, Y., Tomelleri, E., Trudinger, C. M., and Wang, Y.-P.: Improving land surface models with FLUXNET data, Biogeosciences, 6, 1341–1359, doi:10.5194/bg-6-1341-2009, 2009.

Williams, M., Schwarz, P. A., Law, B. E., Irvine, J., and Kurpius, M. R.: An improved analysis of forest carbon dynamics using data assimilation, Glob. Change Biol., 11, 89–105, 2005.

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., Cavender- Bares, J., Chapin, T., Cornelissen, J. H., Diemer, M., Flexas, J., Garnier, E., Groom, P. K., Gulias, J., Hikosaka, K., Lamont, B. B., Lee,T., Lee, W., Lusk, C., Midgley, J. J., Navas, M.-L., Niinemets, Ü., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V. I., Roumet, C., Thomas, S. C., Tjoelker, M. G., Veneklaas, E. J., and Villar, R.: The worldwide leaf economics spectrum, Nature, 428, 821–827, 2004.

Xu, T., White, L., Hui, D., and Luo, Y.: Probabilistic inversion of a terrestrial ecosystem model: Analysis of uncertainty in parameter estimation and model prediction, Global. Biogeochem. Cy., 20, GB2007, doi:10.1029/2005GB002468, 2006.

Ziehn, T., Scholze, M., and Knorr, W.: On the capability of monte carlo and adjoint inversion techniques

Table 1. DALEC2 model parameters, descriptions, and minimum – maximum parameter values: the corresponding DALEC2 equations are fully described in Appendix A.

Parameter	Description	Range
f_{auto}	autotrophic respiration fraction	0.3–0.7
f_{lab}	fraction of GPP allocated to labile C pool	0.01–0.5
f_{fol}	fraction of GPP allocated to foliage	0.01–0.5
f_{roo}	fraction of GPP allocated to fine roots	0.01–0.5
f_{woo} [1]	fraction of GPP allocated to wood	0.01–0.5
θ_{woo}	Woody C turnover rate	2.5×10^{-5}–$10^{-3}\,\text{d}^{-1}$
θ_{roo}	Fine root C turnover rate	10^{-4}–$10^{-2}\,\text{d}^{-1}$
θ_{lit}	Litter C turnover rate	10^{-4}–$10^{-2}\,\text{d}^{-1}$
θ_{som}	Soil organic C turnover rate	10^{-7}–$10^{-3}\,\text{d}^{-1}$
θ_{min}	Litter mineralisation rate	10^{-2}–$10^{-5}\,\text{d}^{-1}$
Θ	temperature dependence exponent factor	0.018–0.08
d_{onset}	Leaf Onset Day	1–365
d_{fall}	Leaf Fall Day	1–365
c_{eff}	Canopy Efficiency Parameter	10–100
c_{lma}	Leaf Mass per area	10–400 g C m^{-2}
c_{lf}	Annual Leaf Loss Fraction	$\frac{1}{8}$–1
c_{ronset}	Labile C release period	10–100 day
c_{rfall}	Leaf-fall period	~~10~~20–~~100~~150 day
C^t_{lab}	Labile C pool at time t	~~10~~20–~~1000 g C m^{-2}~~2000 g C m^{-2}
C^t_{fol}	Foliar C pool at time t	~~10~~20–~~1000 g C m^{-2}~~2000 g C m^{-2}
C^t_{roo}	Fine root C pool at time t	~~10~~20–~~1000 g C m^{-2}~~2000 g C m^{-2}
C^t_{woo}	Above & Below ground woody C pool at time t	100–10^5 g C m^{-2}
C^t_{lit}	Litter C pool at time t	~~10~~20–~~1000 g C m^{-2}~~2000 g C m^{-2}
C^t_{som}	Soil organic C pool at time t	100–2×10^5 g C m^{-2}

[1] f_{woo} is equivalent to $1 - f_{\text{auto}} - f_{\text{fol}} - f_{\text{lab}}$.

to derive posterior parameter uncertainties in terrestrial ecosystem models, Global. Biogeochem. Cy., 26, GB3025, doi:10.1029/2011GB004185, 2012.

Table 2. Synthetic experiment parameter error reduction, and Ameriflux experiment 3 yr NEE 50% CR and bias reduction for MDF estimates using individual EDCs, relative to the standard MDF estimates.

