https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Biogeosciences
Discussions
Open Access

Resolving scale-variance in the carbon dynamics of fragmented, mixed-use landscapes estimated using Model-Data Fusion

David T. Milodowski[1,2], T. Luke Smallman[1,2], and Mathew Williams[1,2]

[1]School of GeoSciences, University of Edinburgh, Edinburgh UK
[2]National Centre for Earth Observations, University of Edinburgh, UK

Correspondence: D. T. Milodowski (d.t.milodowski@ed.ac.uk)

Abstract. Many terrestrial landscapes are heterogeneous. Mixed land cover and land-use generate a complex mosaic of fragmented ecosystems at fine spatial resolutions with contrasting ecosystem stocks, traits and processes, each differently sensitive to environmental and human factors. Representing spatial complexity within terrestrial ecosystem models is a key challenge for understanding regional carbon dynamics, their sensitivity to environmental gradients, and their resilience in the face of climate change. Heterogeneity underpins this challenge due to the trade-off between the fidelity of ecosystem representation within modelling frameworks and the computational capacity required for fine-scale model calibration and simulation. We directly address this challenge by quantifying the sensitivity of simulated carbon fluxes in a mixed-use landscape in the UK to the spatial resolution of the model analysis. We test two different approaches for combining EO data into the CARDAMOM Model-Data Fusion (MDF) framework, assimilating time series of satellite-based Earth Observation (EO) derived estimates of ecosystem leaf area and biomass stocks to constrain estimates of model parameters and their uncertainty for an intermediate complexity model of the terrestrial C cycle. In the first approach, ecosystems are calibrated and simulated at pixel-level, representing a "community average" of the encompassed land cover and management. This represents our baseline approach. In the second, we stratify each pixel based on land-cover (e.g. coniferous forest, arable/pasture etc.), and calibrate the model independently using EO data specific to each stratum. We test the scale-dependence of these approaches for grid resolutions spanning 1° to 0.05° over a mixed land-use region of the UK. Our analyses indicate that spatial resolution matters for MDF. Under the "community-average" baseline approach biological C fluxes (GPP, R_{eco}) simulated by CARDAMOM are insensitive to resolution. However, disturbance fluxes exhibit scale-variance that increases with greater landscape fragmentation, and for coarser model domains. In contrast, stratification of assimilated data based on fine-resolution land-use distributions resolved the resolution dependence, leading to disturbance fluxes that were approximately double the baseline experiments. The differences in simulated disturbance fluxes were sufficient to drive alternative interpretations of the terrestrial C balance: in the baseline experiment the live C pools suggest a strong C sink, whereas in the stratified experiment, the live C pools were approximately in steady-state as the C gains from NPP were balanced by losses due to the higher simulated harvest fluxes focused in conifer woodlands. We also find that stratifying the model domain based on land-use leads to differences in the retrieved parameters that reflect variations in ecosystem function between neighbouring areas of contrasting land-use. The emergent differences in model parameters between land-use strata give rise to divergent responses to future climate change. Accounting for fine-scale structure in heterogeneous landscapes (e.g. stratification) is therefore vital for ensuring the ecological fidelity of

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

large-scale MDF frameworks. The need for stratification arises because land-use places strong controls on the spatial distribution of carbon stocks and plant functional traits, and on the ecological processes controlling the fluxes of C through landscapes, particularly those related to management and disturbance. Given the importance of disturbance to global terrestrial C fluxes, together with the widespread increase in fragmentation of forest landscapes, these results carry broader significance for the application of MDF frameworks to constrain the terrestrial C-balance at regional and national scales.

1 Introduction

Over the past decade, terrestrial ecosystems have provided a global net carbon (C) sink sequestering $\sim$3.4$\pm$0.9 PgC yr^{-1}, $\sim$30% of anthropogenic CO_2 emissions, despite estimated emissions of $\sim$1.6$\pm$0.7 PgC yr^{-1} associated with land-use and land-cover change (Friedlingstein et al., 2020). The future trajectory of the terrestrial carbon sink will therefore have a significant impact on global efforts to achieve the goal of the UN Framework Convention on Climate Change to avoid dangerous climate change, reaffirmed in the Glasgow Climate Pact (UNFCCC, 2021). Quantification of spatial and temporal variations in exchange magnitude, alongside their associated uncertainties, are therefore essential to understanding the stability of the terrestrial carbon sink in the face of rapid environmental change (Hurlbert et al. , 2019), and prerequisite to robust national reporting of land-based CO_2 emissions and their attribution to different sectors (Grassi et al., 2017; Jones and Friedlingstein, 2020; McGlynn et al., 2022). Terrestrial biosphere models provide a means of quantifying the land carbon balance in a systemic, ecologically coherent way (Bonan et al., 2018). However, the current and future dynamics of terrestrial C exchange are highly uncertain, largely due to uncertainties in the structure and parameter constraints of the biosphere models themselves (Lovenduski and Bonan , 2017; Smallman et al., 2021).

Global land-use and land-cover change has increased the fragmentation of ecosystems, creating highly heterogeneous landscapes that host a mosaic of land-cover and uses (Lindenmayer and Fischer, 2013; Brink et al., 2017; Matricardi et al., 2020). This heterogeneity juxtaposes ecosystems with contrasting C stocks, traits and ecological processes, management, and environmental sensitivity, at length-scales of 10-100m. For example, within the UK, landscapes comprise a patchwork of managed arable land and pasture, semi-natural and plantation woodland, heath and settlements (Figure 1). Insight into the dynamics of this patchwork of ecosystems has been greatly accelerated by the proliferation of Earth Observation (EO) data from satellites that monitor ecosystems with ever-increasing spatial and temporal resolution (Exbrayat et al., 2019). A major challenge is to synthesise this expanding range of EO data to generate systemic understanding of the terrestrial C cycle, thus transforming ecosystem observation into ecological understanding that can inform policy development and facilitate land management (Smallman et al., 2022).

Model-Data Fusion (MDF) frameworks provide the means to integrate EO observations with spatially explicit process-based ecosystem models that encapsulate our understanding of how C flows through ecosystems (Luo et al., 2011), thereby providing key, mass-balanced, constraints on the fluxes of C between the atmosphere and land surface alongside their associated uncertainties (Niu et al., 2014; Bloom et al., 2016; Peylin et al., 2016; MacBean et al., 2018; Smallman et al., 2021). MDF frameworks that exploit intermediate complexity models of the terrestrial C cycle, such as CARDAMOM (Bloom et al., 2016;

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Exbrayat et al., 2018; Lopez-Blanco et al., 2019; Smallman et al., 2021), are able to generate "local" calibrations based on pixel-level inversions of EO and auxiliary data streams. Calibrating ecosystem models to local data is important, because the functional traits of ecosystems vary in space (Smith et al., 2013; Reich et al., 2014; Butler et al., 2017; Exbrayat et al., 2018; Lopez-Blanco et al., 2019; Smallman et al., 2021), with trait differences within biomes often exceeding differences between biomes (Van Bodegom et al., 2012; Butler et al., 2017); failure to account for such variations may lead to biases in the estimated dynamics (Scheiter et al., 2013; Exbrayat et al., 2018). However, the computational intensity of large-scale MDF frameworks limits their spatial resolutions to 10-100 km, several orders of magnitude greater than the length scales relevant to differentiating the ecosystems within landscape mosaics (e.g. Kaminski et al., 2012; Smith et al., 2013; Kuppel et al., 2014; Bloom et al., 2016; Peylin et al., 2016; Yin et al., 2020; Smallman et al., 2021).

The scale disparity between model domains and the ecological fabric those domains represent poses a major challenge to large-scale modelling of terrestrial C dynamics in heterogeneous landscapes (Stoy et al., 2009; Fisher et al., 2020; Levy et al., 2022). In typical spatially distributed MDF applications, available observations are aggregated to pixel-level “community averages”, prior to inversion. There are usually sufficient degrees of freedom in ecological process models to fit the observed temporal changes in aggregated stocks and fluxes, based on available observation constraints (Beven et al., 2006; Famiglietti et al., 2021). Nevertheless their ecological fidelity may be limited in heterogeneous landscapes, for which the parameters retrieved by these “community-average” models provide intermediate representations of the distinct ecosystems present. This limitation compromises efforts to attribute fluxes to specific land-uses and raises potential for significant sources of bias when estimating the terrestrial carbon balance and its environmental sensitivity. Firstly, the C-cycle represents the interplay of a number of nonlinear ecological processes, and therefore upscaling raises the familiar foe of Jensen’s inequality (Jensen, 1906; Levy et al., 2022), whereby for a set of input variables, X, the expectation value for a nonlinear function f (i.e. $E[f(X)]$), will not yield the same estimate as the same nonlinear function applied to the average values of those variables ($f(E[X])$), leading to scale-variance. In the case of terrestrial C fluxes, land-use places strong controls both on the distribution of carbon stocks and plant functional traits within the landscape, but also on the processes controlling the fluxes of C through landscapes, particularly those related to exogenous processes such as management and disturbance. Failing to account for the co-location of stocks and process imposed by land-use in mixed-use landscapes (e.g. concentration of C stocks in woodland, where timber harvest is focused) provides a clear source of potential scale-variant bias in derived flux estimates across large scales. Additionally, "community average" models may miss or poorly represent processes specific to certain land-uses (White et al., 2019; Kondo et al., 2020). To ensure the ecological fidelity of large-scale ecosystem C-cycle models, it is therefore vital to adequately capture the essential processes controlling the fluxes of C through these different ecosystems, and their potentially divergent temporal dynamics and environmental sensitivities (Levy et al., 2022).

In this study, we specifically address the impact of the resolution trade-off in spatially explicit MDF frameworks between ecological fidelity and computational intensity by investigating how simulated carbon cycling in a mixed-use landscape in the UK respond to the spatial resolution of the model grid. We test two different MDF approaches that assimilate EO information of ecosystem characteristics to constrain model parameters and uncertainty for an intermediate complexity model of the terrestrial C cycle, DALEC (Williams et al., 2005; Bloom et al., 2016; Smallman et al., 2017, 2021). In the first approach, ecosystems

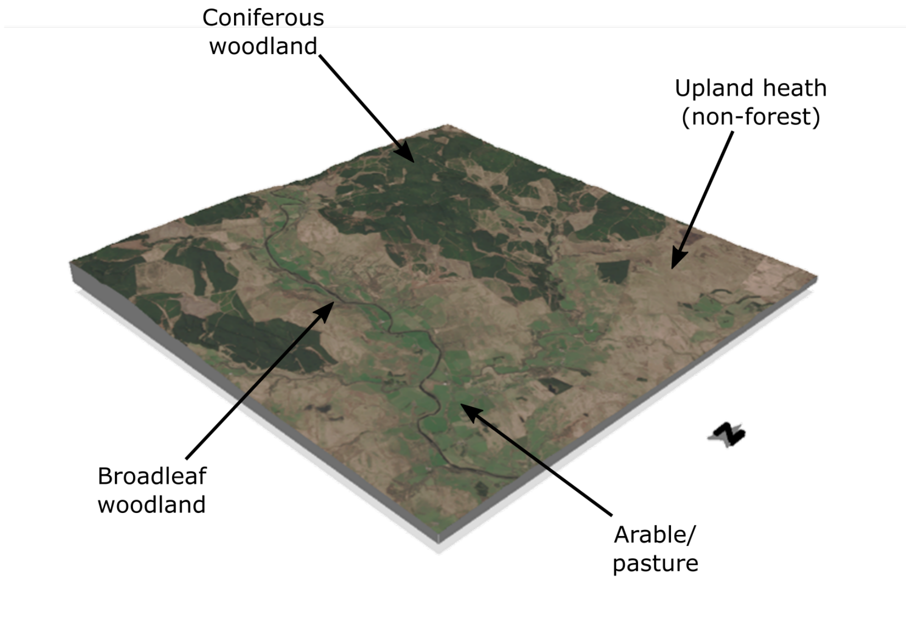

Figure 1. Perspective view of Sentinel 2 imagery over a typical landscape sampled from the study area illustrating the fine-scale mosaic of land-use characteristic of this region. The spatial extent of the displayed domain is 10 km x 10 km, and comprises part of the North River Tyne catchment in Northumberland, with the spatially extensive coniferous woodlands of Kielder Forest encroaching into the NW of the scene (top). Contains modified Copernicus Sentinel data [2021].

are calibrated and modelled at the pixel level, representing a "community average" of the encompassed land-cover and management. This corresponds to the approach commonly employed in large-scale ecosystem MDF frameworks (e.g. Smith et al., 2013; Bloom et al., 2016; Yin et al., 2020; Smallman et al., 2021). In the second, we stratify each pixel based on land-cover, and calibrate the model independently using remotely sensed data specific to each stratum, aligning more closely with the tiled Plant Function Type (PFT) approach employed in many terrestrial biosphere models (e.g. Sitch et al., 2008; Kaminski et al., 2012; Kuppel et al., 2014). The novelty of introducing stratification within a MDF context is that we use fine-scale ecosystem information contained within EO data to retrieve locally calibrated parameter ensembles for the ecosystems represented by each stratum, and therefore retain that key advantage of MDF systems, which enables calibrated traits to vary across environmental gradients, within the constraints of the available observations and ecological knowledge (Smallman et al., 2022). We test the two MDF approaches - the novel sub-pixel stratification approach and the traditional pixel average (baseline) approach - on grid resolutions spanning 1° to 0.05°. Specifically we address the following hypotheses:

– *H1*: estimated C fluxes will be scale variant, with stronger resolution-sensitivity exhibited by exogenous fluxes (i.e. disturbance) compared to biogenic fluxes (e.g. GPP).

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Biogeosciences Discussions

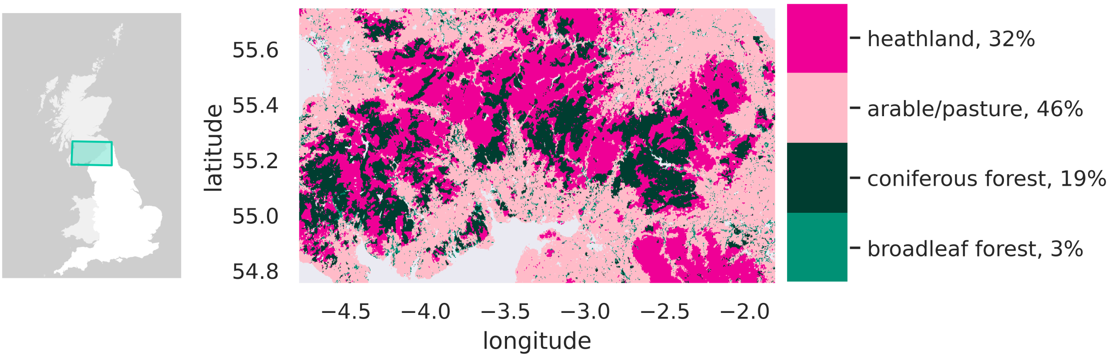

Figure 2. Map of the study area, spanning one° of latitude and three° of longitude across southern Scotland and northern England. The land-cover types displayed are aggregated from the LCM2015 land-cover map of Great Britain (Rowland et al., 2017), regridded to approximately 300m resolution.

- *H2*: C fluxes will be more consistent across grid resolutions when the framework explicitly accounts for sub-pixel heterogeneity in land-use; estimates from the baseline (unstratified) experiments will converge on the stratified estimates at finer spatial resolutions.

- *H3*: the model parameters retrieved for contrasting land-cover types will have contrasting optimal parameters, reflecting fundamental variations in ecosystem function that will drive divergent carbon dynamics in simulations of future trajectories; aggregation degrades the ecological information embedded in the retrieved parameters unless observational constraints are stratified prior to assimilation.

Using MDF to combine models and data at local scale offers huge potential for rigorous quantification of the state and dynamics of the terrestrial C cycle across large spatial scales, with propagation of uncertainty through analyses (Bloom et al., 2016; Smallman et al., 2022). In testing the hypotheses outlined above, we seek to address a key challenge relating to the mismatch between the scales of ecological processes and of large-scale MDF frameworks through the development of a novel stratified MDF framework. Our approach retains the core advantages of MDF, namely local calibration with local information, while also capturing the fine-scale ecosystem heterogeneity common to fragmented or mixed-use landscapes.

2 Methods

2.1 Study Area

The study area for this site covers northern England and the Scottish Borders, spanning approximately 30,000 km across three degree of longitude and one degree of latitude (Figure 2). The region includes nationally significant forestry estates of Kielder Forest, Eskdalemuir Forest and Galloway Forest, includes the Northumberland and Lake District National Parks, and comprises a mosaic of land-cover types, including coniferous plantation forest and fragments of broadleaf woodland, upland heath, arable

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

agriculture and pasture. The longitudinal extent stretches from coast to coast, from the Firth of Clyde in the West to the North Sea in the East. Elevation varies from sea level to a high of 978 m on Scafell Pike in the Lake District. These gradients in longitude and elevation are associated with gradients in both precipitation and temperature (Jenkins et al., 2009). Precipitation decreases from west to east in response to the prevailing westerly wind direction and orographic enhancement of rainfall in areas of high topography; temperature gradients are broadly controlled by elevation.

