Agricultural and Forest Meteorology 385 (2026) 111197

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

First global carbon dynamics from an observational and process-informed hybrid perspective: Oversimplified respiration representation likely drives divergence in terrestrial carbon sequestration across models

Songyan Zhu [a], Wenquan Dong [b,c,d,*], Vasileios Myrgiotis [e], Jian Xu [f], Pablo Reyes-Muñoz [g], Mana Gharun [h], Rui Ma [i], Man Chen [j], Jadu Dash [a]

[a] School of Geography and Environmental Science, University of Southampton, Southampton, UK
[b] Department of Earth and Environmental Sciences, Lund University, Lund, Sweden
[c] School of GeoSciences, University of Edinburgh, Edinburgh, UK
[d] Royal Botanic Garden Edinburgh, Edinburgh, UK
[e] UK Centre for Ecology and Hydrology, Edinburgh, UK
[f] National Space Science Center, Chinese Academy of Sciences, Beijing, China
[g] IPL - Image Processing Laboratory, University of Valencia, Catedrático Agustín Escardino Benlloch 9, Paterna, 46980, Spain
[h] Institute of Landscape Ecology, Department of Geosciences, University of Münster, Münster, Germany
[i] Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France
[j] LIRA, Observatoire de Paris, Sorbonne Université, Paris, France

ARTICLE INFO

Dataset link: https://doi.org/10.5281/zenodo.16202323, https://doi.org/10.5194/essd-15-5301-2023

Keywords:
Carbon cycles
Remote sensing
Eddy covariance
Climate change
Process-based model
Machine learning
Terrestrial carbon dynamics

ABSTRACT

Process-based biosphere models, which formulate biophysical ecosystem processes, provide a mechanistic framework for understanding terrestrial carbon dynamics. In contrast, data-driven approaches, e.g., upscaling eddy covariance (EC) fluxes using satellite observations and machine learning, offer empirical estimates. These complementary methods often diverge significantly, particularly in estimating global photosynthetic uptake and ecosystem respiration, with discrepancies exceeding 50 PgC $yr^{-1}$and highlighting persistent uncertainties. To bridge this gap, we adopt a hybrid strategy that embeds physiological understanding via semi-empirical models, refines it with EC fluxes constrained by machine learning, and integrates process-based allocation to resolve component fluxes. This process-informed hybrid approach links ecological knowledge with predictive models, enabling generalisation beyond flux tower sites and supporting the development of new insights. We assess global carbon dynamics over the past two decades, applying Bayesian inference to evaluate climate impacts on land carbon processes. Our study delivers the first observational and process-informed hybrid assessment of global carbon flux and stock changes. Notably, while gross carbon uptake is consistent across methods (≈ 130 PgC yr^{-1}in 2022, increasing by 0.4 annually), net carbon uptake estimates diverge, from 6 PgC yr^{-1}in process-based models to 26 in conventional upscaling, and 16 in our hybrid model, reflecting structural differences in respiration parameterisation. Improved representation of respiratory processes is essential to capture the competing roles of photosynthesis and respiration under climate change. Despite rising global carbon fluxes and biomass stocks, tipping point risks remain: in tropical regions, increased photosynthesis (0.1 PgC yr^{-1}) is offset by rising respiration (0.05 PgC yr^{-1}).

1. Introduction

Process-based models are crucial for understanding global carbon cycles as they can simulate key biogeochemical processes, such as photosynthesis, respiration, and decomposition, under varying environmental conditions (Sitch et al., 2024; Friedlingstein et al., 2024). However, these models often exhibit divergent carbon uptake and interannual variability (Figure A.11), primarily due to factors such as parameterisation that smooths short-term fluctuations and structural simplifications that omit stochastic disturbances (e.g., extreme weather events or pest outbreaks) (Wang et al., 2014; Graven et al., 2024). For instance, many process-based models estimate net ecosystem exchange (NEE) with standard deviations of less than 10% across decades, whereas observational datasets often report year-to-year variations exceeding 20%, particularly in ecosystems sensitive to droughts or

* Corresponding author at: Department of Earth and Environmental Sciences, Lund University, Lund, Sweden.
E-mail address: wenquan.dong@mgeo.lu.se (W. Dong).

https://doi.org/10.1016/j.agrformet.2026.111197
Received 2 February 2026; Accepted 29 March 2026
Available online 27 April 2026

temperature anomalies (Friedlingstein et al., 2023; Zeng et al., 2020). This discrepancy could lead to an underestimation of ecosystem responses to climate extremes, ultimately reducing the models' predictive power for carbon-climate feedbacks.

Upscaling EC measurements offers a data-driven approach to understanding terrestrial carbon dynamics by extending site-level flux observations to regional and global scales (Jung et al., 2020; Nelson et al., 2024). The EC method directly measures land–atmosphere exchanges of CO_2 and other gases at high temporal resolution, capturing ecosystem responses to climate variability, disturbances, and land management. When integrated with remote sensing and machine learning, these flux tower observations can be scaled up to estimate photosynthetic carbon uptake as gross primary productivity (GPP) and NEE across broad spatial extents. For example, global upscaling studies have revealed interannual variations in GPP could exceed 15 PgC yr^{-1} (Joiner and Yoshida, 2020; Zhu et al., 2023b), highlighting climate-induced fluctuations that many process-based models could be improved to capture. This data-driven method provides a critical benchmark for improving model constraints, reducing uncertainties in carbon budgets, and refining predictions of ecosystem responses to climate change.