EDC	Synthetic experiment parameter error reduction ([1] $I_{EDC(n)}$)	AmeriFlux experiments [2]NEE 50% CR reduction (bias reduction)		
		US-Syv	US-Ho1	US-MMS
1	−0%	27% (−11%)	19% (−13%)	3% (−11%)
2	−1%	39% (−26%)	29% (−25%)	14% (−19%)
3	0%	13% (−0%)	1% (3%)	0% (−7%)
4	−1%	30% (−14%)	22% (−14%)	9% (−17%)
5	8%	3% (−3%)	0% (−4%)	1% (−11%)
6	2%	10% (3%)	−2% (6%)	−1% (−3%)
7	−13%	−15% (52%)	−28% (76%)	−25% (95%)
8	3%	34% (−36%)	37% (−9%)	16% (−66%)
9	1%	−39% (89%)	−50% (57%)	−31% (100%)
10	2%	10% (19%)	6% (25%)	5% (18%)
11	−1%	10% (−0%)	1% (11%)	3% (1%)
12	2%	8% (−1%)	2% (0%)	3% (−6%)
ALL EDCs	**34%**	**43%(69%)**	**48%(93%)**	**32%(93%)**

[1]The parameter error reduction metric, $I_{EDC(n)}$, is described in section 2.4. [2]The derivations of 3 yr NEE 50% CR and bias reductions are described in section 2.6

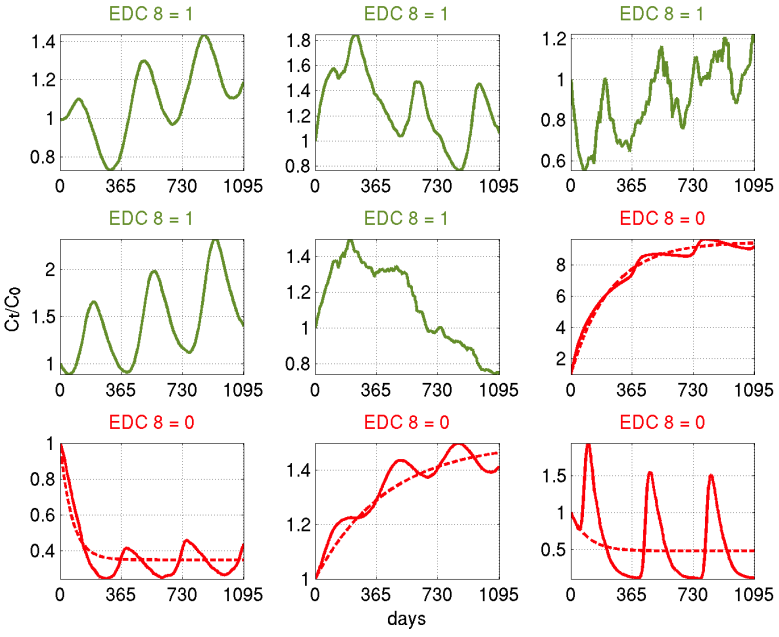

Fig. 1. Exponential decay test (EDC 8) performed on nine example normalised C_{pool} trajectories over a 3 yr time-span. The C_{pool} trajectories are normalised such that $C_{\text{pool}} = 1$ at $t = 0$. Examples 1–5 were accepted (EDC 8 $= 1$) and examples 6–9 were rejected (EDC 8 $= 0$). The exponential decay fit (dashed line) is shown for pool trajectories where EDC 8 $= 0$.