2.2 Model-data fusion with CARDAMOM (CARbon DAta MOdel fraMework)

2.2.1 DALEC

At the core of our model-data fusion framework sits DALEC, an intermediate complexity model of the terrestrial C cycle (Williams et al., 2005; Bloom and Williams, 2015; Smallman et al., 2017; Famiglietti et al., 2021). DALEC is a mass balance model of the C cycle with carbon moving through different pools based on parameterised fluxes (Figure 3). A number of variants of DALEC have been created representing ecosystem carbon dynamics with varying degrees of complexity (Famiglietti et al., 2021; Smallman et al., 2021). The specific version of DALEC used here corresponds to the C6 model outlined in Famiglietti et al. (2021) which combines the C-cycle structure from Bloom and Williams (2015) with the revised photosynthesis model from Smallman and Williams (2019). There are four live biomass pools, specifically relating to carbon stored in foliage, labile carbon, fine roots and wood, and two dead organic carbon pools: litter and soil organic carbon. Carbon enters the system through GPP, modelled using the photosynthesis model ACM2 (Smallman and Williams, 2019), wherein GPP is simulated as a function of modelled leaf area, estimated canopy photosynthetic efficiency, absorbed solar radiation, atmospheric CO_2 concentration, air temperature and a stomatal conductance model that balances potential water supply from the soil (assumed to be at field capacity) through the roots with atmospheric demand, determined by absorbed solar radiation and VPD. Carbon is lost from the system via autotrophic and heterotrophic respiration. NPP is allocated between autotrophic respiration and the live pools based on fixed fractions. Canopy growth is driven by a combination of direct allocation from GPP and transfer of carbon from the labile pool. The flux of carbon from the labile pool to foliage and canopy senescence, driving litter-fall, are controlled by a simple day-of-year phenology model with a parameterised leaf life span (Bloom and Williams, 2015). Carbon flows from the roots and wood into the litter and soil organic carbon pools respectively based on first order turnover rates. Heterotrophic respiration fluxes also follow first order kinetics, but with an exponential temperature sensitivity. A full list of model parameters is provided in the appendix (Table A1). The relative simplicity compared to other terrestrial biosphere models make DALEC amenable to calibration in model-data fusion frameworks, and allows propagation of uncertainties through large ensemble simulations (Bloom et al., 2016; Exbrayat et al., 2018; Famiglietti et al., 2021; Smallman et al., 2021).

2.2.2 Model-Data Fusion

Our model-data fusion framework, CARDAMOM (Bloom et al., 2016), uses a Bayesian approach within an Adaptive Proposal Markov Chain Monte Carlo framework (Haario et al., 2001) that can assimilate a range of information (Bloom et al., 2016), including remotely sensed LAI and aboveground biomass (see Section 2.3). The premise of the approach is to take driving data

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Biogeosciences Discussions

describing the meteorology and disturbances such as forest clearance and fire, and search the model parameter space to find parameter combinations that provide simulated dynamics that are consistent with the available data. Specifically, given a set of observations, O, with uncertainty σ, the probability of a given parameter set x, P($x|O$), is calculated as a function of the likelihood of the observations given the current parameters, P($O|x$), and any prior knowledge on the parameter distributions, P(x):

$$\mathrm{P}(x|O) \propto \mathrm{P}(O|x) \cdot \mathrm{P}(x) \tag{1}$$

The likelihood P($O|x$) is calculated based on the misfit between the N available observations and the equivalent simulated state variables and fluxes for each parameter set, M:

$$\mathrm{P}(O|x) = exp\left(-0.5 \cdot \sum_{n=1}^{N} \left(\frac{O_n - M_n}{\sigma_n}\right)^2\right) \tag{2}$$

To facilitate the calibration process, we employ a series of Ecological Dynamic Constraints, EDCs (Bloom and Williams, 2015; Smallman et al., 2017). EDCs comprise a series of rules that ensure ecological “realism” in the accepted parameter sets. For example, turnover of the wood carbon pool must be slower than foliage turnover. Where EDCs are not satisfied, the likelihood is set to zero. By restricting the acceptable parameter space, the EDCs therefore reduce the effective model complexity (Famiglietti et al., 2021). The resulting ensemble of parameter sets encapsulate the uncertainty in the calibration within the available observational constraints.

2.2.3 Stratification approach for mosaic landscapes

Our approach to handling fine-scale heterogeneity during the model-data fusion process is based on sub-pixel stratification based on land-use (Figure 3). Stratification is achieved by sampling the spatially gridded EO data products at their native spatial resolution based on a reference land-cover map, resampled to the same resolution using the modal category. The specific land-cover product used is the LCM2015 land-cover map produced by the UK Centre for Ecology and Hydrology (CEH) (Rowland et al., 2017), which we aggregate to four classes: coniferous woodland, broadleaf woodland, arable/pasture and heathland, which includes semi-natural grasslands and widespread areas of non-wooded upland (Figure A1). Urban and coastal areas are masked from all analyses. For each pixel, separate ensembles are calibrated independently, yielding a suite of ensembles that maintain the ecological fidelity of the calibrated parameters. This is a contrast to the "traditional" model-data fusion approach, which aggregates the data constraints into pixel-level "community-averaged" prior to calibration, yielding calibrated parameter combinations that may be attempting to account for a multitude of distinct ecological processes. The stratification approach is very flexible. The number of categories can be refined as necessary, within the data constraints. Importantly, different ecosystems can be modelled with distinct, ecosystem-specific models that better capture their functional process dynamics, for example woodlands (Smallman et al., 2017), pasture (Myrgiotis et al., 2021) and arable agriculture (Revill et al., 2021). However, for simplicity of comparison across the experiments in this study against the baseline (i.e. no stratification), we use only one model structure across all strata. For strata where woody tissues are not part of the dominant vegetation

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Biogeosciences Discussions
Open Access
EGU

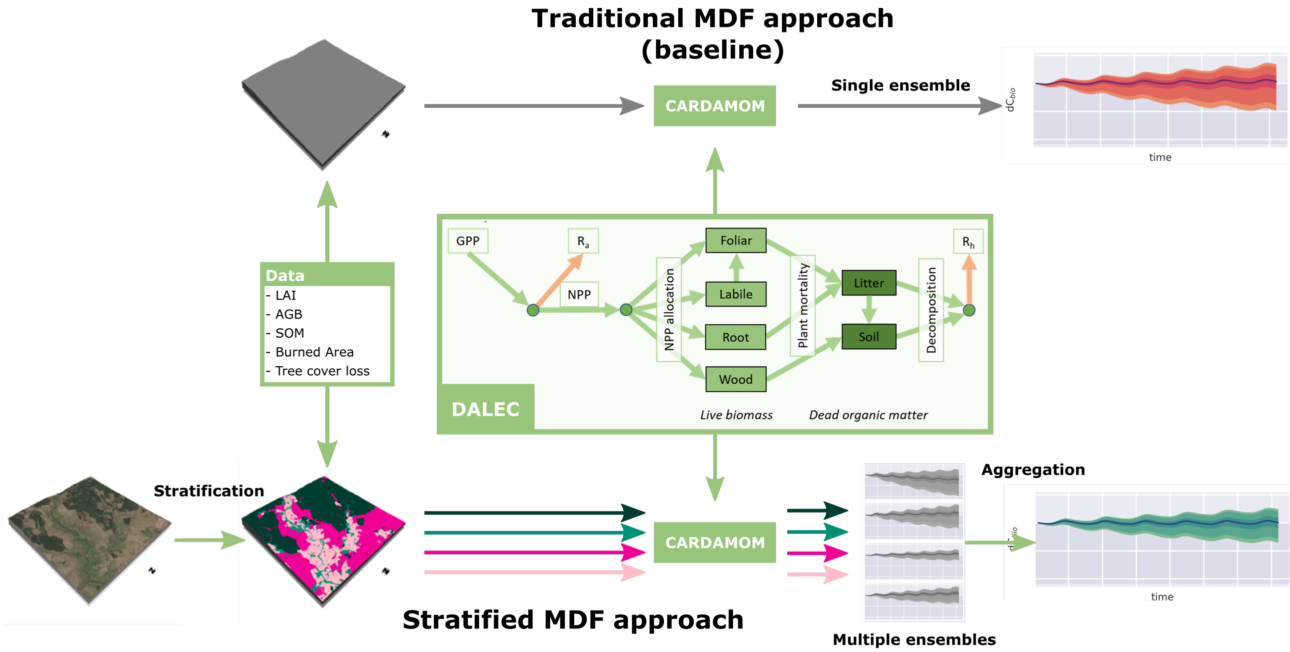

Figure 3. Schematic flow diagrams illustrating the different model-data fusion approaches employed in this study: the "traditional" model-data fusion approach whereby the input data are aggregated to pixel-level "community averages" and stratification based on the land-use leading to calibration of a suite of land-use specific ensembles. At the heart of both MDF approaches sits DALEC, an intermediate complexity model of the terrestrial C cycle (Bloom and Williams, 2015). While the presented time series show specifically the change in live biomass, dC_{bio}, it is important to note that similar information is retrieved for all fluxes and stocks within DALEC, alongside pixel-specific parameter ensembles.

190 types, the C_{Wood} pool is also a reservoir for non-woody structural tissue, for example in areas covered by crop and pasture. Regarding aggregation of uncertainties, we do not have constraints on the extent to which pixel uncertainties are correlated in space. Therefore for each stratum, spatial aggregation of uncertainty conservatively assumes correlated uncertainties (Exbrayat et al., 2018). However, we assume that individual strata, representing different ecosystems, are uncorrelated with other strata when aggregating sub-pixel ensembles to pixel-level.

195 ## 2.3 Data

2.3.1 Meteorological drivers

Meteorological drivers, comprising temperature, shortwave radiation, vapour pressure deficit (VPD) and wind speed, are drawn from the CRU-JRAv1.1 dataset, a 6-hourly 0.5 × 0.5° reanalysis (CRU, 2019). Atmospheric CO_2 concentration is taken from the Mauna Loa global CO_2 concentration (www.esrl.noaa.gov/gmd/ccgg/trends/, accessed: 22/08/2020).

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

2.3.2 Copernicus LAI 300 m

LAI data is obtained from the 300 m Copernicus LAI product v1.0 (Fuster et al., 2020) for period 2014-2019. The LAI estimates in the Copernicus 300 m product represent 10-day composites from daily estimates of LAI that are generated from to daily Top-of-Atmosphere input reflectances detected by the PROBA-V satellite by applying a neural network. These 10-day LAI estimates were aggregated to monthly averages prior to assimilation. Pixel-wise uncertainty estimates are also provided with this product, calculated as the root-mean-square difference between the individual daily neural network estimates and the 10-day average. Previous work has indicated that these uncertainty estimates underestimate the true uncertainties associated with this product (Zhao et al., 2020). We therefore used a more conservative temporal aggregation approach based on the maximum uncertainty within the aggregation period.

2.3.3 ESA Biomass CCI Aboveground Biomass 2017, 2018

Aboveground biomass (AGB) estimates and associated uncertainty were extracted the global maps published within the ESA Biomass CCI collection (Version 2), comprising two estimates for the years 2017 and 2018 with as spatial resolution of 100 m (Santoro et al., 2021). The ESA CCI Biomass data are derived from Synthetic Aperture Radar (SAR) backscatter data, specifically ALOS PALSAR L-band SAR backscatter combined with Sentinel-1 C-band SAR backscatter. Uncertainty estimates are provided with this product, calculated as the standard deviation associated with the AGB estimate after propagating errors through the SAR measurement, SAR-AGB modelling framework and merging of L-band and C-band estimates into an overall AGB estimate (Santoro et al., 2021).

The DALEC wood carbon pool represents the combination of above- and below-ground carbon (i.e. including the coarse root component). The contribution from below-ground biomass (BGB) to the woody biomass pool, alongside the associated uncertainty, is modelled using an allometric relationship following Saatchi et al. (2011):

$$\mathrm{BGB} = 0.489 \cdot \mathrm{AGB}^{0.89} \tag{3}$$

2.3.4 SoilGrids2 Soil Organic Carbon (SOC)

Soil organic carbon estimates and associated uncertainties were obtained from SoilGrids2, which provide 250m resolution spatial maps of depth profiles for various soil properties (Poggio et al., 2021). These maps were produced using EO and auxiliary spatial data within a machine learning framework trained on over 230,000 individual soil profile observations. The extracted SOC estimates are used to set a prior constraint on the initial SOC stock.

2.3.5 Disturbance

Disturbance is imposed on DALEC based on satellite observations of tree cover loss and burned area. Disturbances related to tree cover loss are driven by observations from the Global Forest Watch (GFW) dataset (Hansen et al., 2013), which provides annual constraints on tree cover loss at 30 m resolution based on Landsat data. Note that other mechanisms of disturbance, such as agricultural harvests and pasture management, are not considered in the current analysis. Fire is imposed based on monthly aggregated burnt area fractions in the MODIS MCD64A1 product (Giglio et al., 2018), which maps fire-affected areas at 500m resolution based on changes in surface reflectance, although the occurrence of MODIS-detected fires throughout the model domain was very low. Emissions from fire are estimated by assuming a fraction of simulated biomass either undergoes combustion, therefore immediately released to the atmosphere, or is transferred to the litter pool, based on tissue specific combustion-completeness factors (Exbrayat et al., 2018).

2.4 Experimental setup

To test how simulated C fluxes varied with grid resolution we calibrated DALEC across the target domain at four different grid resolutions: 0.05°, 0.25°, 0.50° and 1.00°, at a monthly time-step spanning the period 2014-2019. We compared the retrieved parameters and simulated C fluxes for two MDF approaches: the proposed stratified CARDAMOM calibration that explicitly accounts for sub-pixel heterogeneity in land-use, and the traditional pixel aggregate CARDAMOM calibration. The latter serves as a baseline. In all cases we use the same underlying DALEC model structure within the MDF framework (Figure 3). Emergent differences in the retrieved parameters, stocks and fluxes between experimental runs are therefore a consequence of the resolution at which land-cover and land-use are aggregated, rather than ecosystem-specific differences in model structure. We characterise the calibration performance for each ensemble based on the RMSE and the bias with respect to the N assimilated observations:

$$\text{RMSE} = \sqrt{\frac{\sum_{n=1}^{N}(O_n - M_n)^2}{N}} \tag{4}$$

$$\text{Bias} = \frac{\sum_{n=1}^{N}(O_n - M_n)}{N} \tag{5}$$

In both cases, we weight the contributions from individual pixels when aggregating across the domain based on the fractional coverage contributed by each stratum. As the observations are also associated with significant uncertainty, we also consider the ratio of the RMSE and Bias to the product uncertainty as a measure of agreement within the uncertainty constraints provided by the assimilated data.

We are able to address our first two hypotheses (*H1*, *H2*), relating to the impact of resolution and sub-pixel stratification on diagnostic analyses of C cycle dynamics, by comparing the changes in C stocks and fluxes over the data assimilation period. *H3* is addressed by comparing the retrieved parameters for each run, including the individual land-use classes in the stratified analysis. To understand the potential impact of any emergent resolution dependence of the retrieved parameters on future trajectories, we then ran forward simulations of our DALEC ensembles to 2100 under the SSP2-4.5W m^{-2} scenario extracted from the UK Earth System Model (UKESM; Sellar et al., 2019) contribution to CMIP6 (Eyring et al., 2016), which corresponds to a middle-of-the-road scenario with a projected mean global warming of 2.7^{o}C (O'Neill et al., 2016). We do not impose future disturbance fluxes, so emergent differences in C dynamics will be driven by the interactions between climate and the retrieved parameters for each ensemble. To avoid step-changes in meteorology between the historical meteorology (from

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

observations) and future meteorology (simulated by UKESM) we apply the future trajectories for each scenario based on the anomaly in the UKESM forecast relative to 2019 (following Smallman et al., 2021).

3 Results

Table 1. Summary of calibration performance, aggregated across the domains for the baseline and stratified experiments. σ represents the uncertainty of the assimilated observation data, thus RMSE / σ provides the ratio of the RMSE to the uncertainty attached to the observation constraint. For an equivalent breakdown of calibration performance of the individual strata in the stratified experiment, see Table A2.