The EC captures the net exchange of CO_2 that can be partitioned into GPP and ecosystem respiration (RECO) (Reichstein et al., 2005; Lasslop et al., 2010). However, this approach does not fully account for the full carbon dynamics (Aubinet et al., 2012). Specifically, the EC method does not measure the carbon stored in soil, vegetation, or other pools, which are critical for understanding the potential long-term sequestration. To accurately assess ecosystem resilience and sustainability in the face of disturbances and a changing climate, it is essential to understand the full carbon dynamics, including how carbon uptake is distributed among different pools — such as soil organic carbon and vegetation biomass — along with detailed allocation fluxes like net primary production (NPP) (Kirschbaum et al., 2001). For instance, studies have shown that local ecosystem carbon stocks can fluctuate by as much as 50% over decadal timescales (Xu et al., 2016; Palviainen et al., 2020), depending on land use and climatic conditions.

Incorporating carbon allocation from process-based models, or partially process-based, semi-empirical models, into upscaled EC carbon uptake preserves the robustness of predictive machine learning models while also offering an observational perspective to address the sources of existing uncertainty in global carbon dynamics (Sitch et al., 2005; Reichstein et al., 2019). For example, the TRENDY (trends and drivers of the regional scale terrestrial sources and sinks of carbon dioxide) models are a suite of process-based dynamic global vegetation models used to simulate historical and future land carbon dynamics in response to climate, CO_2, and land-use changes. These models suggest some ecosystems showing a 20%–30% difference in how carbon is distributed between biomass and soil pools under changing climate scenarios (Sitch et al., 2024). This integration of observations, machine learning, and process-based knowledge helps reduce the uncertainty in carbon dynamics and minimise the reliance of traditional upscaling on the spatiotemporal coverage of EC towers. In addition, machine learning tools (Lundberg and Lee, 2017), like Bayesian causal inference (Li et al., 2023), contribute to generating new ecological knowledge into ecosystem carbon dynamics in response to land management (Zhu et al., 2023b) and climate change (Zhu et al., 2022b, 2023c). Therefore, these hybrid solutions can eventually aid in steering sustainable ecosystem management and preparing ecosystems to be resilient in an uncertain future, as we are falling short of meeting our climate commitments (Organisation for Economic Co-operation and Development (OECD), 2024).

Furthermore, integrating EC flux measurements and their upscaling with process-based models, such as those from TRENDY, offers a powerful approach not merely to replicate TRENDY's global outputs, but to infer critical parameters governing carbon allocation, including the allocation of photosynthetically fixed carbon (i.e. GPP) among respiration, growth, and carbon pools such as aboveground and belowground biomass. Such information is inaccessible from EC observations alone, yet it is essential for informing land management strategies and climate mitigation actions (Jung et al., 2020). TRENDY models typically operate at coarse spatial resolutions, for example 0.5° (Friedlingstein et al., 2024), which are inadequate for fragmented landscapes, where 27% of European land consists of habitats smaller than 0.02 km^2 on average (Bugnot et al., 2021). Observation-driven, upscaled EC data provide higher spatial resolution and capture stronger interannual variability, revealing ecological processes and climate-driven events probably missing in process-based models (Nelson et al., 2024) and highlighting the value of observational insights. Nevertheless, even high-resolution upscaled EC fluxes, such as 10–100 m grids (Zhu et al., 2023b, 2025b), cannot fully resolve carbon dynamics needed to evaluate the effectiveness of climate actions, including nature-based solutions (Griscom et al., 2017). By combining EC upscaling with TRENDY-model relationships, it is possible to estimate comprehensive carbon dynamics, enhancing the predictive power of models and supporting future research and policy-relevant applications.

In this study, we employed the hybrid modelling approach with satellite-derived vegetation proxies, climate reanalysis data, and TRENDY carbon allocation relationships to investigate the full carbon dynamics estimated from upscaled EC data. We also examined the consistency of these results with process-based models and satellite biomass products to assess the observational perspective on the spatiotemporal variability of carbon dynamics, as well as to better understand ecosystem responses to climate change and the uncertainties in our current knowledge of the global terrestrial carbon cycle.

2. Methodology

2.1. Hybrid model overview

The process-informed machine learning model employed here is Unified FLUXes (UFLUX) v2e. The UFLUX model, initially designed to gap-fill carbon and water fluxes from EC towers, improves accuracy by 30%–70% (Zhu et al., 2022a). It also extends to spatial 'gap-filling' (i.e., upscaling), harmonising the upscaling workflow for various flux types (Zhu et al., 2023b). This results in consistently strong performance across different scales (10–500 m) and ecosystem types, including natural pristine forests, managed grasslands, plantations, and drylands (Zhu et al., 2023a, 2024a). Additionally, it accounts for land management, climate effects such as CO_2 fertilisation (Zhu et al., 2023c), and the impact of air pollution on carbon uptake (Zhu et al., 2022b). The release of the UFLUX ensemble products ensures consistency across multiple satellites and proxies at different spatial scales, linking empirical, reflectance-based vegetation indices with physically based like solar-induced fluorescence observations, and delivering long-term (over 20 years) global datasets at resolutions ranging from 100 m to 0.25° (Zhu et al., 2025b). The updated version 'v2' incorporates plant physiological mechanisms (e.g., Farquhar photosynthesis) as the baseline model to constrain the machine-learning framework, moving beyond a purely data-driven approach (Dong et al., 2025). The 'e' in UFLUX v2e stands for 'emulator', reflecting its ability to emulate carbon allocation in TRENDY models and apply carbon uptake driven by both EC and satellite observations. In this context, unless otherwise specified, 'UFLUX' refers to the latest hybrid version, 'UFLUX v2e', while 'upscaling' refers to traditional, pure machine learning-based upscaling methods, such as those used in the older UFLUX v1 and the existing literature (Xiao et al., 2012; Ueyama et al., 2013; Tramontana et al., 2015, 2016; Joiner and Yoshida, 2020; Jung et al., 2020; Zeng et al., 2020).