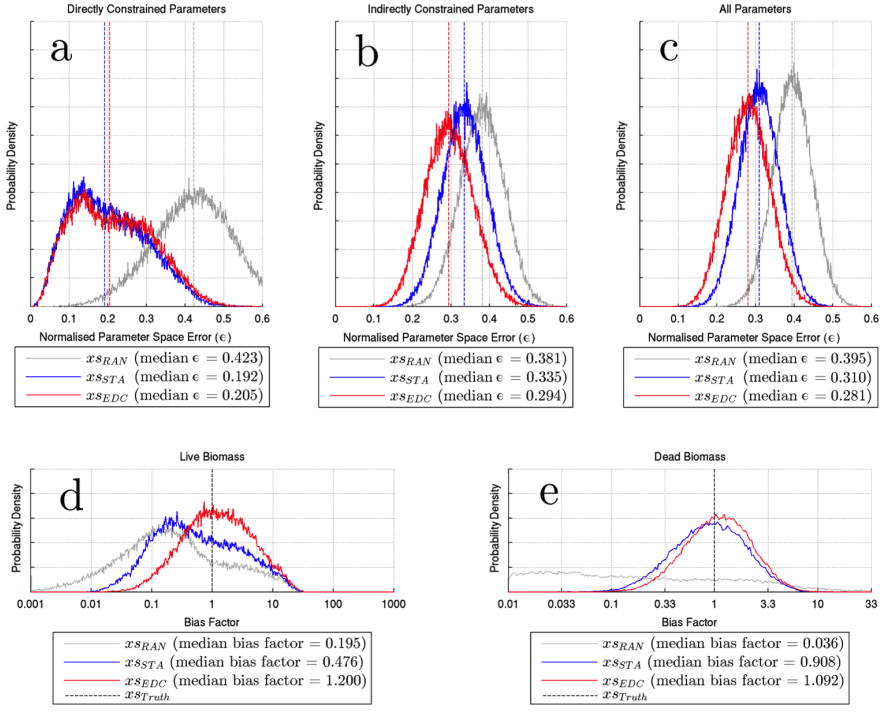

Fig. 2. Aggregated parameter estimates $\boldsymbol{xs}_{\mathrm{STA}}$ (standard sampling – blue) and $\boldsymbol{xs}_{\mathrm{EDC}}$ (EDC sampling - red) from deciduous and evergreen synthetic LAI and soil organic carbon observations – these are compared against observation and EDC independent parameter samples $\boldsymbol{xs}_{\mathrm{RAN}}$ (light grey). Panels (**a**–**c**) Normalised parameter space error (ϵ) ~~normalised likelihood~~ probability density functions for (a) Group A (directly inferable) parameters, (b) Group B (indirectly inferable) parameters, and (c) all DALEC2 parameters. ϵ values for each parameter group were ~~estimated~~ derived using Eq. (15). In panels (d) and (e) the ~~likelihood~~ probability density functions of live carbon stock (foliar labile wood and roots) and dead carbon stock (litter and soil carbon) biases against the synthetic truth parameters $\boldsymbol{s}$ are shown for $\boldsymbol{xs}_{\mathrm{RAN}}$, $\boldsymbol{xs}_{\mathrm{STA}}$ and $\boldsymbol{xs}_{\mathrm{EDC}}$ parameter estimates.

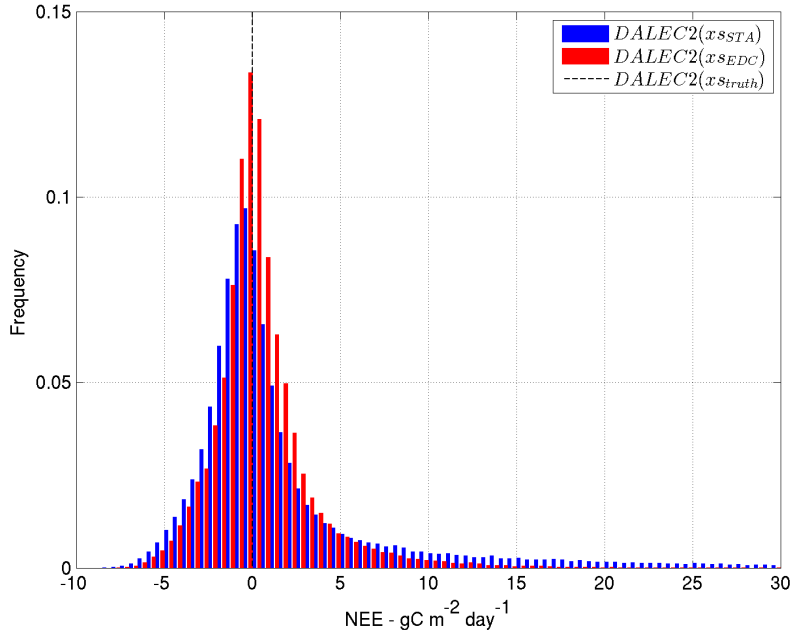

Fig. 3. Three year mean DALEC2 net ecosystem exchange (NEE) biases (relative to synthetic truth) aggregated across 40 synthetic experiments at $0.5\,\mathrm{g\,C\,m^{-2}\,d^{-1}}$ intervals. The bias ~~likelihoods~~ frequencies are shown for DALEC2($\boldsymbol{xs}_{\mathrm{STA}}$) (standard sampling – blue) and DALEC2($\boldsymbol{xs}_{\mathrm{EDC}}$) (EDC sampling – red) relative to the synthetic truth DALEC2($\boldsymbol{s}$) (black dashed line).