Variable	Version	Metric	1.00°	0.50°	0.25°	0.05°
C_{Wood}	Baseline	RMSE / gCm^{-2}	696 (14.9 %)	735 (14.9 %)	716 (13.7 %)	696 (12.5 %)
C_{Wood}	Baseline	RMSE / σ	0.27	0.27	0.25	0.22
C_{Wood}	Baseline	Bias / gCm^{-2}	-695 (-14.9 %)	-728 (-14.8 %)	-688 (-12.5 %)	-584 (-6.6 %)
C_{Wood}	Baseline	Bias / σ	-0.27	-0.27	-0.23	-0.13
C_{Wood}	Baseline	Median gCm^{-2}	3951	3921	3962	4058
C_{Wood}	Stratified	RMSE / gCm^{-2}	632 (13.6 %)	706 (14.6 %)	678 (13.4 %)	694 (12.7 %)
C_{Wood}	Stratified	RMSE / σ	0.25	0.27	0.24	0.23
C_{Wood}	Stratified	Bias / gCm^{-2}	-631 (-13.6 %)	-702 (-14.6 %)	-665 (-12.8 %)	-627 (-8.8 %)
C_{Wood}	Stratified	Bias / σ	-0.25	-0.27	-0.23	-0.17
C_{Wood}	Stratified	Median gCm^{-2} 4062	3979	4003	4027	
LAI	Baseline	RMSE / m^{2}m^{-2}	0.36 (15.4 %)	0.37 (15.9 %)	0.39 (16.9 %)	0.44 (19.1 %)
LAI	Baseline	RMSE / σ	0.34	0.35	0.37	0.42
LAI	Baseline	Bias / m^{2}m^{-2}	-0.02 (-0.7 %)	-0.02 (-0.7 %)	-0.02 (-0.9 %)	-0.02 (-1.0 %)
LAI	Baseline	Bias / σ	-0.01	-0.02	-0.02	-0.02
LAI	Baseline	Median m^{2}m^{-2}	2.19	2.19	2.18	2.19
LAI	Stratified	RMSE / m^{2}m^{-2}	0.34 (14.9 %)	0.36 (15.8 %)	0.39 (16.8 %)	0.44 (19.2 %)
LAI	Stratified	RMSE / σ	0.33	0.35	0.37	0.42
LAI	Stratified	Bias / m^{2}m^{-2}	-0.02 (-0.8 %)	-0.02 (-0.9 %)	-0.02 (-0.9 %)	-0.03 (-1.4 %)
LAI	Stratified	Bias / σ	-0.02	-0.02	-0.02	-0.03
LAI	Stratified	Median m^{2}m^{-2}	2.19	2.18	2.18	2.18

3.1 Calibration performance

The two MDF approaches tested provided comparable fits to the calibration data. Both the baseline and stratified CARDAMOM calibrations were able to fit the assimilated C_{Wood} and LAI observations to well within the levels of observation uncertainty, with the RMSE between simulated and observed variables less than 50% of the uncertainties attached to the assimilated observations (Table 1, Figures A2, A3). In general the RMSE values were comparable between the stratified and baseline experiments (mean RMSE for C_{Wood} across spatial resolutions: 14.0% for the baseline experiment; 13.6% for the stratified experiment). For the C_{Wood} estimates, the RMSE tended to be lowest at the finest spatial resolution in both experiments, as did the bias (Table 1). In the stratified experiment, the bias in C_{Wood} was dominated by the contributions from the woodland strata

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Table 2. Summary of domain-aggregated carbon budgets for the baseline and stratified experiments. Fluxes are gross primary productivity (GPP), total ecosystem respiration (R_{eco}), cumulative changes in live (dC_{bio}) and dead (dC_{soil}) organic C pools integrated over the six year assimilation period (2014-2019), and carbon losses due to harvest and other tree cover loss (harvest). Values represent the median pixel level estimates averaged across the domain, alongside the 5% and 95% percentiles, i.e. assuming fully correlated uncertainties.

Flux	Version	1.00°	0.50°	0.25°	0.05°
GPP / $gCm^{-2}d^{-1}$	baseline	5.04 (3.69 - 6.32)	5.05 (3.69 - 6.34)	5.03 (3.67 - 6.35)	5.00 (3.62 - 6.33)
GPP / $gCm^{-2}d^{-1}$	stratified	4.98 (4.17 - 5.77)	4.98 (4.08 - 5.84)	4.99 (4.03 - 5.92)	4.94 (3.84 - 6.02)
R_{eco} / $gCm^{-2}d^{-1}$	baseline	4.59 (3.03 - 6.65)	4.61 (3.07 - 6.68)	4.58 (3.06 - 6.69)	4.59 (3.06 - 6.71)
R_{eco} / $gCm^{-2}d^{-1}$	stratified	4.59 (3.64 - 5.74)	4.57 (3.55 - 5.85)	4.58 (3.51 - 5.97)	4.58 (3.34 - 6.23)
cumulative dC_{bio} / gCm^{-2}	baseline	282.6 (-1755.1 - 1708.0)	214.9 (-1931.0 - 1698.7)	237.0 (-1855.9 - 1613.0)	164.9 (-1937.9 - 1521.7)
cumulative dC_{bio} / gCm^{-2}	stratified	-5.9 (-957.9 - 823.7)	21.4 (-1090.1 - 873.8)	22.9 (-1160.3 - 914.9)	30.0 (-1449.6 - 1075.8)
cumulative dC_{soil} / gCm^{-2}	baseline	501.6 (-2207.4 - 2582.1)	497.7 (-2247.6 - 2605.4)	507.8 (-2234.9 - 2518.8)	520.6 (-2121.9 - 2517.1)
cumulative dC_{soil} / gCm^{-2}	stratified	425.7 (-957.5 - 1595.2)	428.2 (-1159.8 - 1640.2)	426.3 (-1244.3 - 1723.4)	477.1 (-1515.9 - 1990.2)
harvest / $gCm^{-2}d^{-1}$	baseline	0.052 (0.023 - 0.091)	0.055 (0.023 - 0.096)	0.064 (0.025 - 0.110)	0.085 (0.031 - 0.147)
harvest / $gCm^{-2}d^{-1}$	stratified	0.130 (0.040 - 0.228)	0.127 (0.040 - 0.226)	0.129 (0.043 - 0.227)	0.129 (0.042 - 0.225)

(Table A2), corresponding to their much greater C_{Wood} stocks, which were over four times higher in coniferous woodland than arable/pasture, and six times higher than in the heathland class (Table A2, Figure A3). In contrast to C_{Wood}, the relative RMSE attached to the LAI estimates tended to increase at finer resolutions in both experiments (Table 1; Table A2). Notably, all strata, including the coniferous woodland class, contained a strong seasonal cycle of monthly LAI in both the assimilated observations and the simulations (Figure A2).

3.2 Terrestrial C budget and impact of spatial resolution on C flux estimates

Both the baseline and stratified CARDAMOM analyses estimated the net ecosystem exchange of C to be most likely a net sink of C over the calibration period. C uptake from the atmosphere via GPP was ∼5 gC $m^{-2}d^{-1}$ (Table 2; for stratified experiment, 0.05°, GPP = 4.94 (3.84 - 6.02) gC $m^{-2}d^{-1}$) while C returned to the atmosphere via autotrophic and heterotrophic respiration (R_{eco}) was ∼4.6 gC $m^{-2}d^{-1}$ (Table 2; for stratified experiment, 0.05°, GPP = 4.58 (3.34 - 6.23) gC $m^{-2}d^{-1}$). At increasingly fine spatial resolutions the model reveals greater spatial variability, reflecting the impact of landscape features and topography that are only adequately resolved at the finest grid scale (Figure A4). However, median estimates of GPP and R_{eco} were relatively insensitive to the spatial resolution once aggregated across the spatial domain (Figure A5), varying by ≤0.05 gC $m^{-2}d^{-1}$ (Table 2). The simulated uncertainties were smaller in the stratified experiment compared to the baseline for both GPP and R_{eco} (Figure 7). The reduced uncertainty with stratification is a result of assuming independence between strata. If the uncertainty in these fluxes is assumed fully correlated across strata, then the combined uncertainty in the stratified experiment is comparable to that of the baseline experiment. The degree to which stratification reduces uncertainty is therefore determined by the extent to which strata are considered independent.

In contrast to GPP and R_{eco}, the disturbance flux exhibited resolution dependence in the baseline experiment (1.00°: 0.05 (0.02 - 0.09) gC $m^{-2}d^{-1}$; 0.05°: 0.09 (0.03 - 0.15) gC $m^{-2}d^{-1}$), while the disturbance flux is insensitive to resolution in the

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

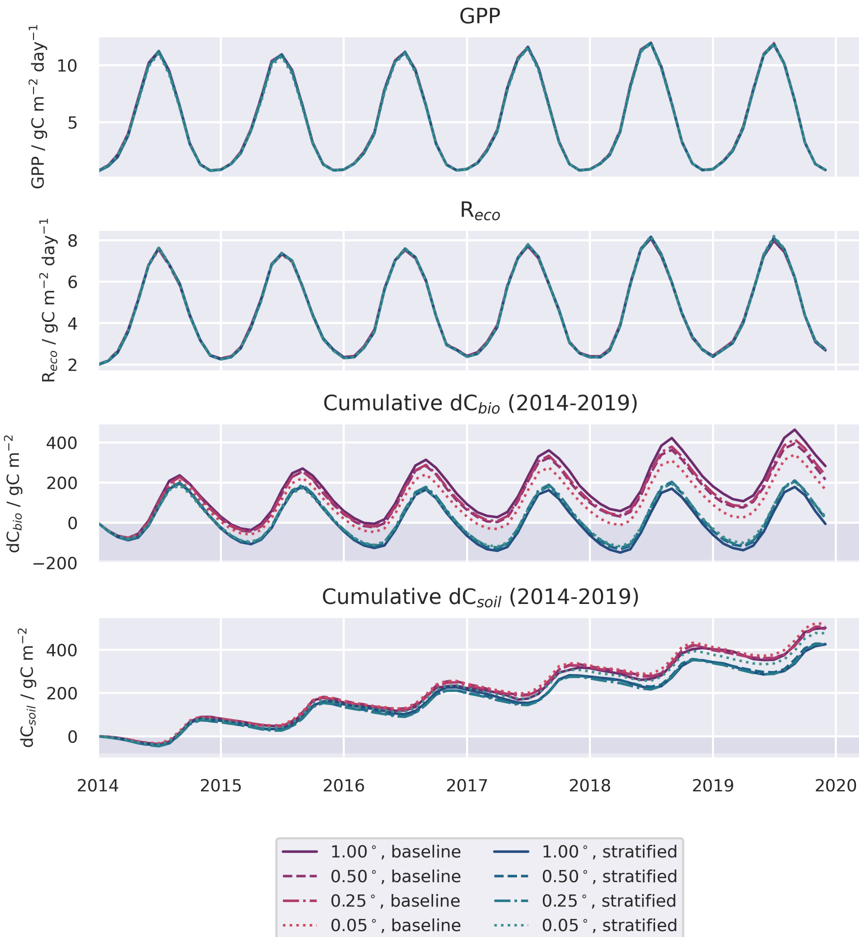

Figure 4. Spatially aggregated time series for GPP and ecosystem respiration (R_{eco} =R_a+R_{het}),and the cumulative change in carbon stocks in the live (dC_{bio}) and soil (dC_{soil}) pools, shown for the baseline and stratified ensembles for four spatial resolution domains. Only the median estimates are shown for clarity. Confidence levels for the 1° and 0.05° domains for the same time series are provided in Figure S2.

stratified experiment (0.13 (0.04 - 0.23) gC m^{-2}d^{-1}; see Table 2). In this set of experiments, the disturbance flux is driven by forest harvest (i.e. tree cover loss), as fire was negligible. The differences in simulated disturbance fluxes were sufficient to drive alternative interpretations of the terrestrial C balance: in the baseline experiment the live C pools suggest a strong C sink, whereas in the stratified experiment, C gains from NPP were balanced by losses due to the higher simulated harvest fluxes focused in conifer woodlands (Table 2, Figure A5). The baseline experiments simulate a tendency towards increasing C_{bio} stocks for all four spatial resolutions, with the median simulated sink strength increasing at coarser grid resolutions. This emergent scale-dependent sensitivity of dC_{bio} is not shared by the stratified experiment, for which median dC_{bio} is more consistent across the range of spatial resolutions. The disagreement in C accumulation between the two approaches is reduced

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

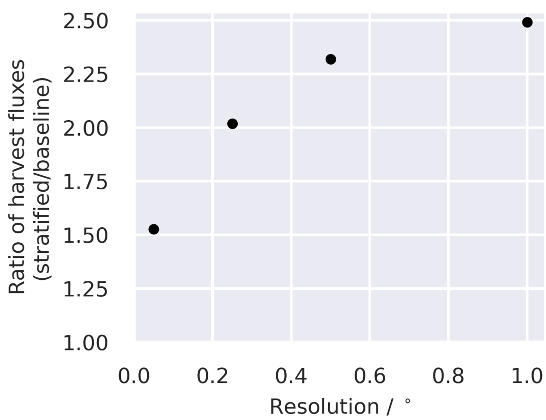

Figure 5. Ratio of the temporal mean harvest flux simulated in the stratified and baseline experiments, i.e (stratified/baseline), as a function of domain resolution. Points represent the median estimate aggregated across the spatial domain. Values >1 indicate baseline fluxes that are lower than the stratified case.

at finer resolution model domains. Considering the median cumulative dC$_{bio}$, the resolution-dependent bias for the baseline experiment compared to the stratified experiment declines by $>50\%$ moving from the 1.00° to the 0.05° domain (Table 2). However, the two approaches to not reach convergence even at 0.05° resolution (Figure A5).

Harvest fluxes in the stratified ensemble setup were between 2.5 and 1.5 times higher than the baseline ensemble, with the difference increasing systematically at coarser grid resolutions (Figure 5). Areas declining in dC$_{bio}$ are generally focused around the primary commercial forestry regions, where timber harvest is most abundant, for both the baseline and stratified experiments (Figure 6). However, it is evident that across the regions hosting significant areas of conifer woodland, the magnitudes of simulated net losses in the live carbon pools are generally greater for the stratified ensemble. The differences in dC$_{bio}$ between the stratified and baseline approaches are dominated by differences in simulated harvest fluxes between these experiments (R^2=0.71, Figure 6). In comparison, there are relatively weak relationships between the difference in dC$_{bio}$ to differences in either GPP (R^2=0.01) or R$_{eco}$ fluxes (R^2=0.07). In turn, differences in harvest flux between the stratified and baseline experiments are strongly influences by the level of sub-pixel heterogeneity (Figure 6), with the difference in the harvest flux between the two experiments declining as the fraction of pixels covered by the coniferous woodland class (where both harvest and live C stocks are concentrated) approaches full coverage. In other words, there is little difference between the two experiments where pixels have homogeneous land-use.

3.3 Impact of stratification on calibrated parameters and future C dynamics

The stratified data assimilation scheme reveals emergent differences between ecosystems, while traits retrieved for the baseline experiments characterised intermediate values (Figures 8, A6, A7, A8). Comparing the retrieved parameters for the different spatial resolution domains, it is evident that stratification leads to preservation of ecological information across resolutions,

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

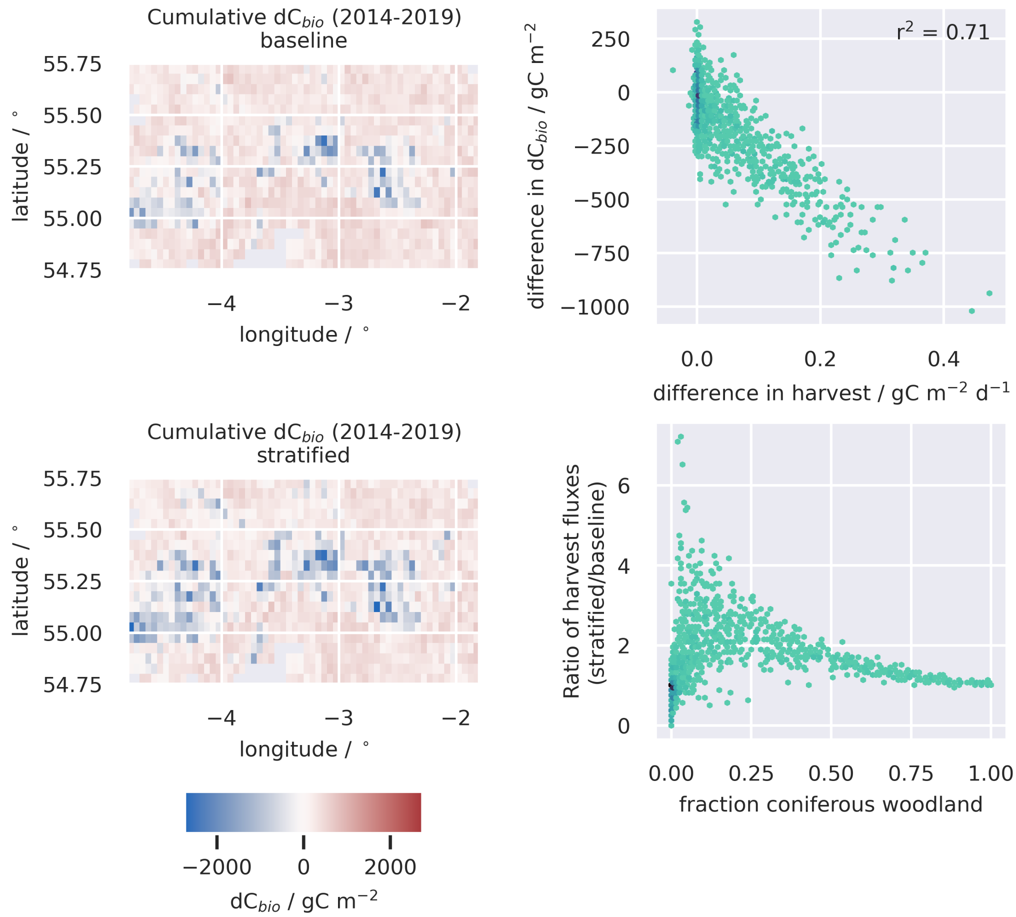

Figure 6. A comparison of changes in the C stocks aggregated across the live C pools for the baseline and stratified ensembles for the 0.05° domain (left); the relationship between the differences in simulated changes in live C stocks (dC_{bio}) and the difference in simulated harvest fluxes for the two approaches (upper right); and the ratio of the simulated harvest fluxes, i.e. (stratified/baseline), as a function of the fraction of the pixel covered by coniferous woodland (lower right). Large differences in simulated stock changes were generally associated with corresponding differences in the simulated timber harvest fluxes (R^2=0.71), rather than differences in GPP (R^2=0.04) or R_{eco} (R^2=0.01). The difference between the stratified and baseline harvest fluxes are greater in pixels with a mix of coniferous woodland and other land-cover classes and diminish in more homogeneous pixels.