The UFLUX workflow (Fig. 1) consists of three main steps: (1) estimating fluxes using baseline models, (2) refining the baseline estimates with EC fluxes through machine learning, and (3) emulating TRENDY

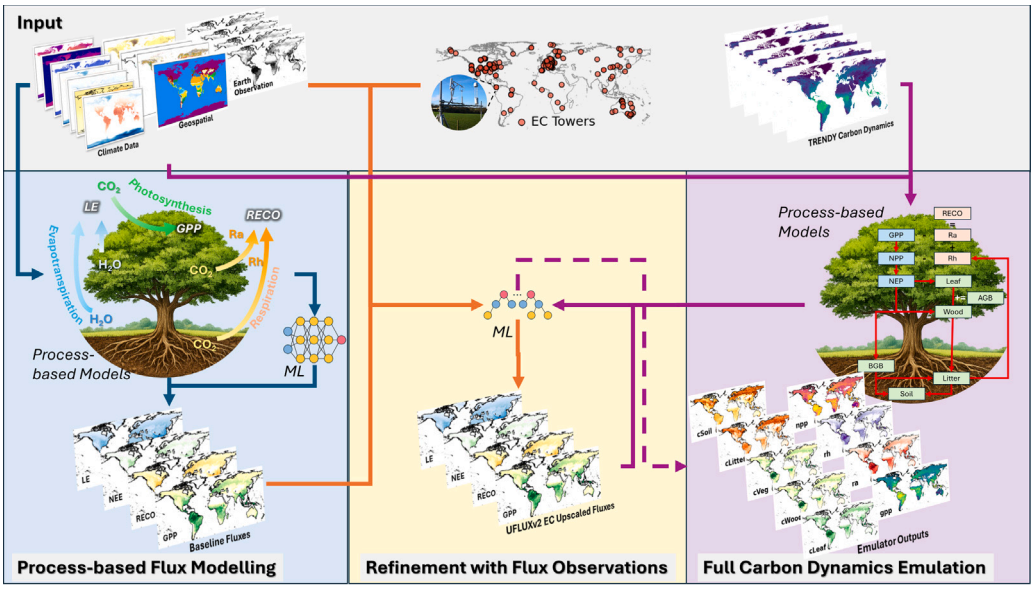

Fig. 1. Overview of the three-step UFLUX workflow for (1) process-based baseline modelling of ecosystem fluxes, (2) refinement of modelled fluxes using EC network observations, and (3) estimation of carbon allocations learned from TRENDY models. Grey background blocks denote inputs, while the coloured blocks from left to right represent the three workflow steps, with outputs shown at the bottom of each coloured block. The two machine-learning algorithms (labelled 'ML') correspond to neural networks (left) and decision-tree-based methods (right). Climate data, Earth observations, and auxiliary geospatial information (e.g. plant functional type, terrain) are used as inputs in all three steps. EC tower networks are applied only in Step 2 to refine baseline fluxes, while the 21 TRENDY models are used exclusively in Step 3 to learn carbon allocations. Further details are provided in the relevant sections.

models to distribute carbon uptake across different stocks. There are three types of process-based machine learning approaches. The first integrates ecological laws (e.g., GPP is always positive, and respiration increases with temperature between roughly 0–35 °C Lloyd and Taylor, 1994) directly into the machine learning loss function (Raissi et al., 2019). The second uses machine learning to correct biases between baseline model estimates and observations and to perform scaling, as demonstrated in Dong et al. (2025). The third learns key parameters, such as the ratio of GPP to autotrophic respiration, rather than learning autotrophic respiration, thereby ensuring that the machine learning respects the underlying ecological mechanisms.

These approaches can be implemented using either neural networks or decision-tree-based methods. Neural networks can incorporate ecological loss functions through end-to-end differentiability, capturing complex, high-dimensional nonlinear relationships at a high computational cost (Praveen, 2021). Decision-tree methods cannot encode ecological laws directly, as they are non-differentiable, but are robust, fast, and interpretable for tabular data like EC upscaling, particularly effective for smaller datasets and requiring minimal computation (Grinsztajn et al., 2022). UFLUX supports both approaches, leaving the choice to the user. By default, UFLUX employs a decision-tree-based framework. The decision-tree implementation is SmartForest4D (Zhu et al., 2025a,b), a deep-layered eXtreme Gradient Boosting (XGBoost) model (Chen and Guestrin, 2016). This choice exploits XGBoost's strengths in computational efficiency, scalability, predictive accuracy, regularisation, parallel processing, and feature-importance assessment (Chen and Guestrin, 2016). SmartForest4D is the first model to jointly integrate deep image-processing networks for extracting ecosystem structural information from satellite imagery (He et al., 2016); sequence-learning neural networks to capture temporal dynamics (Hochreiter and Schmidhuber, 1997); and deep-layered decision-tree models (e.g., XGBoost) to effectively fuse multi-modal inputs, including imagery, time series, and discrete land-management information (Zhou and Feng, 2019). The complete codebase is publicly available at the GitHub repository. In addition, a neural-network-based alternative is implemented using PyTorch, which provides flexible, dynamic computation graphs that facilitate rapid prototyping and the development of custom loss functions (Imambi et al., 2021). The study covers a 20-year period from 2003 to 2022, and all data are at a monthly temporal resolution. All datasets used in this study are available via Zenodo. Further details are provided in the following subsections.