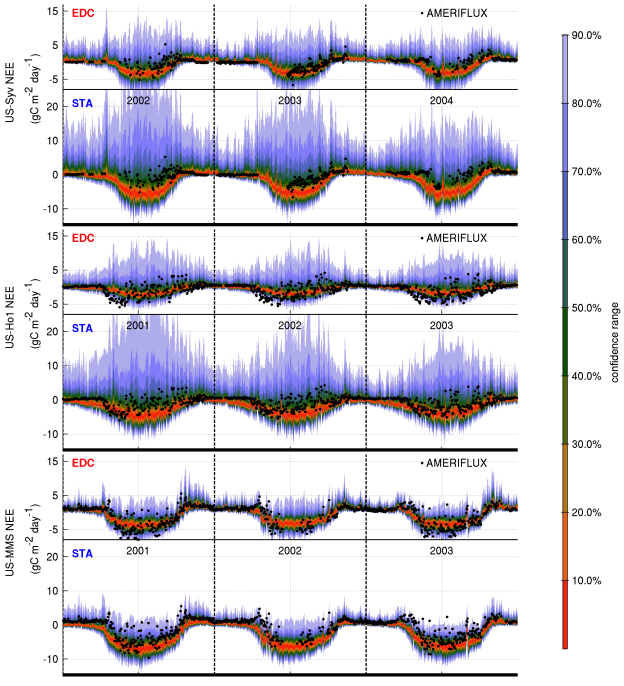

Fig. 4. DALEC2 daily NEE ensemble estimates at three AmeriFlux sites: Sylvania Wilderness (US-Syv, mixed forest, top two rows), Howland Forest (US-Ho1, evergreen needleleaf, middle two rows), and Morgan Monroe State Forest (US-MMS, deciduous broadleaf, bottom two rows). For each site the DALEC2($\boldsymbol{xa}_{\mathrm{EDC}}$) and the DALEC2($\boldsymbol{xa}_{\mathrm{STA}}$) ensemble confidence intervals are denoted as EDC and STA, respectively. The DALEC2 analyses - based on MODIS LAI retrievals, HWSD soil organic carbon estimates and ERA interim meteorological drivers - are completely independent from all AmeriFlux site measurements.

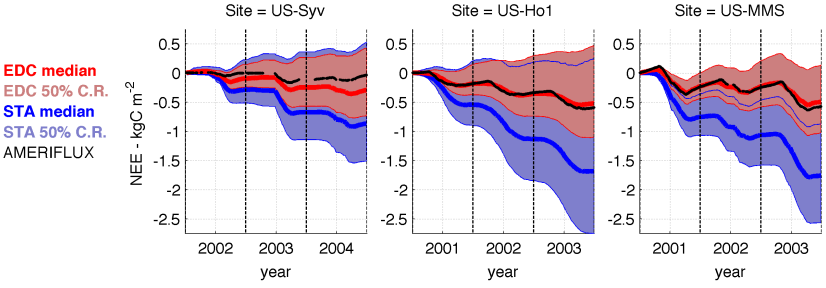

Fig. 5. Three year mean DALEC2 cumulative NEE ($\mathrm{kg\,C\,m^{-2}}$) compared against cumulative measured NEE at three AmeriFlux sites:Sylvania Wilderness (US-Syv, mixed forest, left), Howland Forest (US-Ho1, evergreen neddleleaf, middle), and Morgan Monroe State Forest (US-MMS, deciduous broadleaf, right). The standard analysis median and 50 % confidence ranges (CR) are shown in blue, and the corresponding analyses with EDCs are shown in red. AmeriFlux NEE measurements are denoted as a black line. The DALEC2 analyses - based on MODIS LAI retrievals, HWSD soil organic carbon estimates and ERA interim meteorological drivers – are completely independent from all AmeriFlux site measurements.

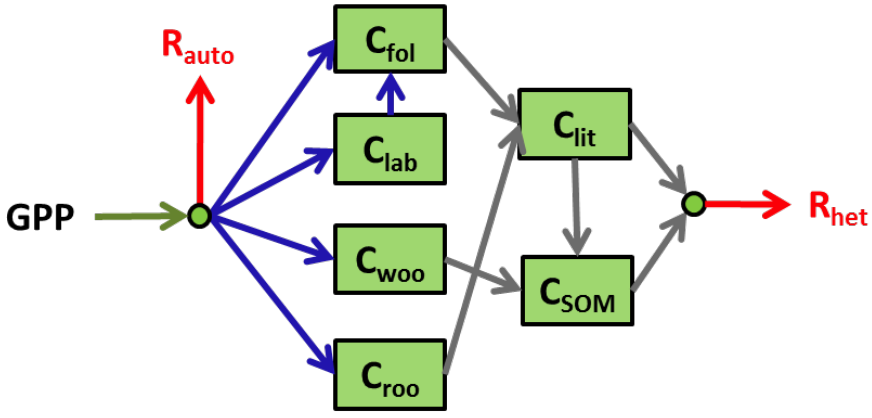

Fig. 6. Schematic of the carbon fluxes in DALEC2. The green arrow indicates the gross primary production (GPP). Red arrows represent respiration fluxes: autotrophic respiration (R_{auto}) and heterotrophic respiration (R_{het}). Blue arrows represent C allocation to the labile (C_{lab}), foliar (C_{fol}), wood (C_{woo}) and fine root (C_{roo}) pools. Grey arrows represent the litterfall and decomposition fluxes to the litter (C_{lit}) and soil organic matter (C_{som}) pools.