such as longer-lived C_{Wood} pools in the woodland classes (e.g. Figure 8). In contrast, in the baseline experiments, aggregation to coarser spatial resolutions reduces the range of retrieved traits, and results in the shredding of information that relates to fundamental aspects of ecosystem function. The most pronounced differences in the calibrated parameters are associated

Biogeosciences Discussions

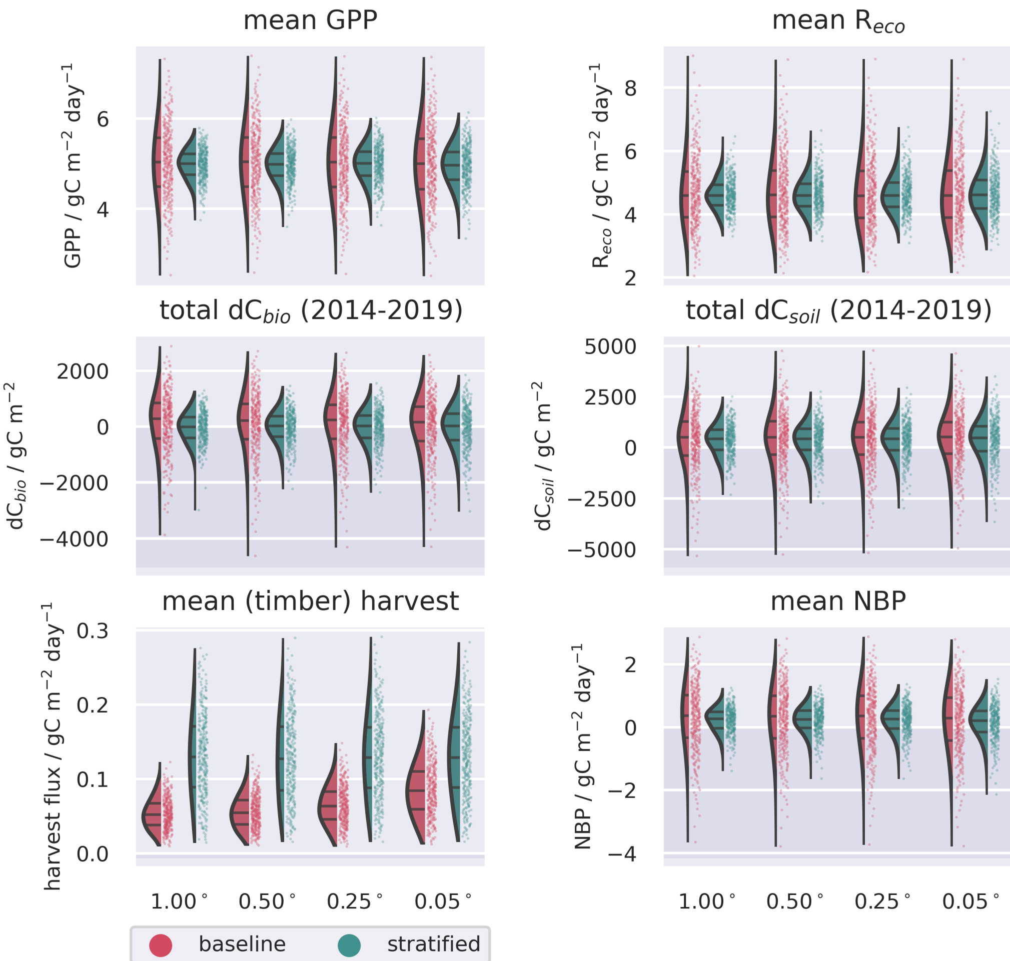

Figure 7. Ensemble distributions for the spatially aggregated ecosystem C fluxes and stock changes simulated with the 0.05° resolution domain, contrasting the baseline and stratified analyses. Fluxes represent temporal averages between 2014-2019; stock changes represent cumulative differences over the same period. Uncertainties are assumed to be fully correlated in space, but independent across the different land-cover strata in the stratified ensemble.

with parameters closely related to the dynamics of the C_{Wood} pool (Figure 8). Within the study landscape, live C stocks are concentrated in the woodland classes, particularly the conifer woodlands. Higher C stocks are reflected in greater woody

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

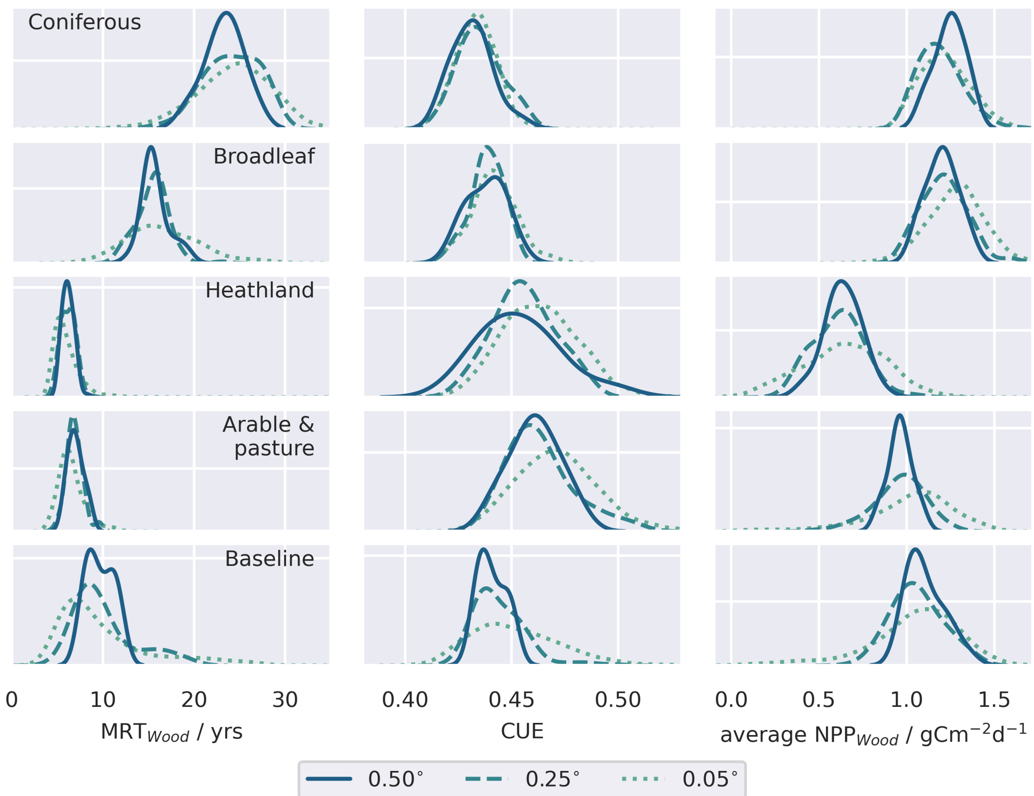

Figure 8. Comparison of a subset of retrieved parameters for the individual strata compared to the baseline retrieval for the different resolution domains. Distributions represent the pixel-level median parameter estimates weighted by the pixel-fraction estimated associated with each stratum. Similar distributions of the retrieved pool residence times for the live organic carbon pools and dead organic carbon pools are provided in figure A6 and figure A7 respectively. Allocation fractions of NPP to the live carbon pools are presented in figure A8.

productivity (Figure 8), higher residence times for the wood pool (MRT$_{Wood}$; Figure A6) and higher allocation fractions of NPP to the woody pool (Figure A8). In the baseline experiment, it is notable that large MRT$_{Wood}$ estimates are not represented within the distribution of median parameter estimates; this effect is magnified in the coarser domains, where the range of the simulated MRT$_{Wood}$ distribution contracts (Figures 8). The longest soil C residence times (MRT$_{Soil}$ were retrieved for

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

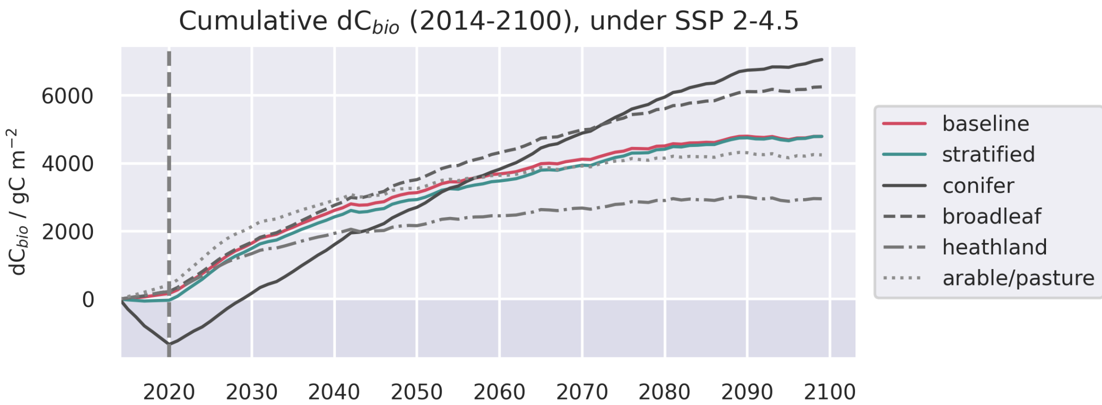

Figure 9. Evolution of live C pools aggregated across the baseline and stratified domains, alongside the individual strata, under three future trajectories of climate change. Climate trajectories were extracted from the UK Earth System Model (UKESM; Sellar et al., 2019) contribution to CMIP6 (Eyring et al., 2016). No disturbance was simulated after the end of the calibration period. The presented trajectories were taken from the 0.50° domain.

the heathland class (Figure A7), which includes upland areas underlain by peat deposits (Tanneberger et al., 2017). Again, the longer MRT_{Soil} estimates of the heathland stratum are not well represented within the baseline ensemble, and as with MRT_{Wood}, this loss of ecological information is exacerbated at coarser resolutions. Notably, there was little variation in leaf lifespan between classes (Figure A6), reflecting the strong seasonal cycle in the assimilated LAI observations for all classes, including coniferous woodland (Figure A2).

Differences in the parameter ensembles calibrated for each stratum lead to divergent C dynamics in response to future climate change (Figure 9). Both the stratified and baseline ensembles simulated increasing live C stocks up to 2100, assuming no disturbance (i.e. responding to climate and CO_2 only), with the rate of growth tailing off in the latter half of the forecast period. While the trajectories of the baseline and stratified ensembles were similar, the individual strata evolved along divergent paths (Figure 9). The increase, and rate of increase, of live C density (dC_{bio}) tracked differences in the retrieved MRT_{Wood} (Figure 8). Accumulation was greatest for the the coniferous and broadleaf forest strata, for which the median dC_{bio} increased throughout the simulation period. In contrast, the heathland class and arable/pasture class accrued C in live biomass more slowly, and had stopped accumulating C by the end of the forecast scenarios. Of the the two forest strata, the median rate of increase in C density was greatest for coniferous forest (Figure 9).

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

4 Discussion

4.1 Scale-variance of simulated carbon balance in heterogeneous ecosystems, *H1, H2*

Stratification of land-use is a prerequisite to scalable modelling of the C balance due to interaction of uneven distribution of C stocks, and spatially correlated disturbance fluxes. For our study spanning approximately 30,000 km^2 of mixed land-use in northern England and southern Scotland, we found contrasting resolution dependence for spatially distributed biogenic fluxes (GPP, R_{eco}) and disturbance fluxes, typically restricted to specific land-uses (e.g. timber harvest). In our baseline experiments, where we aggregated the assimilated data to the domain resolution without considering the underlying land-use, GPP and R_{eco} were insensitive to the grid resolution for model domains spanning a resolution range of 1° to 0.05° (Table 2, Figure A5). While we expected the biogenic fluxes to be less sensitive than exogenous fluxes, the invariance in biogenic fluxes with respect to both resolution and method was surprising. Within the version of DALEC used, in the absence of moisture stress GPP is closely controlled by the time series of LAI, which did not show strong variations between strata (Figure A2). Regions with greater variance in phenology, or more severe moisture stress might exhibit greater scale-dependence and a greater impact from stratification. Our model also ignored changes to litter pools associated with harvest; a more complex treatment incorporating coarse and fine residues might lead to greater sensitivity in R_h. Conversely, disturbance fluxes, dominated in this case by timber harvest, showed a clear resolution-dependence, with flux estimates approaching the stratified estimates at the finest grid resolutions (Figure 5), although the baseline estimates did not reach convergence even at 0.05° resolution (Table 2, Figure A5). The emergent resolution-dependence is a consequence of the fact that AGB is not distributed evenly within the landscape, but concentrated in woodlands, areas which unsurprisingly correlate strongly with the distribution of timber harvest. Failing to account for this localisation of disturbance within distinct ecosystems when aggregating to coarse spatial domains therefore results in a significant negative bias in the simulated C losses from the live C pools. In contrast, in the stratified framework, disturbance fluxes were insensitive to resolution, giving a consistent estimate of the carbon balance across the resolution ranges considered.

4.2 Impact of heterogeneity on model parameters and ecosystem response to future climate *H3*

We found that by stratifying the landscape prior to MDF, the variability in ecosystem function exhibited between ecosystems, manifest in their retrieved parameters, was retained across the range of spatial resolutions considered (Figure 8). Conversely this ecological information is shredded if data are aggregated without considering *a priori* the underlying distribution of land-use and land-cover. This shredding of ecological information is exemplified by our attempts to constrain mean residence times in the long-lived wood and soil pools, which are critical for understanding the potential carbon sink of terrestrial ecosystems (Luo et al., 2015; Smallman et al., 2021): in the baseline experiment, where data were the relationship between stocks and land-use was ignored, the longer residence times specific to woodland ecosystems (MRT_{Wood}; Figure 8) and heathland areas supporting C-rich peat deposits (MRT_{Soil}; Figure A7) were not well-represented by the posterior parameter estimates. Critically, this misrepresentation of ecosystem function was exacerbated at the coarser grid resolutions commonly employed in large scale MDF applications. Prior research has demonstrated that CARDAMOM can retrieve trait differences across biomes (Bloom et al., 2016; Smallman et al., 2021). We demonstrate that CARDAMOM can also retrieve ecosystem-specific traits in mosaic landscapes when the assimilated data is stratified based on prior knowledge of land-use. Stratification therefore ensures that the ecological fidelity of the model ensembles is maintained when aggregating to the coarser spatial domains for which large-scale forecasts can be generated. Given the relative importance of parameter uncertainty on future trajectories of the terrestrial C cycle (Smallman et al., 2021), stratification also presents significant opportunities to take advantage of the Bayesian framework embedded within CARDAMOM by taking advantage of the prior information on land-cover and land-use, for example using global trait databases (e.g. Kattge et al., 2020), to inform parameter prior estimates.

In our prognostic experiments, we explored the potential future response of our calibrated ecosystem model under different levels of climate forcing (Figure 9), highlighting divergent future C accumulation, which we suggest is largely related to differences in MRT between strata (Luo et al., 2015; Smallman et al., 2021). Ecosystem responses to climate and disturbance are modulated by their functional traits (e.g. Greenwood et al., 2017). In our experiments, the differences in simulated future C accumulation reflect differences in the calibrated parameters between strata, characterising the differing functional traits of the represented ecosystems. In our study landscape, differences in forecast C accumulation between strata counteracted to give overall trajectories that were similar to the baseline experiments. However, in practical terms, understanding the differential dynamics between ecosystems is likely to be important for the utility of forecasts for land managers and policy makers operating in heterogeneous landscapes (Smallman et al., 2022). Ecosystem-specific calibrations may also be particularly important for understanding the future trajectories of heterogeneous landscapes in regions close to thresholds of abrupt ecological change (Turner et al., 2020).

4.3 Limitations of current approach and future work

A key advantage of spatially distributed MDF approaches, such as CARDAMOM, is that model parameters are calibrated locally based on the available observations of the ecosystem. Nevertheless, deficiencies in the assimilated observation streams and/or their uncertainty estimates will propagate to affect the calibrated parameters (Zhao et al., 2020). Excessive seasonality in needle-leaf forests has been documented previously in boreal forests (e.g. Heiskanen et al., 2012). In our stratified experiments, it was notable that the assimilated LAI time series for all strata, including coniferous woodlands, exhibited strong seasonality, and consequently the calibrated leaf lifespans coniferous woodlands were indistinguishable from those calibrated for deciduous systems (Figure A6). Secondly, repeated estimates of AGB have been demonstrated to significantly reduce uncertainties by helping to constrain the residence times of C within the ecosystem (Smallman et al., 2017). In our experiments we used observations of AGB for two time points, but spaced only a year apart (i.e. 2017,2018; Santoro et al., 2021). The spacing of these estimates is short compared to the residence times of aboveground C in forests and woodlands (Figure A6. Improved constraints on the residence times and trajectories of the long-lived C pools may potentially be possible with additional, comparable, AGB estimates with greater temporal separation, and reduced uncertainty.

Our stratified CARDAMOM framework provides flexibility to improve the process representation for important fluxes in arable and pasture landscapes by incorporating ecosystem-specific sub-models that account for important C fluxes associated with crop harvest (Revill et al., 2021), grazing and mowing Myrgiotis et al. (2021). Top-down estimates of the terrestrial C balance for the UK suggest a pulse of emissions late in the summer, coincident with the main harvest season that is not observed in bottom-up CARDAMOM simulations based on a similarly simple DALEC model structure employed here (White et al., 2019). Stratification by itself does not resolve this discrepancy. In contrast, we find that the temporal patterns of simulated R_{eco} were indistinguishable between the baseline and stratified experiments across all resolutions (Figure A5). However, stratification provides the basic framework within which to add ecosystem sub-models that explicitly model land-use in agricultural settings (Revill et al., 2021; Myrgiotis et al., 2021) thereby providing an avenue to resolving this discrepancy. This remains a target for future work.