2.2. Baseline estimation

To operate the baseline models, satellite Earth observations, climate variables, and ancillary geospatial data including International Geosphere-Biosphere Programme (IGBP) land covers (Friedl and Sulla-Menashe, 2019) were used to estimate ecosystem carbon (GPP and RECO) and water (latent energy, LE) fluxes. The net flux NEE was derived as the difference: $NEE = RECO - GPP$.

GPP was modelled using the light-use efficiency (LUE) approach (Yuan et al., 2007; Stocker et al., 2020):

$$GPP = \epsilon \cdot fAPAR \cdot PAR \tag{1}$$

where ϵ is the light-use efficiency, $fAPAR$ is the fraction of absorbed photosynthetically active radiation (PAR) which can be estimated using satellite remote sensing vegetation indices; ϵ is computed as (Prentice et al., 2014):

$$\epsilon = \phi_o \cdot M_c \cdot m_j \tag{2}$$

Here, ϕ_o denotes the quantum yield efficiency of photosynthesis, which is temperature-dependent. The quadratic temperature-scaling of ϕ_o differs between C3 and C4 plants. The term M_c represents the molar mass of carbon (approximately 12 g), and m_j is the light-use efficiency limitation factor. This factor is derived from the optimal ratio (χ) of internal leaf CO_2 partial pressure (C_i) to ambient CO_2 partial pressure (C_a), defined as $\chi = \frac{C_i}{C_a}$. The formulation of m_j differs between C3 and C4 species. For C3 plants:

$$m_j = (Ci - \Gamma^*)/(Ci + 2 \cdot \Gamma^*) \tag{3}$$

where Γ^* is the photorespiratory compensation point, influenced by temperature and air pressure and linked to enzyme kinetics through an

Arrhenius-type scaling (Stocker et al., 2020). For C4 plants, m_j is equal to 1. Here, ϵ is computed based on climate data (see Section 2.5).

RECO was estimated using the Lloyd–Taylor model (Lloyd and Taylor, 1994):

$$RECO = R_{\text{ref}} \times e^{B\cdot(T_a - T_{\text{ref}})} \tag{4}$$

$$B = \frac{\ln(Q_{10})}{10} \tag{5}$$

where T_a is the ambient temperature, R_{ref} is reference respiration at T_{ref} (default 10 °C), and Q_{10} is the temperature sensitivity factor, with a default value of 1.5.

The LE was estimated using a practical variant of the Priestley–Taylor potential evapotranspiration model (Priestley and Taylor, 1972) by Jet Propulsion Laboratory (JPL) (Fisher et al., 2008):

$$ET = \alpha \frac{\Delta R_n}{(\Delta + \gamma)\cdot \lambda_v} \tag{6}$$

$$LE = ET \times \lambda_v \tag{7}$$

where α is the Priestley–Taylor coefficient related to the Normalised Difference Vegetation Index (NDVI) as $\alpha = 1.26 \cdot (NDVI - 0.2)/0.8$, λ_v is the latent heat of vaporisation equal to 2.45×10^6 (J/kg), γ is the psychrometric constant of 0.066 (kPa/°C), and Δ is the slope of the saturation vapour pressure curve: $\Delta = (4098 \cdot e_s)/((T_a - 273.15) + 237.3)^2$, where e_s is vapour pressure deficit (VPD).

The baseline outputs amount to 4.8 GB, comprising monthly data from 2003 to 2022 for four fluxes: GPP, RECO, NEE, and LE, and NEE is calculated as RECO minus GPP.

2.3. Machine learning residual refinement

After running baseline models to estimate GPP, RECO, NEE, and LE separately for all 206 open-access FLUXNET2015 EC towers (Pastorello et al., 2020) (for a list of sites and tower meta information, see FLUXNET.org), machine learning was applied to model residuals between EC flux measurements ($\boldsymbol{\Phi}_{EC}$) and the baseline modelled ($\boldsymbol{\Phi}_p$) (Dong et al., 2025):

$$\hat{f}(\mathbf{x}) = \boldsymbol{\Phi}_{EC} - \boldsymbol{\Phi}_p \tag{8}$$

A trained model $\hat{f}(\mathbf{x})$ was used to correct the flux estimates as $\boldsymbol{\Phi}_u$:

$$\boldsymbol{\Phi}_u = \boldsymbol{\Phi}_p + \hat{f}(\mathbf{x}) \tag{9}$$

Specifically, we employed SmartForest4D, which is built upon deep-layer XGBoost. Beyond the introduction of XGBoost in Section 2.1, additional advantages of this approach include the following: XGBoost enhances gradient boosting by introducing regularisations, tree pruning, and sparse data handling to reduce overfitting and improve performance (Zhu et al., 2024a; Nelson et al., 2024). It employs a second-order approximation for optimisation, speeding up training and enhancing accuracy. Additionally, its parallelisation capabilities make it highly efficient, enabling faster computation and scalability, which is ideal for large datasets (Chen and Guestrin, 2016). While UFLUX was designed to be inclusive, supporting machine learning algorithms like Random Forest, CatBoost, and deep neural networks (including Long Short-Term Memory and Deep Forest), XGBoost was selected as the default due to its balance of performance, time efficiency, and computational resource requirements (Zhu et al., 2024b).