Finally, in our stratified framework, CARDAMOM retrieves model ensembles that best represent the observations within the calibration period based on a specified set of land-use classes extracted from the LCM2015 land-cover map (Rowland et al., 2017). However, land-use is not static. The terrestrial biosphere is a dynamic environment, with environmental and anthropogenic change driving temporal shifts in land-use and cover that we need to be able to account for. The UK has a relatively stable land-use configuration; however, landscapes with significant land-use change present a challenge, as the current implementation of the MCMC within CARDAMOM assimilates the full time series of calibration data to calibrate a set of time-invariant parameters. This carries the advantage of constraining the model parameters using as many observations as possible. Other frameworks employing sequential approaches to data assimilation provide scope for parameters to shift through time as new data are assimilated, such as four-dimensional ensemble variational data assimilation (e.g. Pinnington et al., 2020) and particle filters (e.g. Montzka et al., 2011). Adapting CARDAMOM to deal with land-use change is therefore an important target for future development, although such efforts will be reliant on reliable, frequently repeated information on shifting land-use (e.g. Souza et al., 2020).

4.4 Broader implications for constraining the terrestrial C balance

Landscapes frequently host a mosaic of contrasting land-uses and land-cover types, with heterogeneity exhibited at scales of 10s to 100s of metres. Modelling the land surface carbon balance across large scales requires the aggregation of ecosystems to resolutions that are computationally feasible. Adequately handling this scaling challenge is critical to avoid the introduction of biases into the simulated C balance, particularly when dealing with heterogeneous landscapes. Our results indicate that localised, ecosystem-specific fluxes, such as those related to disturbance, are particularly sensitive to resolution in mixed-use landscapes (Figure 6). Disturbances through logging, clearance for agriculture, or fire are important determinants of the carbon balance in many forest landscapes, globally (Gatti et al., 2021; Harris et al., 2021). Deforestation and degradation are often concentrated at forest edges and coincident with forest fragmentation (Brink et al., 2017; Matricardi et al., 2020). These hot-spots of environmental change are therefore characterised by the juxtaposition of contrasting ecosystems in fragmented landscapes, highlighting the importance of stratification within MDF frameworks constraining the carbon balance in forested regions globally.

The C pools within the comparatively simple C-cycle model structure of DALEC (Bloom and Williams, 2015; Smallman et al., 2017) map well onto the IPCC-recommended pools for reporting greenhouse gas emissions in the Agriculture, Forestry and Other Land-Use (AFOLU) sector (IPCC, 2019). By combining multiple, spatially explicit observation streams with eco-

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

logical theory embedded in models, MDF approaches such as CARDAMOM ensure conservation of mass balance, ecological "common sense" in the retrieved parameter sets and simulated temporal dynamics, as well as transparent propagation of uncertainties when quantifying ecosystem C fluxes (Smallman et al., 2021). In addition, sub-pixel stratification within CARDAMOM enables sector-specific interrogation of the terrestrial C budget even in mosaic landscapes. For example, in the stratified experiment it is evident that C losses driven by timber harvest over the simulation period result in coniferous woodlands tending to lose C over time. The magnitude and spatial extent of these losses is only readily apparent when we stratified CARDAMOM based on land-use. CARDAMOM, facilitated by land-use stratification, therefore has clear potential to feed into Tier 3 greenhouse gas emissions reporting to the UNFCCC (IPCC, 2019).

5 Conclusions

Quantifying the current and future terrestrial C balance is essential to understanding the stability of terrestrial ecosystems facing rapid environmental change, and to support robust national reporting of land-based CO_2 emissions. Bayesian MDF frameworks like CARDAMOM integrate the ecological knowledge embedded within models with constraints provided from a range of observation sources and their associated uncertainties, thus providing self-consistent, mass-balanced estimates of systemic C cycling (Luo et al., 2011; Bloom et al., 2016; Smallman et al., 2017). However, when applying MDF across large regions, we are confronted by the challenge of capturing the innate complexity of terrestrial ecosystem with spatial and temporal resolutions that are computationally feasible. This challenge is particularly severe in heterogeneous landscapes with a mosaic of land-uses. Failure to account for sub-pixel ecosystem heterogeneity within MDF inversions leads to bias in flux estimates and shredding of the ecological information embedded within the calibrated model ensembles. We explored the carbon balance for a region of ~30,000 km^2 in the UK using a range of spatial resolutions (0.05° - 1.0°). In our baseline experiment (ignoring sub-pixel heterogeneity), disturbance fluxes in particular exhibited a resolution dependent negative bias that was exacerbated both at coarser grid resolutions and as landscape fragmentation increased. Accounting for fine-scale structure of land-use through through stratification resolved this scale dependence, and yielded higher disturbance fluxes. Stratification also enabled CARDAMOM to retrieve parameter ensembles that preserved the differences in ecological function between different land-uses, thus maintaining ecological fidelity at coarse resolutions.

Stratification within CARDAMOM therefore provides three key benefits: (i) stratification improves flux estimates; these appear to be relatively insensitive to resolution, facilitating scaling of CARDAMOM applications across larger spatial domains; (ii) stratification provides transparency for sector-level estimates of the terrestrial carbon balance that could be integrated into Tier 3 national emissions reporting frameworks; and (iii) by separately analysing distinct ecosystems within fragmented landscapes, the ecological fidelity of the calibrated model parameters is enhanced, enabling more robust ecological forecasting and raising the prospect of mapping spatial variations of ecosystem functional traits based on a diverse range of EO data. Future work will build on this stratification framework to build in more detailed process representation to better account for C fluxes in managed arable (Revill et al., 2021) and pasture landscapes (Myrgiotis et al., 2021). Finally, landscape fragmentation and disturbance, whether driven by logging, agriculture or fire, are important determinants of the carbon balance globally.

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Therefore, while the focus of this study is a temperate landscape within the UK, these results carry broader significance for the application of MDF frameworks to constrain the terrestrial C-balance at regional and national scales.

Code and data availability. The driving data, and selected carbon cycle outputs have been archived (Milodowski et al., 2022) at Edinburgh DataShare: https://datashare.ed.ac.uk/handle/10283/4491.
The code to regrid the EO datasets is archived at https://github.com/GCEL/DAREUK_EOregrid.
The code to drive the model and analyse the model output and generate the paper figures is archived at:
https://github.com/GCEL/DAREUK_scale_variance_paper.
The specific version of the CARDAMOM code used for the analysis presented here is archived at
https://github.com/GCEL/DAREUK_CARDAMOM_scale_variance_paper.
Registration to the github repositories is provided on request to either T. L. Smallman or M. Williams.

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

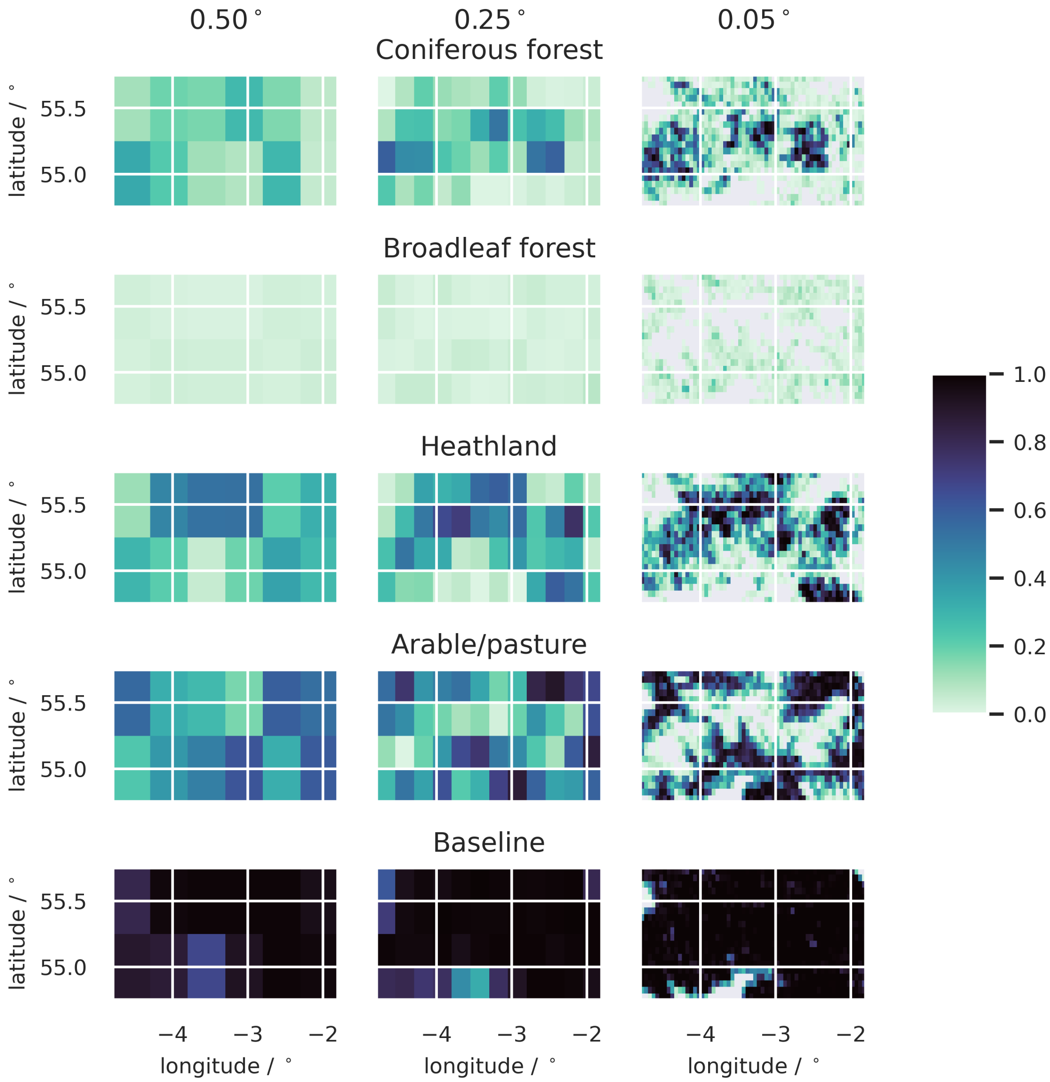

Figure A1. Pixel fractions occupied by the different land-cover strata, presented for the 0.50°, 0.25° and 0.05° model domains. The 1.00° domain is not shown. At 0.05° resolution, Recognisable landscape units are picked out with much more clarity.

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Biogeosciences Discussions
Open Access

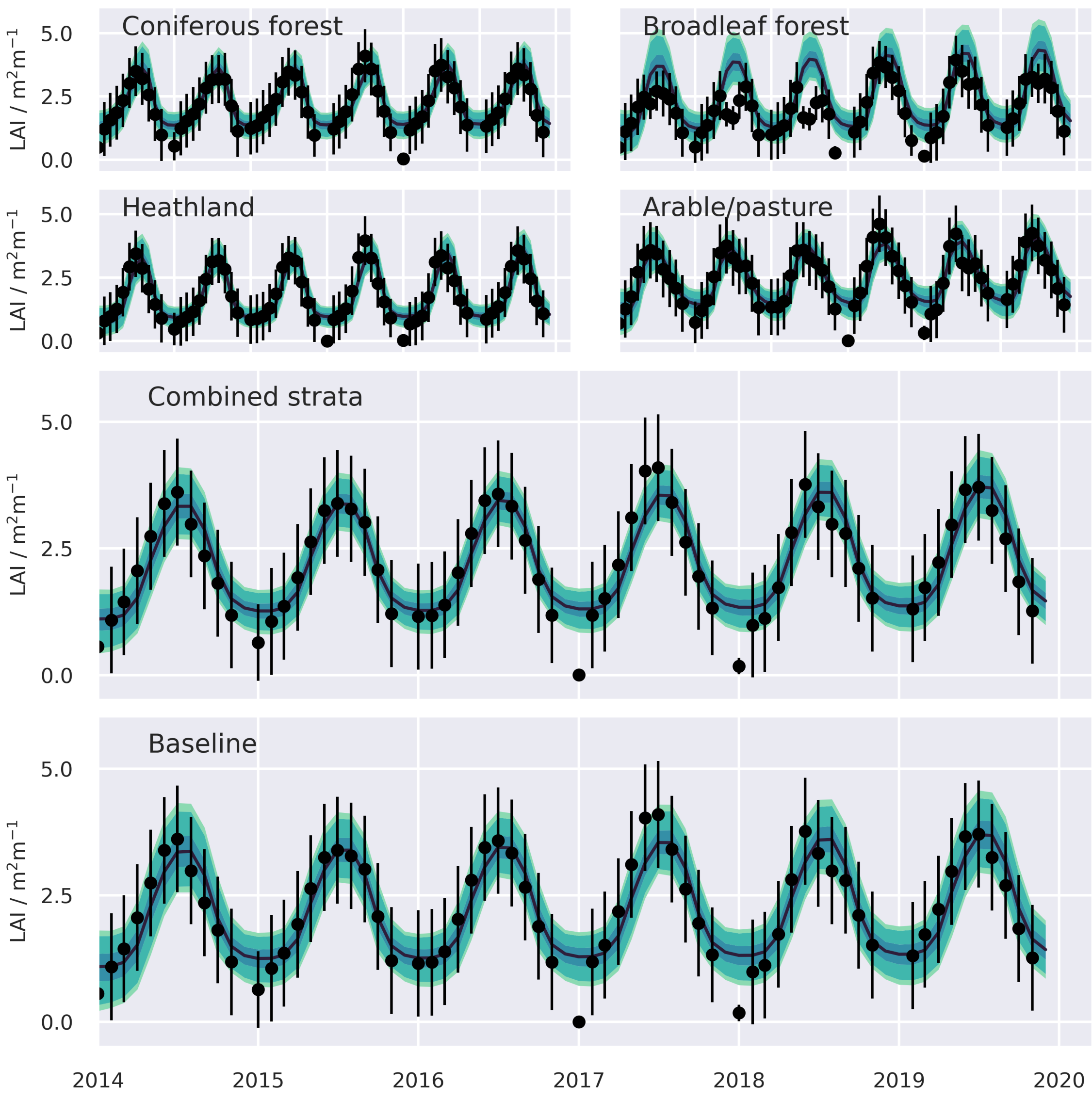

Figure A2. Simulated LAI time series and assimilated LAI observations, presented with uncertainty, aggregated from the 0.05° domain. Shaded bands represent the 50%, 90% and 95% confidence intervals, assuming full spatial correlation of uncertainties across the domain. Calibration performance statistics for all resolution domains are summarised in Table 1 and for individual strata in Table A1

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Biogeosciences
Discussions

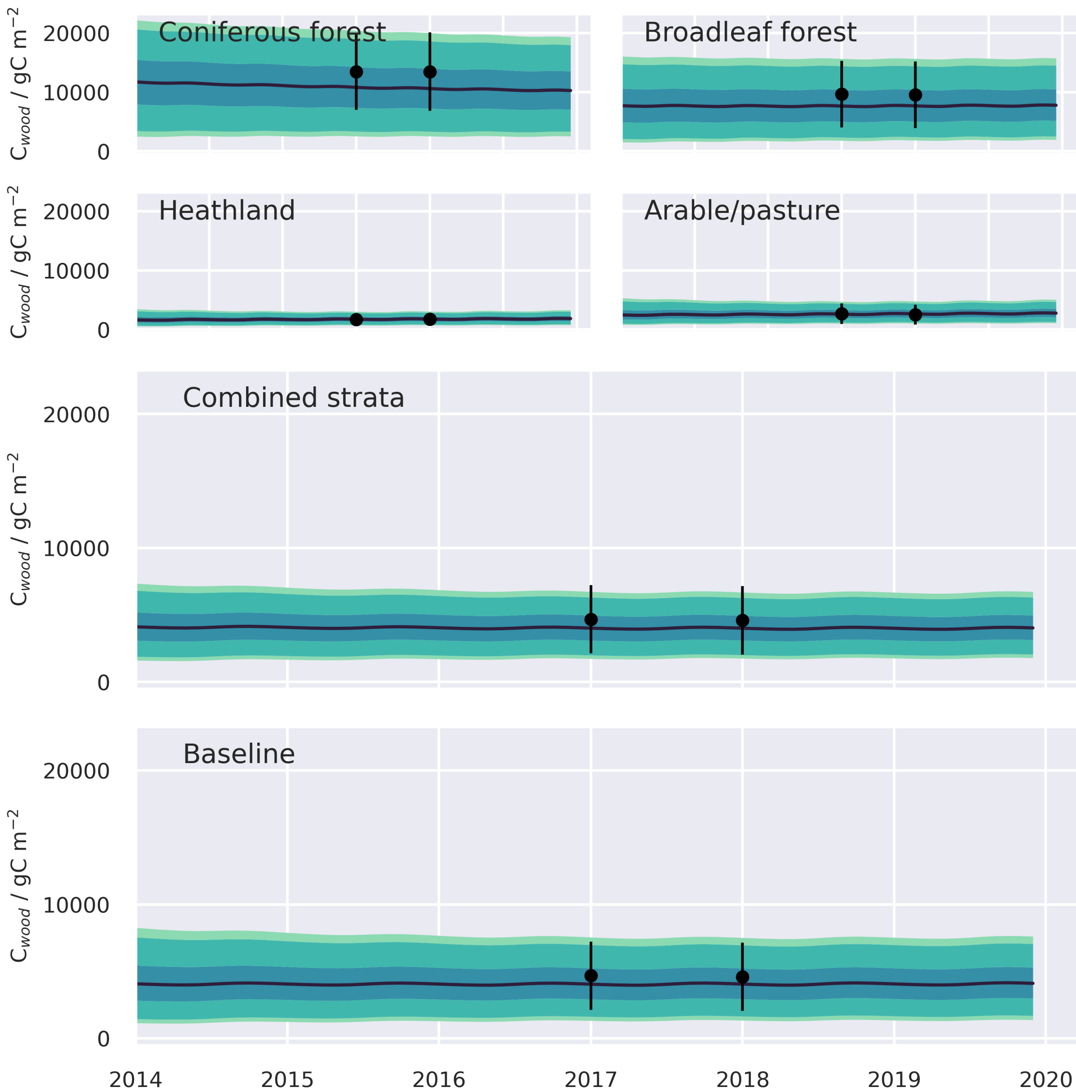

Figure A3. Simulated C_{Wood} time series and assimilated observations, presented with uncertainty, aggregated from the 0.05° domain. Shaded bands represent the 50%, 90% and 95% confidence intervals, assuming full spatial correlation of uncertainties across the domain. Calibration performance statistics for all resolution domains are summarised in Table 1 and for individual strata in Table A1.

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

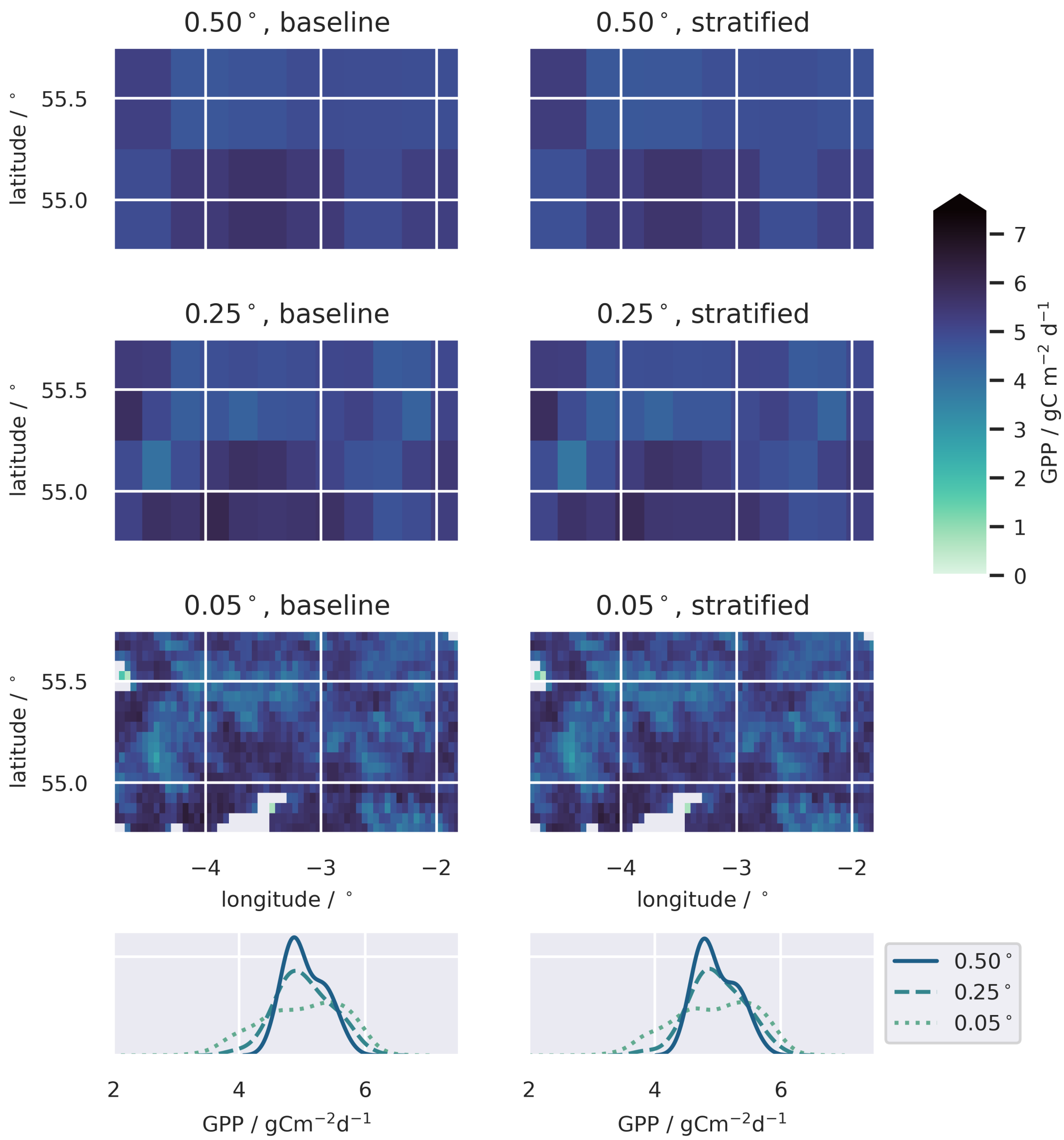

Figure A4. Temporally averaged GPP for the baseline and stratified ensembles, for domains of 0.50°, 0.25° and 0.05° resolution. Note the 1.00° domain is not shown. The bottom row shows the distributions of median GPP across the domain for the baseline and stratified ensembles, and demonstrates the diminishing variability in simulated GPP when data are aggregated to spatial domains with coarser resolution.

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Biogeosciences
Discussions
Open Access
EGU

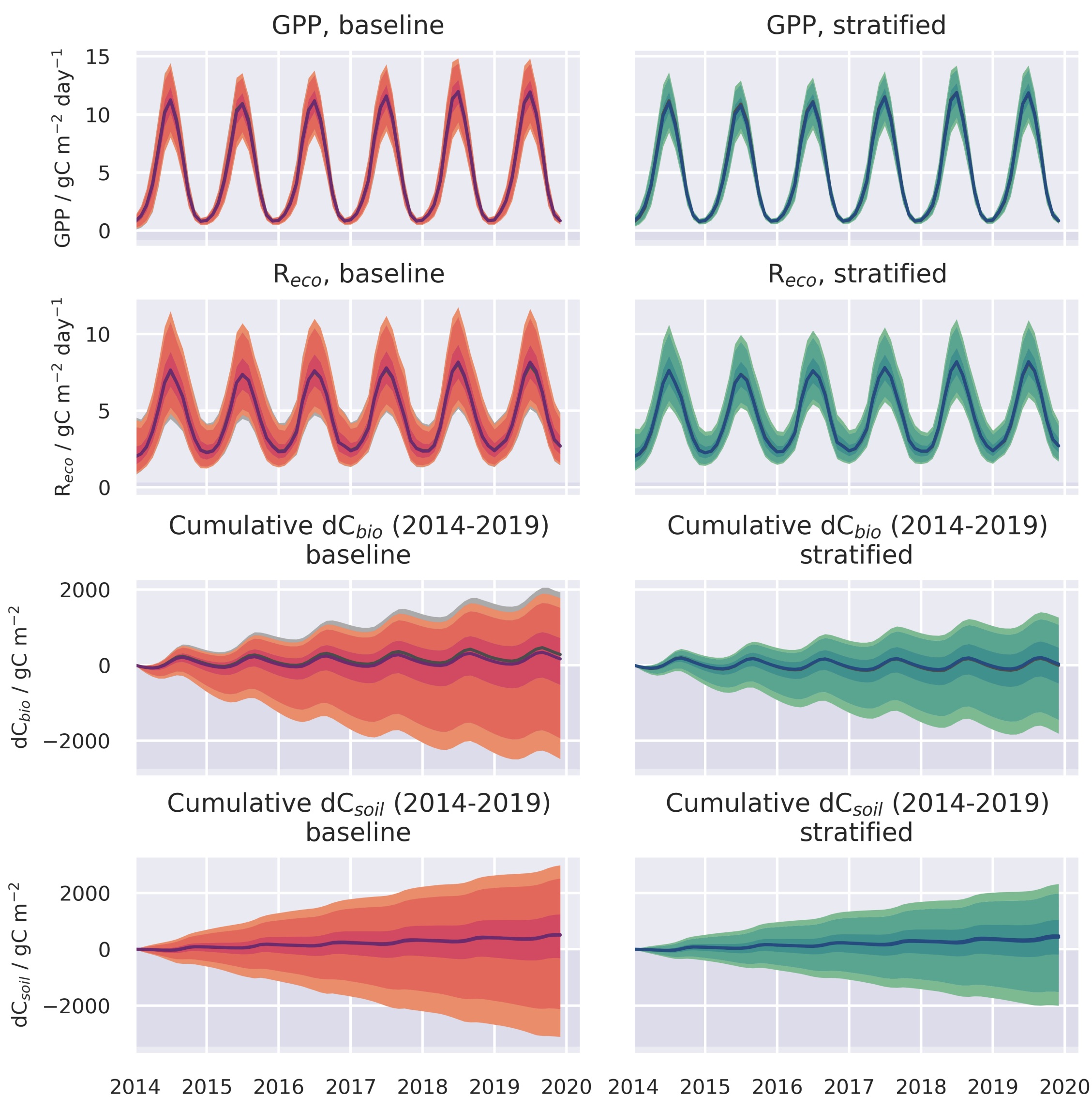

Figure A5. Spatially aggregated time series for GPP and ecosystem respiration ($R_{eco} = R_a + R_{het}$),and the cumulative change in carbon stocks in the live (dC_{bio}) and soil (dC_{soil}) pools, shown for the baseline and stratified ensembles for the 0.05° and 1.00° resolution domains. The median estimates are plotted with the shaded regions representing the 50%, 90% and 95% confidence intervals. The 1.00° simulation results are plotted in grey-scale, but for the most part, the ensemble ranges overlap.

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

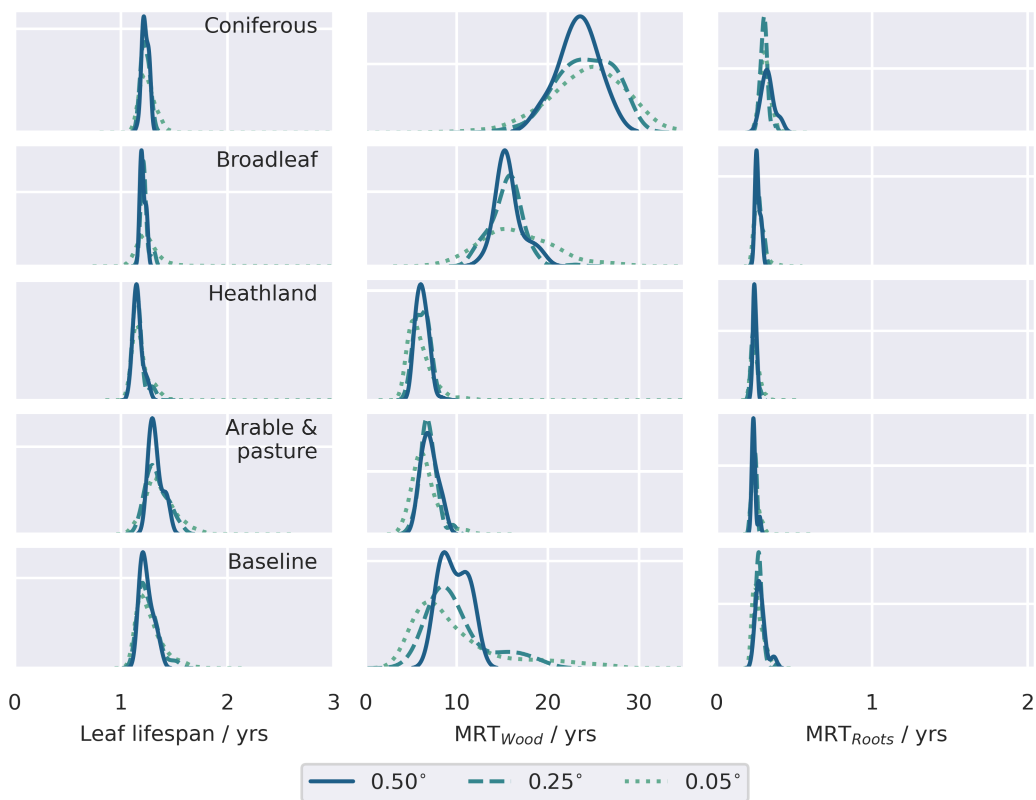

Figure A6. Comparison of the retrieved residence times for the live carbon pools, for the individual strata and the baseline retrieval for the different resolution domains. Distributions represent the pixel-level median parameter estimates weighted by the pixel-fraction estimated associated with each stratum.

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

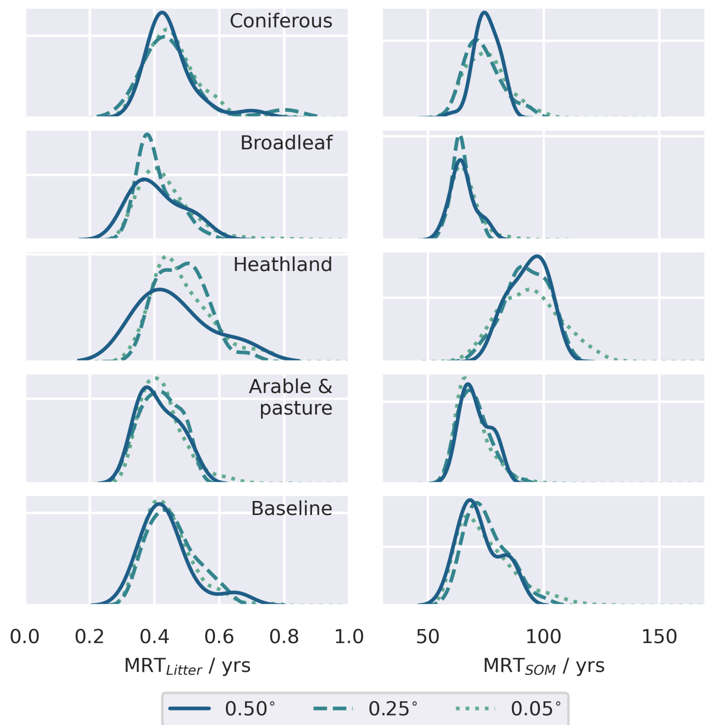

Figure A7. Comparison of the retrieved residence times for the dead organic matter carbon pools, for the individual strata and the baseline retrieval for the different resolution domains. Distributions represent the pixel-level median parameter estimates weighted by the pixel-fraction estimated associated with each stratum.

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Biogeosciences Discussions
Open Access
EGU

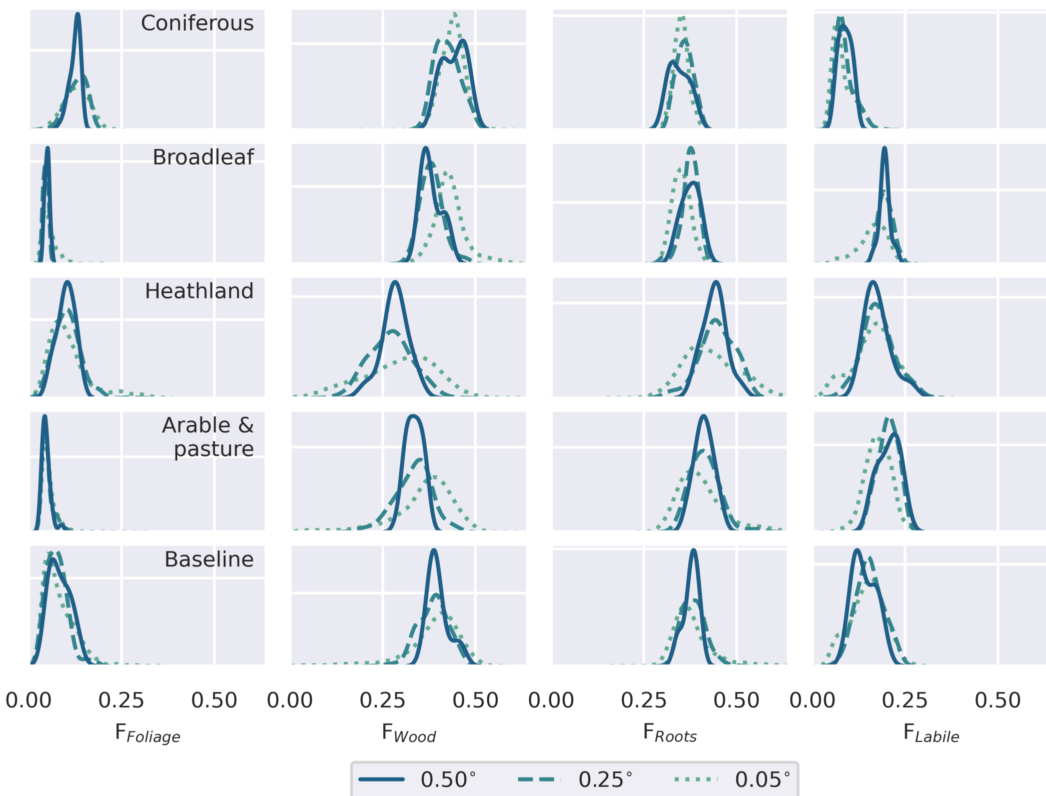

Figure A8. Comparison of the retrieved allocation fractions the partitioning of NPP between the live organic carbon pools, for the individual strata and the baseline retrieval for the different resolution domains. Distributions represent the pixel-level median parameter estimates weighted by the pixel-fraction estimated associated with each stratum. Note that labile carbon is subsequently allocated to foliage 7

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Biogeosciences Discussions

Table A1. Description of parameters estimated for the DALEC model used in this study (Bloom and Williams, 2015). Each parameter is given a name, unit, description. Note: Gross Primary Productivity = GPP, Autotrophic respiration = R_a, Autotrophic maintenance respiration = R_m, heterotrophic respiration = Rh. Litter is assumed to be the combined foliage and fine root litter pools. Note that GPP allocation fractions are applied sequentially such that GPP allocation to C_{wood} = GPP - (GPP·R_a:GPP) - (GPP·GPP_{lab}) - (GPP·GPP_{root}).