In addition to the residual refinement implemented using decision-tree-based SmartForest4D, UFLUX also realises an alternative process-informed machine-learning approach based on neural networks, which are differentiable and therefore allow ecological constraints to be incorporated directly into the training loss. Under this framework, carbon mass conservation is explicitly enforced by constraining net carbon flux NEE to equal the difference between respiration RECO and photosynthetic productivity GPP. The total neural network loss, $\mathcal{L}$, is thus composed of the sum of data-misfit losses for each predicted flux ($\sum_i \mathcal{L}^i_{\text{data}}$) and ecological constraint penalties ($\mathcal{L}^i_C$).

$$\mathcal{L} = \sum_i \mathcal{L}^i_{\text{data}} + \lambda^i_C \mathcal{L}^i_C \tag{10}$$

Where λ^i_C refers to hyperparameters controlling the strength of the ecological constraints relative to data fitting. The ecological constraint penalties include, first, the carbon mass conservation penalty, where the hat notation indicates modelled variables:

$$\mathcal{L}_{C\,mass} = \left\| \hat{\text{NEE}} - (\hat{\text{RECO}} - \hat{\text{GPP}}) \right\| \tag{11}$$

Furthermore, GPP and RECO are constrained to remain non-negative, reflecting fundamental ecological knowledge (Chapin III et al., 2011). This is implemented as a hard constraint, meaning that negative values are never allowed, by modifying the neural network output layer to fully guarantee physically meaningful predictions (Pan, 2021).

$$\hat{\text{GPP}} = \text{ReLU}(f_\theta(\mathbf{x})) \quad , \quad \hat{\text{RECO}} = \text{ReLU}(g_\theta(\mathbf{x})) \tag{12}$$

Where $ReLU$ denotes the Rectified Linear Unit activation function, which enforces non-negativity (He et al., 2020), and $\mathbf{x}$ represents the input data. f_θ and g_θ are neural network functions that map the inputs outputs for GPP and RECO, respectively. $theta$ represents all trainable parameters (weights and biases) of the network. It is also possible to include a soft penalty to discourage negative GPP or RECO predictions:

$$\mathcal{L}_{\text{positivity}} = \lambda_{\text{GPP}} \left\langle \max(0, -\hat{\text{GPP}}) \right\rangle + \lambda_{\text{RECO}} \left\langle \max(0, -\hat{\text{RECO}}) \right\rangle \tag{13}$$

This term allows the neural network to produce negative outputs while penalising them in the loss function. The penalty grows smoothly as the prediction becomes more negative. This approach is flexible and easy to tune, but small negative values may still occur during training. Therefore, a hard constraint enforcing non-negativity is generally recommended as the default in UFLUX. We also incorporate the ecological principle that GPP is positively related to photosynthetically active radiation (PAR), such that any predictions that are negative are penalised:

$$\mathcal{L}_{\text{light-GPP}} = \left\langle \max\left(0, -\frac{\partial \hat{\text{GPP}}}{\partial \text{PAR}}\right) \right\rangle \tag{14}$$

However, to account for the saturation effect (Bao et al., 2022), where GPP no longer increases at higher PAR, this relationship is only applied up to 200 W m^{-2}, which represents a relatively broad limitation (Zhu et al., 2024a). A similar approach is applied to the relationship between temperature (Ta) and RECO, considering only temperatures below 40 °C, to account for limitations due to enzyme kinetics, substrate availability, and thermal stress (Heskel et al., 2016):

$$\mathcal{L}_{\text{temperature-RECO}} = \left\langle \max\left(0, -\frac{\partial \hat{\text{RECO}}}{\partial \text{Ta}}\right) \right\rangle \tag{15}$$

To address the potential issue regarding the completeness of sampling space — i.e., the FLUXNET2015 may not capture all global ecosystems and climate conditions, we introduced an ensemble strategy to estimate uncertainty. We trained an ensemble of 10 independent models, each using 80% of the training data and 80% of the features each time, aggregating the outputs to calculate confidence intervals (Zhu et al., 2024a).

For validation, we employed Leave-One-Out Cross-Validation (LOOCV) (Geisser, 1975), a rigorous resampling method where each data point is used as a test set once, with the remaining $n - 1$ samples used for training. This approach ensures an unbiased performance estimate, with each EC tower in FLUXNET2015 serving as the test set once, providing comprehensive validation as utilised in previous studies (Zhu et al., 2023b, 2024a; Dong et al., 2025).

Furthermore, UFLUX supports feature importance techniques and explainable machine learning methods to capture the partial independence of carbon/water cycles from climate variables, advancing ecological knowledge (Zhu et al., 2023c).

After training and validating the machine learning models, we calculated global GPP, RECO, NEE, and LE using the baseline models and then upscaled the trained residual learning model to adjust these baseline estimates. The refined outputs also amount to 4.8 GB, comprising monthly data from 2003 to 2022 for the four fluxes.

2.4. Carbon allocation emulation

The UFLUX allocates the refined carbon uptake GPP estimates into ecosystem stocks using all the 21 TRENDY v12 models (Friedlingstein et al., 2023). Herein, the carbon dynamics (fluxes and stocks) are as follows (Table 1): GPP represents the total carbon fixed through photosynthesis, forming the primary input into terrestrial ecosystems. A portion of this carbon is lost via RECO, which includes Autotrophic Respiration (Ra) from plants and Heterotrophic Respiration (Rh) from microbial activity ($RECO = Ra + Rh$). The remaining carbon after Ra is Net Primary Production (NPP, $NPP = GPP - Ra$), contributing to plant biomass growth, while Net Ecosystem Production (NEP, $NEP = -NEE$), defined as the balance of GPP and RECO, determines whether an ecosystem acts as a carbon sink or source. This carbon is further allocated into various pools: the vegetation (cVeg) pool, which stores carbon in aboveground biomass (AGB = cLeaf + cWood) and belowground biomass (BGB); the litter pool (cLitter), composed of decomposing plant material that contributes to nutrient cycling and carbon storage; and the soil pool (cSoil), which stores carbon in soil organic matter, including decomposed plant and animal material, all of which regulate carbon cycling over different timescales (Fig. 1).