Name	units	Description
R_a:GPP	fraction	Fraction of GPP allocated to R_a
GPP_{fol}	fraction	Fraction of GPP allocated to foliage
GPP_{lab}	fraction	Fraction of GPP allocated to labile
GPP_{root}	fraction	Fraction of GPP allocated to fine root
Leaf lifespan	years	Maximum natural leaf lifespan
Leaf growth day	day of year	Julian day on which max labile turnover to foliage as defined by the phenology model
Leaf growth period	days	Standard deviation defining the period over which labile turnover to foliage occurs
Leaf fall day	day of year	Julian day on which max foliar turnover to litter as defined by the phenology model
Leaf fall period	days	Standard deviation defining the period over which foliar turnover to litter occurs
Wood turnover	day^{-1}	Fraction of wood loss per day
Fine root turnover	day^{-1}	Fraction of fine root loss per day
Litter decomposition	day^{-1} at 0^oC	Fraction of fine root loss per day
Litter mineralisation	day^{-1} at 0^oC	Baseline litter turnover to R_{het}
Soil mineralisation	day^{-1} at 0^oC	Baseline soil turnover to R_{het}
R_{het} coefficient	-	Exponential temperature response coefficient for R_{het}
LMA	$g(C)m^{-2}$	Leaf mass per unit leaf area
Ceff	$g(C)m^{-2}day^{-1}$	Potential photosynthetic activity per unit leaf area
Initial labile	$g(C)m^{-2}$	Size of the labile C pool at time step 1
Initial foliage	$g(C)m^{-2}$	Size of the foliar C pool at time step 1
Initial fine root	$g(C)m^{-2}$	Size of the fine root C pool at time step 1
Initial wood	$g(C)m^{-2}$	Size of the wood C pool at time step 1
Initial litter	$g(C)m^{-2}$	Size of the litter C pool at time step 1
Initial soil	$g(C)m^{-2}$	Size of the soil C pool at time step 1

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Table A2. Summary of calibration performance for the individual strata in the stratified CARDAMOM experiment, aggregated across the domains. σ represents the uncertainty of the assimilated observation data, thus RMSE / σ provides the ratio of the RMSE relative to the uncertainty attached to the observation constraint.

Variable	Version	Metric	1.00°	0.50°	0.25°	0.05°
C_{Wood}	Coniferous forest	RMSE / gCm^{-2}	2764 (20.6 %)	3007 (22.5 %)	2856 (21.2 %)	2785 (20.6 %)
C_{Wood}	Coniferous forest	RMSE / σ	0.42	0.46	0.43	0.42
C_{Wood}	Coniferous forest	Bias / gCm^{-2}	-2762 (-20.6 %)	-3003 (-22.5 %)	-2849 (-21.2 %)	-2758 (-20.4 %)
C_{Wood}	Coniferous forest	Bias / σ	-0.42	-0.46	-0.43	-0.41
C_{Wood}	Coniferous forest	Median gCm^{-2}	10807	10572	10730	10887
C_{Wood}	Broadleaf forest	RMSE / gCm^{-2}	2407 (26.5 %)	1948 (21.1 %)	1987 (21.6 %)	1969 (20.5 %)
C_{Wood}	Broadleaf forest	RMSE / σ	0.45	0.36	0.37	0.35
C_{Wood}	Broadleaf forest	Bias / gCm^{-2}	-2404 (-26.5 %)	-1945 (-21.1 %)	-1979 (-21.5 %)	-1948 (-20.2 %)
C_{Wood}	Broadleaf forest	Bias / σ	-0.45	-0.36	-0.37	-0.34
C_{Wood}	Broadleaf forest	Median gCm^{-2}	6712	7173	7220	7672
C_{Wood}	Non-forest	RMSE / gCm^{-2}	25 (1.5 %)	70 (4.3 %)	102 (6.3 %)	180 (11.3 %)
C_{Wood}	Non-forest	RMSE / σ	0.02	0.07	0.10	0.18
C_{Wood}	Non-forest	Bias / gCm^{-2}	13 (0.7 %)	24 (2.2 %)	6 (1.8 %)	19 (5.2 %)
C_{Wood}	Non-forest	Bias / σ	0.01	0.03	0.03	0.07
C_{Wood}	Non-forest	Median gCm^{-2}	1727	1742	1722	1720
C_{Wood}	Arable/pasture	RMSE / gCm^{-2}	191 (7.4 %)	174 (6.7 %)	199 (8.0 %)	255 (10.4 %)
C_{Wood}	Arable/pasture	RMSE / σ	0.11	0.10	0.12	0.16
C_{Wood}	Arable/pasture	Bias / gCm^{-2}	45 (2.2 %)	24 (1.4 %)	14 (1.7 %)	11 (3.0 %)
C_{Wood}	Arable/pasture	Bias / σ	0.03	0.02	0.02	0.04
C_{Wood}	Arable/pasture	Median gCm^{-2}	2615	2593	2590	2596
LAI	Coniferous forest	RMSE / m^2m^{-2}	0.38 (16.1 %)	0.38 (16.2 %)	0.39 (16.6 %)	0.42 (17.9 %)
LAI	Coniferous forest	RMSE / σ	0.36	0.36	0.37	0.39
LAI	Coniferous forest	Bias / m^2m^{-2}	-0.00 (-0.1 %)	-0.00 (-0.1 %)	-0.00 (-0.1 %)	-0.01 (-0.4 %)
LAI	Coniferous forest	Bias / σ	-0.00	-0.00	-0.00	-0.01
LAI	Coniferous forest	Median m^2m^{-2}	2.22	2.21	2.21	2.21
LAI	Broadleaf forest	RMSE / m^2m^{-2}	0.47 (17.6 %)	0.48 (17.9 %)	0.49 (18.4 %)	0.53 (22.0 %)
LAI	Broadleaf forest	RMSE / σ	0.42	0.43	0.43	0.48
LAI	Broadleaf forest	Bias / m^2m^{-2}	-0.03 (-1.1 %)	-0.04 (-1.3 %)	-0.03 (-1.3 %)	-0.05 (-2.2 %)
LAI	Broadleaf forest	Bias / σ	-0.03	-0.03	-0.03	-0.05
LAI	Broadleaf forest	Median m^2m^{-2}	2.51	2.50	2.51	2.40
LAI	Non-forest	RMSE / m^2m^{-2}	0.36 (20.3 %)	0.37 (20.3 %)	0.39 (21.0 %)	0.42 (22.3 %)
LAI	Non-forest	RMSE / σ	0.40	0.40	0.41	0.44
LAI	Non-forest	Bias / m^2m^{-2}	-0.01 (-0.7 %)	-0.02 (-0.9 %)	-0.01 (-0.7 %)	-0.02 (-1.2 %)
LAI	Non-forest	Bias / σ	-0.01	-0.02	-0.01	-0.02
LAI	Non-forest	Median m^2m^{-2}	1.69	1.73	1.75	1.77
LAI	Arable/pasture	RMSE / m^2m^{-2}	0.41 (16.2 %)	0.43 (16.8 %)	0.45 (17.4 %)	0.49 (19.0 %)
LAI	Arable/pasture	RMSE / σ	0.38	0.39	0.40	0.43
LAI	Arable/pasture	Bias / m^2m^{-2}	-0.03 (-1.1 %)	-0.04 (-1.4 %)	-0.04 (-1.5 %)	-0.04 (-1.5 %)
LAI	Arable/pasture	Bias / σ	-0.03	-0.03	-0.03	-0.04
LAI	Arable/pasture	Median m^2m^{-2}	2.44	2.43	2.43	2.44

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Biogeosciences Discussions

Author contributions. DTM, TLS and MW created the experimental design. DTM ran CARDAMOM and analysed the outputs. DTM led the writing with inputs from TLS and MW.

Competing interests. The authors have no competing interests.

Acknowledgements. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission. D.T. Milodowski and M. Williams were funded primarily by the NERC DARE-UK project (NE/S003819/1). T. L Smallman and M. Williams were additionally supported by the UK's National Centre for Earth Observation. M Williams also received funding from the Royal Society. This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF) (http://www.ecdf.ed.ac.uk/). LAI information was generated by the Global Land Service of Copernicus, the Earth Observation programme of the European Commission. Aboveground biomass maps for 2017 and 2018 were provided by the European Space Agency through their Climate Change Initiative (CCI) Biomass project. Soilgrids2 soil organic carbon maps were made available by the International Soil Reference and Information Centre (ISRIC). The University of Edinburgh CARDAMOM framework is available on Github (https://github.com/GCEL/CARDAMOM) with permissions provided on request to the corresponding author. The Jet Propulsion Laboratory CARDAMOM framework is also available on Github (github.com/CARDAMOM-framework/CARDAMOM_2.1.6c - contact abloom@jpl.nasa.gov for access).

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Biogeosciences Discussions

References

Beven, K.: A manifesto for the equifinality thesis, Journal of Hydrology, 320, 18–36, https://doi.org/10.1016/j.jhydrol.2005.07.007, 2006.

Bloom, A. A., Exbrayat, J.-F., van der Velde, I. R., Feng, L., and Williams, M.: The decadal state of the terrestrial carbon cycle: Global retrievals of terrestrial carbon allocation, pools, and residence times, P. Natl. Acad. Sci. USA, 113, 1285-1290, https://doi.org/10.1073/pnas.1515160113, 2016.

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

Bonan, G. B. and Doney, S. C.: Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models, 359, eaam8328, https://doi.org/10.1126/science.aam8328, 2018.

Brinck, K., Fischer, R., Groeneveld, J., Lehmann, S., Dantas De Paula, M., Pütz, S., Sexton, J. O., Song, D., and Huth, A.: High resolution analysis of tropical forest fragmentation and its impact on the global carbon cycle, 8, 14855, https://doi.org/10.1038/ncomms14855, 2017.

Butler, E. E., Datta, A., Flores-Moreno, H., Chen, M., Wythers, K. R., Fazayeli, F., Banerjee, A., Atkin, O. K., Kattge, J., Amiaud, B., Blonder, B., Boenisch, G., Bond-Lamberty, B., Brown, K. A., Byun, C., Campetella, G., Cerabolini, B. E. L., Cornelissen, J. H. C., Craine, J. M., Craven, D., de Vries, F. T., Díaz, S., Domingues, T. F., Forey, E., González-Melo, A., Gross, N., Han, W., Hattingh, W. N., Hickler, T., Jansen, S., Kramer, K., Kraft, N. J. B., Kurokawa, H., Laughlin, D. C., Meir, P., Minden, V., Niinemets, Ü., Onoda, Y., Peñuelas, J., Read, Q., Sack, L., Schamp, B., Soudzilovskaia, N. A., Spasojevic, M. J., Sosinski, E., Thornton, P. E., Valladares, F., van Bodegom, P. M., Williams, M., Wirth, C., and Reich, P. B.: Mapping local and global variability in plant trait distributions, 114, E10937–E10946, https://doi.org/10.1073/pnas.1708984114, 2017.

Chaplin-Kramer, R., Ramler, I., Sharp, R., Haddad, N. M., Gerber, J. S., West, P. C., Mandle, L., Engstrom, P., Baccini, A., Sim, S., Mueller, C., and King, H.: Degradation in carbon stocks near tropical forest edges, Nat Commun, 6, 10158, https://doi.org/10.1038/ncomms10158, 2015.

University of East Anglia Climatic Research Unit; Harris, I.C.: CRU JRA v1.1: A forcings dataset of gridded land surface blend of Climatic Research Unit (CRU) and Japanese reanalysis (JRA) data; Jan.1901 - Dec.2017. Centre for Environmental Data Analysis, 25 February 2019. https://doi.org/10.5285/13f3635174794bb98cf8ac4b0ee8f4ed, 2019.

Exbrayat, J-F., Smallman, T. L., Bloom, A. A., Huntley, L. B. and Williams, M.: Inverse Determination of the Influence of Fire on Vegetation Carbon Turnover in the Pantropics. Global Biogeochem. Cy., 32, 1776-1789, https://doi.org/10.1029/2018GB005925, 2018.

Exbrayat, J.-F., Bloom, A. A., Carvalhais, N., Fischer, R., Huth, A., MacBean, N., and Williams, M.: Understanding the Land Carbon Cycle with Space Data: Current Status and Prospects, Surv Geophys, 40, 735–755, https://doi.org/10.1007/s10712-019-09506-2, 2019.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Harper, A. B., Cox, P. M., Friedlingstein, P., Wiltshire, A. J., Jones, C. D., Sitch, S., Mercado, L. M., Groenendijk, M., Robertson, E., Kattge, J., Bönisch, G., Atkin, O. K., Bahn, M., Cornelissen, J., Niinemets, Ü., Onipchenko, V., Peñuelas, J., Poorter, L., Reich, P. B., Soudzilovskaia, N. A., and Bodegom, P. van: Improved representation of plant functional types and physiology in the Joint UK Land Environment Simulator (JULES v4.2) using plant trait information, 9, 2415–2440, https://doi.org/10.5194/gmd-9-2415-2016, 2016.

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

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

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

Fisher, R. A. and Koven, C. D.: Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems, 12, e2018MS001453, https://doi.org/10.1029/2018MS001453, 2020.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S., Aragão, L. E. O. C., Arneth, A., Arora, V., Bates, N. R., Becker, M., Benoit-Cattin, A., Bittig, H. C., Bopp, L., Bultan, S., Chandra, N., Chevallier, F., Chini, L. P., Evans, W., Florentie, L., Forster, P. M., Gasser, T., Gehlen, M., Gilfillan, D., Gkritzalis, T., Gregor, L., Gruber, N., Harris, I., Hartung, K., Haverd, V., Houghton, R. A., Ilyina, T., Jain, A. K., Joetzjer, E., Kadono, K., Kato, E., Kitidis, V., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Liu, Z., Lombardozzi, D., Marland, G., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pierrot, D., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Smith, A. J. P., Sutton, A. J., Tanhua, T., Tans, P. P., Tian, H., Tilbrook, B., van der Werf, G., Vuichard, N., Walker, A. P., Wanninkhof, R., Watson, A. J., Willis, D., Wiltshire, A. J., Yuan, W., Yue, X., and Zaehle, S.: Global Carbon Budget 2020, 12, 3269–3340, https://doi.org/10.5194/essd-12-3269-2020, 2020.

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., Woodward, F. I.: Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO_2. P. Natl. Acad. Sci. USA, 111 (9) 3280-3285, https://doi.org/10.1073/pnas.1222477110, 2014.

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

Gatti, L. V., Basso, L. S., Miller, J. B., Gloor, M., Gatti Domingues, L., Cassol, H. L. G., Tejada, G., Aragão, L. E. O. C., Nobre, C., Peters, W., Marani, L., Arai, E., Sanches, A. H., Corrêa, S. M., Anderson, L., Von Randow, C., Correia, C. S. C., Crispim, S. P., and Neves, R. A. L.: Amazonia as a carbon source linked to deforestation and climate change, 595, 388–393, https://doi.org/10.1038/s41586-021-03629-6, 2021.

Giglio, L., Boschetti, L., Roy, D. P., Humber, M. L. and Justice, C. O.: The Collection 6 MODIS burned area mapping algorithm and product, Remote Sens. Environ., 217, 72-85, https://doi.org/10.1016/j.rse.2018.08.005, 2018.

Grassi, G., House, J., Dentener, F., Federici, S., den Elzen, M., and Penman, J.: The key role of forests in meeting climate targets requires science for credible mitigation, Nature Clim Change, 7, 220–226, https://doi.org/10.1038/nclimate3227, 2017.

Greenwood, S., Ruiz-Benito, P., Martínez-Vilalta, J., Lloret, F., Kitzberger, T., Allen, C. D., Fensham, R., Laughlin, D. C., Kattge, J., Bönisch, G., Kraft, N. J. B., and Jump, A. S.: Tree mortality across biomes is promoted by drought intensity, lower wood density and higher specific leaf area, 20, 539–553, https://doi.org/10.1111/ele.12748, 2017.

Haario, H., Saksman, E. and Tamminen, J.: An adaptive Metropolis algorithm. *Bernoulli*, 7 (2), 223-242, 2001.

Harris, N. L., Gibbs, D. A., Baccini, A., Birdsey, R. A., de Bruin, S., Farina, M., Fatoyinbo, L., Hansen, M. C., Herold, M., Houghton, R. A., Potapov, P. V., Suarez, D. R., Roman-Cuesta, R. M., Saatchi, S. S., Slay, C. M., Turubanova, S. A., and Tyukavina, A.: Global maps of twenty-first century forest carbon fluxes, Nat. Clim. Chang., 11, 234–240, https://doi.org/10.1038/s41558-020-00976-6, 2021.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., and Townshend, J. R. G.: High-Resolution Global Maps of 21st-Century Forest Cover Change, Science, 342, 850–853, https://doi.org/10.1126/science.1244693, 2013.

Huntingford, C., Atkin, O. K., Martinez-de la Torre, A., Mercado, L. M., Heskel, M. A., Harper, A. B., Bloomfield, K. J., O'Sullivan, O. S., Reich, P. B., Wythers, K. R., Butler, E. E., Chen, M., Griffin, K. L., Meir, P., Tjoelker, M. G., Turnbull, M. H., Sitch, S., Wiltshire, A., and Malhi, Y.: Implications of improved representations of plant respiration in a changing climate, Nat Commun, 8, 1602, https://doi.org/10.1038/s41467-017-01774-z, 2017.

Hurlbert, M., J. Krishnaswamy, E. Davin, F.X. Johnson, C.F. Mena, J. Morton, S. Myeong, D. Viner, K. Warner, A. Wreford, S. Zakieldeen, Z. Zommers, 2019: Risk Management and Decision making in Relation to Sustainable Development. In: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems [P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D.C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, J. Malley, (eds.)]. In press.

IPCC: Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Calvo Buendia, E., Tanabe, K., Kranjc, A., Baasansuren, J., Fukuda, M., Ngarize S., Osako, A., Pyrozhenko, Y., Shermanau, P. and Federici, S. (eds)., Intergovernmental Panel on Climate Change (IPCC), Switzerland, 2019.

Jenkins, G. J., Perry, M. C., and Prior, M. J.: The climate of the UK and recent trends, Met Office Hadley Centre, Exeter, 2009

Jensen, J. L. W. V.: Sur les fonctions convexes et les inégalités entre les valeurs moyennes, Acta Math., 30, 175–193, https://doi.org/10.1007/BF02418571, 1906.

Jones, C. D. and P. Friedlingstein: Quantifying process-level uncertainty contributions to TCRE and carbon budgets for meeting Paris Agreement climate targets, Environ. Res. Lett., 15, 074019, https://doi.org/10.1088/1748-9326/ab858a, 2020.

Kaminski, T., Rayner, P. J., Voßbeck, M., Scholze, M., and Koffi, E.: Observing the continental-scale carbon balance: assessment of sampling complementarity and redundancy in a terrestrial assimilation system by means of quantitative network design, 12, 7867–7879, https://doi.org/10.5194/acp-12-7867-2012, 2012.

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Tautenhahn, S., Werner, G. D. A., Aakala, T., Abedi, M., Acosta, A. T. R., Adamidis, G. C., Adamson, K., Aiba, M., Albert, C. H., Alcántara, J. M., Alcázar C, C., Aleixo, I., Ali, H., Amiaud, B., Ammer, C., Amoroso, M. M., Anand, M., Anderson, C., Anten, N., Antos, J., Apgaua, D. M. G., Ashman, T.-L., Asmara, D. H., Asner, G. P., Aspinwall, M., Atkin, O., Aubin, I., Baastrup-Spohr, L., Bahalkeh, K., Bahn, M., Baker, T., Baker, W. J., Bakker, J. P., Baldocchi, D., Baltzer, J., Banerjee, A., Baranger, A., Barlow, J., Barneche, D. R., Baruch, Z., Bastianelli, D., Battles, J., Bauerle, W., Bauters, M., Bazzato, E., Beckmann, M., Beeckman, H., Beierkuhnlein, C., Bekker, R., Belfry, G., Belluau, M., Beloiu, M., Benavides, R., Benomar, L., Berdugo-Lattke, M. L., Berenguer, E., Bergamin, R., Bergmann, J., Bergmann Carlucci, M., Berner, L., Bernhardt-Römermann, M., Bigler, C., Bjorkman, A. D., Blackman, C., Blanco, C., Blonder, B., Blumenthal, D., Bocanegra-González, K. T., Boeckx, P., Bohlman, S., Böhning-Gaese, K., Boisvert-Marsh, L., Bond, W., Bond-Lamberty, B., Boom, A., Boonman, C. C. F., Bordin, K., Boughton, E. H., Boukili, V., Bowman, D. M. J. S., Bravo, S., Brendel, M. R., Broadley, M. R., Brown, K. A., Bruelheide, H., Brumnich, F., Bruun, H. H., Bruy, D., Buchanan, S. W., Bucher, S. F., Buchmann, N., Buitenwerf, R., Bunker, D. E., et al.: TRY plant trait database – enhanced coverage and open access, 26, 119–188, https://doi.org/10.1111/gcb.14904, 2020.

Kondo, M., Patra, P. K., Sitch, S., Friedlingstein, P., Poulter, B., Chevallier, F., Ciais, P., Canadell, J. G., Bastos, A., Lauerwald, R., Calle, L., Ichii, K., Anthoni, P., Arneth, A., Haverd, V., Jain, A. K., Kato, E., Kautz, M., Law, R. M., Lienert, S., Lombardozzi, D., Maki, T.,

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Nakamura, T., Peylin, P., Rödenbeck, C., Zhuravlev, R., Saeki, T., Tian, H., Zhu, D., and Ziehn, T.: State of the science in R_{eco}nciling top-down and bottom-up approaches for terrestrial CO2 budget, 26, 1068–1084, https://doi.org/10.1111/gcb.14917, 2020.

Kuppel, S., Peylin, P., Maignan, F., Chevallier, F., Kiely, G., Montagnani, L., and Cescatti, A.: Model–data fusion across ecosystems: from multisite optimizations to global simulations, 7, 2581–2597, https://doi.org/10.5194/gmd-7-2581-2014, 2014.

Levy, P., Clement, R., Cowan, N., Keane, B., Myrgiotis, V., van Oijen, M., Smallman, T. L., Toet, S., and Williams, M.: Challenges in Scaling Up Greenhouse Gas Fluxes: Experience From the UK Greenhouse Gas Emissions and Feedbacks Program, 127, e2021JG006743, https://doi.org/10.1029/2021JG006743, 2022.

López-Blanco, E., Exbrayat, J.-F., Lund, M., Christensen, T. R., Tamstorf, M. P., Slevin, D., Hugelius, G., Bloom, A. A., and Williams, M.: Evaluation of terrestrial pan-Arctic carbon cycling using a data-assimilation system, 10, 233–255, https://doi.org/10.5194/esd-10-233-2019, 2019.

Lindenmayer, D. B. and Fischer, J.: Habitat fragmentation and landscape change: an ecological and conservation synthesis, Island Press, 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, 21, 1429–1442, https://doi.org/10.1890/09-1275.1, 2011.

Luo, Y., Keenan, T. F., and Smith, M.: Predictability of the terrestrial carbon cycle, 21, 1737–1751, https://doi.org/10.1111/gcb.12766, 2015.

Lovenduski, N. S. and Bonan, G. B.: Reducing uncertainty in projections of terrestrial carbon uptake, Environ. Res. Lett., 12, 044020, https://doi.org/10.1088/1748-9326/aa66b8, 2017.

MacBean, N., Maignan, F., Bacour, C., Lewis, P., Peylin, P., Guanter, L., Köhler, P., Gómez-Dans, J., and Disney, M.: Strong constraint on modelled global carbon uptake using solar-induced chlorophyll fluorescence data, Sci Rep, 8, 1973, https://doi.org/10.1038/s41598-018-20024-w, 2018.

Matricardi, E. A. T., Skole, D. L., Costa, O. B., Pedlowski, M. A., Samek, J. H., and Miguel, E. P.: Long-term forest degradation surpasses deforestation in the Brazilian Amazon, 369, 1378–1382, https://doi.org/10.1126/science.abb3021, 2020.

McGlynn, E., Li, S., F. Berger, M., Amend, M., and L. Harper, K.: Addressing uncertainty and bias in land use, land use change, and forestry greenhouse gas inventories, Climatic Change, 170, 5, https://doi.org/10.1007/s10584-021-03254-2, 2022.

Milodowski, D. T., Smallman, T. L., Williams, M.: CARDAMOM driving data and C-cycle model outputs to accompany "Resolving scale-variance in the carbon dynamics of fragmented, mixed-use landscapes estimated using Model-Data Fusion", 2014-2019 [dataset]. Global Change Institute, School of GeoSciences, University of Edinburgh, https://doi.org/10.7488/ds/3509, 2022.

Montzka, C., Moradkhani, H., Weihermüller, L., Franssen, H.-J. H., Canty, M., and Vereecken, H.: Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter, Journal of Hydrology, 399, 410–421, https://doi.org/10.1016/j.jhydrol.2011.01.020, 2011.

Myrgiotis, V., Smallman, T. L., and Williams, M.: The carbon budget of the managed grasslands of Great Britain constrained by earth observations, 1–35, https://doi.org/10.5194/bg-2021-144, 2021.

Niu, S., Luo, Y., Dietze, M. C., Keenan, T. F., Shi, Z., Li, J., and Iii, F. S. C.: The role of data assimilation in predictive ecology, 5, art65, https://doi.org/10.1890/ES13-00273.1, 2014.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, 9, 3461–3482, https://doi.org/10.5194/gmd-9-3461-2016, 2016.

Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P., Kuppel, S., Koffi, E., Kane, A., Maignan, F., Chevallier, F., Ciais, P., and Prunet, P.: A new stepwise carbon cycle data assimilation system using multiple data streams to constrain the simulated land surface carbon cycle, 9, 3321–3346, https://doi.org/10.5194/gmd-9-3321-2016, 2016.

Pinnington, E., Quaife, T., Lawless, A., Williams, K., Arkebauer, T., and Scoby, D.: The Land Variational Ensemble Data Assimilation Framework: LAVENDAR v1.0.0, 13, 55–69, https://doi.org/10.5194/gmd-13-55-2020, 2020.

Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., and Rossiter, D.: SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty, SOIL, 7, 217–240, https://doi.org/10.5194/soil-7-217-2021 2021.

Reich, P. B., Rich, R. L., Lu, X., Wang, Y.-P., and Oleksyn, J.: Biogeographic variation in evergreen conifer needle longevity and impacts on boreal forest carbon cycle projections, PNAS, 111, 13703–13708, https://doi.org/10.1073/pnas.1216054110, 2014.

Revill, A., Myrgiotis, V., Florence, A., Hoad, S., Rees, R., MacArthur, A., and Williams, M.: Combining Process Modelling and LAI Observations to Diagnose Winter Wheat Nitrogen Status and Forecast Yield, 11, 314, https://doi.org/10.3390/agronomy11020314, 2021.

Rowland, C.S.; Morton, R.D.; Carrasco, L.; McShane, G.; O'Neil, A.W.; Wood, C.M.: Land Cover Map 2015 (25m raster, GB). NERC Environmental Information Data Centre. https://doi.org/10.5285/bb15e200-9349-403c-bda9-b430093807c7, 2017

Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., 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. USA, 108 (24), 9899-9904, https://doi.org/10.1073/pnas.1019576108, 2011.

Santoro, M.; Cartus, O.: ESA Biomass Climate Change Initiative (Biomass_cci): Global datasets of forest above-ground biomass for the years 2010, 2017 and 2018, v2. Centre for Environmental Data Analysis, 17 March 2021.

Scheiter, S., Langan, L., and Higgins, S. I.: Next-generation dynamic global vegetation models: learning from community ecology, 198, 957–969, https://doi.org/10.1111/nph.12210, 2013.

Scholze, M., Buchwitz, M., Dorigo, W., Guanter, L., and Quegan, S.: Reviews and syntheses: Systematic Earth observations for use in terrestrial carbon cycle data assimilation systems, 14, 3401–3429, https://doi.org/10.5194/bg-14-3401-2017, 2017.

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., de Mora, L., Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews, T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N., Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker, R. J., Predoi, V., Robertson, E., Siahaan, A., Smith, R. S., Swaminathan, R., Woodhouse, M. T., Zeng, G., and Zerroukat, M.: UKESM1: Description and Evaluation of the U.K. Earth System Model, 11, 4513–4558, https://doi.org/10.1029/2019MS001739, 2019.

Sitch, S., Huntingford, C., Gedney, N., Levy, P. E., Lomas, M., Piao, S. L., 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), 14, 2015–2039, https://doi.org/10.1111/j.1365-2486.2008.01626.x, 2008.

Smallman, T. L., Exbrayat, J-F., Mencuccini, M., Bloom, A. A., & Williams, M.: Assimilation of repeated woody biomass observations constrains decadal ecosystem carbon cycle uncertainty in aggrading forests. J Geophys Res-Biogeo, https://doi.org/10.1002/2016JG003520, 2017.

Smallman, T. L., and Williams, M.: Description and validation of an intermediate complexity model for ecosystem photosynthesis and evapotranspiration: ACM-GPP-ETv1, Geosci. Model Dev., 12 (6), https://doi.org/10.5194/gmd-12-2227-2019, 2019.

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

Smallman, T. L., Milodowski, D. T., and Williams, M.: From Ecosystem Observation to Environmental Decision-Making: Model-Data Fusion as an Operational Tool, Front. For. Glob. Change, 4:818661, https://doi.org/10.3389/ffgc.2021.818661, 2022.

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

Souza, C. M., Z. Shimbo, J., Rosa, M. R., Parente, L. L., A. Alencar, A., Rudorff, B. F. T., Hasenack, H., Matsumoto, M., G. Ferreira, L., Souza-Filho, P. W. M., de Oliveira, S. W., Rocha, W. F., Fonseca, A. V., Marques, C. B., Diniz, C. G., Costa, D., Monteiro, D., Rosa, E. R., Vélez-Martin, E., Weber, E. J., Lenti, F. E. B., Paternost, F. F., Pareyn, F. G. C., Siqueira, J. V., Viera, J. L., Neto, L. C. F., Saraiva, M. M., Sales, M. H., Salgado, M. P. G., Vasconcelos, R., Galano, S., Mesquita, V. V., and Azevedo, T.: Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine, 12, 2735, https://doi.org/10.3390/rs12172735, 2020.

Stoy, P. C., Williams, M., Disney, M., Prieto-Blanco, A., Huntley, B., Baxter, R., and Lewis, P.: Upscaling as ecological information transfer: a simple framework with application to Arctic ecosystem carbon exchange, Landscape Ecol, 24, 971–986, https://doi.org/10.1007/s10980-009-9367-3, 2009.

Tanneberger, F., Moen, A., Joosten, H., and Nilsen, N.: The peatland map of Europe, https://doi.org/10.19189/MaP.2016.OMB.264, 2017.

Turner, M. G., Calder, W. J., Cumming, G. S., Hughes, T. P., Jentsch, A., LaDeau, S. L., Lenton, T. M., Shuman, B. N., Turetsky, M. R., Ratajczak, Z., Williams, J. W., Williams, A. P., and Carpenter, S. R.: Climate change, ecosystems and abrupt change: science priorities, 375, 20190105, https://doi.org/10.1098/rstb.2019.0105, 2020.

UNFCCC. Glasgow Climate Pact. Report No. FCCC/PA/CMA/2021/L.16, 2021, https://unfccc.int/sites/default/files/resource/cma2021_L16_adv.pdf.

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

van der Laan-Luijkx, I. T., van der Velde, I. R., van der Veen, E., Tsuruta, A., Stanislawska, K., Babenhauserheide, A., Zhang, H. F., Liu, Y., He, W., Chen, H., Masarie, K. A., Krol, M. C., and Peters, W.: The CarbonTracker Data Assimilation Shell (CTDAS) v1.0: implementation and global carbon balance 2001–2015, 10, 2785–2800, https://doi.org/10.5194/gmd-10-2785-2017, 2017.

White, E. D., Rigby, M., Lunt, M. F., Smallman, T. L., Comyn-Platt, E., Manning, A. J., Ganesan, A. L., O'Doherty, S., Stavert, A. R., Stanley, K., Williams, M., Levy, P., Ramonet, M., Forster, G. L., Manning, A. C., and Palmer, P. I.: Quantifying the UK's carbon dioxide flux: an atmospheric inverse modelling approach using a regional measurement network, Atmos. Chem. Phys., 19, 4345–4365, https://doi.org/10.5194/acp-19-4345-2019, 2019.

Williams, M., Rastetter, E. B., Fernandes, D. N., Goulden, M. L., Shaver, G. R., & Johnson, L. C.: Predicting gross primary productivity in terrestrial ecosystems. Ecol Appl, 7 (3), 882-894, 1997.

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

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

Zhao, Y., Chen, X., Smallman, T. L. Flack-Prain, S., Milodowski, D. T., Williams, M.: Characterizing the Error and Bias of Remotely Sensed LAI Products: An Example for Tropical and Subtropical Evergreen Forests in South China, Remote Sens., 12 (19), 3122, https://doi.org/10.3390/rs12193122, 2020.

https://doi.org/10.5194/bg-2022-160
Preprint. Discussion started: 29 August 2022

Zhou, X., Zhou, T., and Luo, Y.: Uncertainties in carbon residence time and NPP-driven carbon uptake in terrestrial ecosystems of the conterminous USA: a Bayesian approach, 64, 17223, https://doi.org/10.3402/tellusb.v64i0.17223, 2012.