In this step, the process-informed nature of UFLUX is realised by learning key ecological carbon allocation parameters rather than directly predicting fluxes or stocks. Unlike approaches that directly predict carbon fluxes (Zhu et al., 2024a) or stocks like aboveground biomass (Lamahewage et al., 2025) using climate and satellite observations, UFLUX infers parameters that follow the natural carbon allocation pathways of ecosystems, as illustrated in the purple block of Fig. 1. UFLUX emulates parameters for both terms on the right-hand side of the allocation equations, not only enabling carbon allocation but also enforcing mass conservation. For example, considering the relationship GPP = NPP + Ra, the model is trained using climate and satellite inputs to learn the allocation ratios of GPP into autotrophic respiration Ra and net primary production NPP, against ratios from TRENDY models:

$$\frac{\mathrm{GPP}}{\mathrm{Ra}} = \hat{f}_1(\mathbf{x}), \frac{\mathrm{GPP}}{\mathrm{NPP}} = \hat{f}_2(\mathbf{x}), \tag{16}$$

Where the machine-learning functions $\hat{f}_1$ and $\hat{f}_2$ represent learned nonlinear relationships between the input drivers **x** (e.g. climate and remote-sensing variables) and the corresponding allocation parameters. This formulation enables the estimation of NPP as the residual between GPP and Ra, while ensuring mass conservation. These learnt parameters are then applied to upscaled EC GPP data to estimate global Ra and NPP. The model is further trained to learn the allocation ratios between autotrophic and heterotrophic respiration Rh, which are applied to upscaled EC RECO to estimate global Rh and verify mass balance consistency. Consequently, the accumulation of carbon in different pools is explicitly learned. For instance, UFLUX treats the NPP-to-biomass conversion rate as a target parameter, informed by climate and satellite data as well as preceding GPP and Ra dynamics, enhancing the guidance provided to the machine learning model:

$$\frac{\sum_1^{12} \mathrm{NPP}}{\mathrm{AGB}} = \hat{f}_3(\mathbf{x}) \tag{17}$$

Where AGB denotes aboveground biomass, and NPP is summed from January to December, as carbon stocks are reported on an annual basis. The net ecosystem production NEP from TRENDY is subsequently used as a check to verify mass conservation is ensured in estimating photosynthetic and respiratory fluxes. This process ultimately results in the development of over ten specialised machine learning models, each dedicated to a specific ecological parameter, rather than a single integrated model. The models were trained using the SmartForest4D approach, with a 70:30 training-to-test split applied to global pixels randomly, following a 5-fold cross-validation scheme.

Table 1
Carbon stocks and fluxes from UFLUX, derived from upscaled EC data and TRENDY model emulation.

Fluxes		
GPP	Gross Primary Productivity	gC m^{-2} d^{-1}
RECO	Ecosystem Respiration	gC m^{-2} d^{-1}
NEE	Net Ecosystem Exchange	gC m^{-2} d^{-1}
LE	Latent Energy (i.e., Latent Heat Flux)	W m^{-2}
NPP	Net Primary Productivity	gC m^{-2} d^{-1}
Ra	Autotrophic Respiration	gC m^{-2} d^{-1}
Rh	Heterotrophic Respiration	gC m^{-2} d^{-1}
Stocks		
cVeg	Vegetation Carbon	gC m^{-2}
cLeaf	Leaf Carbon	gC m^{-2}
cWood	Wood Carbon	gC m^{-2}
cLitter	Litter Carbon	gC m^{-2}
cSoil	Soil Carbon	gC m^{-2}

Bayesian inference was applied to assess the causal impact of climate variables on fluxes, using Bayes' theorem to integrate prior knowledge with observed data; This approach provides probabilistic estimates of model parameters, enhancing uncertainty quantification and model interpretability (Koot et al., 2023).

2.5. Climate, satellite, geospatial, and carbon cycle data

In this study, alongside the 21 TRENDY models (Table 3), we incorporate EC fluxes, satellite observations, climate datasets, and ancillary geospatial information:

- The FLUXNET2015 database is a global collection of EC flux measurements from 206 open-access sites, providing half-hourly records of ecosystem carbon, water, and energy fluxes from the early 1990s to the end of 2014, along with meteorological observations (Pastorello et al., 2020). In this study, we used NEE as well as nighttime-partitioned GPP and RECO (Reichstein et al., 2005). FLUXNET, along with regional EC networks, has been successfully supporting land ecosystem research for decades on like validating new satellite platforms and ecosystem models as well as advancing our understanding of global carbon dynamics, particularly in relation to climate variability and land use change (Baldocchi, 2020).
- ERA5-Land is an hourly climate reanalysis dataset developed by the European Centre for Medium-Range Weather Forecasts (ECMWF), providing essential meteorological variables a 0.1° resolution for investigating land–atmosphere interactions (see variables in Table 2) (Hersbach et al., 2020).
- The satellite vegetation index employed was Near-Infrared Reflectance of vegetation (NIRv), which overcomes the saturation issue in dense vegetation observed in other indices and more effectively captures variations in canopy structure and chlorophyll activity (Badgley et al., 2017, 2019). NIRv is defined as: $NIRv = NIR \cdot (NIR - R)/(NIR + R)$ where NIR and R represent the near-infrared and red band reflectance, respectively, derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) MCD43A4.061 Nadir BRDF-Adjusted Reflectance product. This dataset provides daily surface reflectance at a 500-metre resolution, adjusted for the bidirectional reflectance distribution function (BRDF), and has delivered high-quality Earth observations for over 20 years (Schaaf et al., 2021).

Table 2
ERA5-Land climate variables used to drive the UFLUX model, with dewpoint temperature utilised exclusively for the calculation of VPD.

Code	Full name	Unit
TA	Air temperature at 2 m	°C
Td	*Dewpoint temperature*	°C
TSoil	Soil temperature (0–7 cm)	°C
VPD	Vapour pressure deficit (VPD)	hPa
SWIN	Shortwave incoming solar radiation	W m^{-2}
LWIN	Longwave thermal radiation	W m^{-2}
RNET	Net solar radiation	W m^2
P	Total precipitation	m

- The global land cover data was obtained from the MODIS MCD12C1.061 dataset, which provides yearly updates at a 0.05° resolution. This dataset, based on 17 land cover classes — including forests, croplands, and urban areas — uses MODIS multi-spectral data to monitor land cover dynamics and ecosystem changes (Sulla-Menashe and Friedl, 2018).
- The Köppen climate classification categorises global climates into five major groups (A: tropical, B: dry, C: temperate, D: cold, E: polar) based on temperature and precipitation patterns. Each group is further subdivided, reflecting how climate influences vegetation types (Beck et al., 2018).
- The Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) provides high-resolution (30 m) global topographic data, covering 80% of Earth's land surface and collected by NASA in 2000 using radar interferometry (Farr et al., 2007).
- The C4/C3 vegetation ratio data, sourced from the global C4 distribution mapping by Luo et al. (2024) (Luo et al., 2024), were incorporated into the light use efficiency module to improve efficiency estimates by reflecting the varying proportions of C3 and C4 photosynthetic pathways across diverse ecosystems and climatic regions (Gowik and Westhoff, 2011).

Except for the EC fluxes, all other datasets were gridded. To ensure consistency, these datasets were resampled to a common grid prior to baseline modelling and emulation. Potential spatial resolution issues have been evaluated and addressed in our previous study (Zhu et al., 2025a). For the refinement of baseline estimates, gridded data were interpolated to the locations of EC towers to train the model, which was then used to upscale predictions globally. The effects of differing data sources and resolutions have also been examined and resolved in our earlier work (Zhu et al., 2025b). The UFLUX output products have a time span of 20 years (2003–2022) with a spatial resolution of 0.5° (monthly for fluxes and yearly for stocks), with TRENDY model outputs from different spatial resolutions resampled to the same 0.5° grid.

2.6. Evaluation metrics

We employed statistical validation metrics to evaluate the performance of UFLUX in estimating carbon dynamics:

(1) The coefficient of determination (r^2) quantifies the proportion of variance in the observed data that is explained by the model, offering insight into how well the model fits the data.

$$r^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{18}$$

where y_i are the observed values, $\hat{y}_i$ are the predicted values, and $\bar{y}$ is the mean of the observed values.

(2) Root Mean Squared Error (RMSE) that measures the average magnitude of errors between observed and predicted values, giving a sense of the model's prediction accuracy in the same units as the original data.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{19}$$

where: y_i are the observed values, $\hat{y}_i$ are the predicted values, and n is the total number of data points.

(3) Normalised Root Mean Squared Error (nRMSE) is a normalised version of RMSE, where the RMSE is divided by the mean value of the observed data.

$$nRMSE = \frac{RMSE}{\text{Mean of observed values}} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}{\text{Mean of } y_i} \tag{20}$$

(4) Slope and intercept of the linear regression line between the observed and predicted values. Ideally, the slope should be 1, and the intercept should be 0, which would indicate a perfect prediction model.

$$\text{slope} = \frac{n\sum y_i\hat{y}_i - \sum y_i \sum \hat{y}_i}{n\sum \hat{y}_i^2 - (\sum \hat{y}_i)^2} \tag{21}$$

$$\text{intercept} = \frac{\sum y_i - \text{slope}\sum \hat{y}_i}{n} \tag{22}$$

(5) Mean Absolute Error (MAE) is the average of the absolute differences between the observed and predicted values. It provides a measure of how close the predictions are to the actual outcomes.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{23}$$

(6) IAV (Inter-Annual Variability) that measures the year-to-year variability in the estimated global carbon dynamics.

$$IAV = \frac{\sigma}{\mu} = \frac{\text{Standard Deviation of yearly means}}{\text{Mean of yearly means}} \tag{24}$$

(7) Conservation of Mass. The estimated carbon fluxes and stocks must adhere to the law of conservation of mass, meaning the estimates should conform to:

$$NEP = -NEE = GPP - RECO \tag{25}$$

$$RECO = Ra + Rh \tag{26}$$

$$NPP = GPP - Ra \tag{27}$$

Where, NEP stands for net ecosystem production, Ra for autotrophic respiration, Rh for heterotrophic respiration, and NPP for net primary productivity. A full list of acronyms and their definitions can also be found in Table A.5. The net changes in carbon stocks should ideally equal the net accumulation ecosystem fluxes, assuming disturbances such as fires are not considered. While turnover may introduce delays, the values of the two should remain within a reasonable range.

(8) Comparison with other datasets. Besides, we validated the UFLUX output product against the European Space Agency's Climate Change Initiative (CCI) AGB dataset. The CCI AGB dataset version 5.01 represents a significant advancement in global biomass monitoring (Santoro and Cartus, 2024). This product provides spatially explicit estimates of forest AGB at 100-metre resolution globally. Version 5.01 incorporates multi-sensor satellite data, including radar backscatter from Advanced Land Observing Satellite (ALOS-1 and ALOS-2), Sentinel-1, and Envisat's ASAR (Advanced Synthetic Aperture Radar) data, calibrated with extensive ground measurements from forest inventory plots (Araza et al., 2022). CCI-AGB v5.01 offers critical reference data for carbon cycle studies, with available estimates for the years 2010 and 2015–2021. This product serves as a valuable validation resource for model-based carbon stock estimations due to its consistent methodology, making it particularly suitable for evaluating spatially explicit carbon dynamics at regional to global scales.

To implement this validation approach, the UFLUX-derived cAGB was computed as the aggregate of two carbon pools: cWood and cLeaf. Concurrently, we established the reference dataset by resampling the

Table 3
List of TRENDY models with their full names and references.

TRENDY model name	Abbreviation	Reference
Community Atmosphere–Biosphere Land Exchange – Population	CABLE-POP	Haverd et al. (2018)
CARbon DAta MOdel FraMework	CARDAMOM	Smallman et al. (2021)
Canadian Land Surface Scheme Including Biogeochemical Cycles	CLASSIC	Melton et al. (2020)
Community Land Model version 5.0	CLM5.0	Lawrence et al. (2019)
Dynamic Land Ecosystem Model	DLEM	You et al. (2022)
Energy Exascale Earth System Model	E3SM	Yang et al. (2023)
Ecosystem Demography Model version 3	EDv3	Ma et al. (2022)
Integrated Biosphere Simulator	IBIS	Xia et al. (2024)
Integrated Science Assessment Model	ISAM	Shu et al. (2020)
Interaction Soil–Biosphere–Atmosphere – Carbon Transport in Rivers and Inland Waters	ISBA-CTRIP	Delire et al. (2020)
Jena Scheme for Biosphere–Atmosphere Coupling in Hamburg	JSBACH	Mauritsen et al. (2019)
Joint UK Land Environment Simulator	JULES	Wiltshire et al. (2021)
Lund-Potsdam-Jena General Ecosystem Simulator	LPJ-GUESS	Smith et al. (2014)
Lund-Potsdam-Jena managed Land model	LPJml	Heinke et al. (2022)
Lund-Potsdam-Jena model adapted by the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL)	LPJwsl	Poulter et al. (2011)
Land surface Processes and eXchanges model – University of Bern version	LPX-Bern	Lienert and Joos (2018)
Organising Carbon and Nitrogen model	OCN	Zaehle et al. (2011)
Organizing Carbon and Hydrology In Dynamic Ecosystems	ORCHIDEE	Vuichard et al. (2019)
Sheffield Dynamic Global Vegetation Model	SDGVM	Walker et al. (2017)
Vegetation Integrative Simulator for Trace gases	VISIT	Kato et al. (2013)
Yale Interactive Biosphere Model	YIBs	Yue and Unger (2015)

CCI-AGB estimates to the same 0.5° grid spatial resolution and converting the CCI AGB estimates to carbon content using a conversion factor of 0.47 (IPCC, 2006). This enables a systematic evaluation of both the spatial congruence and quantitative accuracy between modelled and observation-based carbon stock distributions, thereby providing comprehensive metrics for assessing the fidelity of the UFLUX product in representing terrestrial carbon pools at the global scale.

3. Results

This section presents the results of the study, while the corresponding analysis, interpretation, and discussion are provided in Section 4 Discussion.

3.1. Model validation

3.1.1. Baseline residual refinement

Fig. 2 presents the validation results for both the baseline models and the UFLUX model.

The baseline models exhibited a validation r^2 of approximately 0.5. Among the carbon fluxes, GPP achieved the highest r^2 value (0.63), followed by RECO at 0.48, while NEE had the lowest at 0.29. The water flux, LE, showed an r^2 of 0.54. The RMSE for carbon fluxes was around 2.5 gC m^{-2} d^{-1}, with GPP having the highest RMSE and NEE the lowest, differing by approximately 1 gC m^{-2} d^{-1}. The RMSE for LE was slightly below 30 W m^{-2}. The normalised RMSE (nRMSE) exceeded 50% across fluxes, with RECO surpassing 70% and LE approximately 54%. The slope of the linear regression between the baseline models and EC measurements was greater than 1 for all flux types, with GPP being closest to the one-to-one line (1.07), while RECO and LE deviated by 25%–30%.

In contrast, the UFLUX model demonstrated improved performance over baseline models, with higher r^2 values and lower RMSE. The r^2 values for all four fluxes were close to 0.8, representing an increase of approximately 60% compared to the baseline models. GPP had the highest r^2 (0.84), while NEE had the lowest (0.72), with RECO and LE both close to 0.8. With respect to RMSE, a reduction of approximately 50% was observed relative to the baseline models. The three carbon fluxes had RMSE values slightly above 1 gC m^{-2} d^{-1}, with GPP exhibiting the highest RMSE (1.55 gC m^{-2} d^{-1}), while RECO and NEE remained below 1.2 gC m^{-2} d^{-1}. The RMSE for LE was approximately 16 W m^{-2}. The nRMSE was about one-third for fluxes, e.g., 31%–33% for GPP, RECO, and LE. The slope of the linear regression against EC measurements was below 1 for all fluxes, with GPP and LE being closest to the one-to-one line (0.93–0.94), whereas RECO and NEE deviated by around 10%.

Scatter plots for both the baseline and UFLUX models (Fig. 2) showed largely symmetrical distributions around the one-to-one line for all four fluxes. Compared with the baseline, UFLUX predictions were more tightly clustered around the one-to-one line, with the most pronounced improvement observed for NEE. For GPP, closer agreement with the one-to-one line was evident across both low and high GPP ranges. Improvements in r^2 varied by plant functional type, with a 1.5-fold increase for evergreen broadleaf forests and an 18% increase for mixed forests. Substantial improvements were also observed in managed ecosystems, including croplands and grasslands, where r^2 increased by approximately 40%. For RECO, improvements were more pronounced at higher flux values, particularly in tropical ecosystems and during summer. When RECO values were below 5 gC m^{-2} d^{-1}, r^2 increased by approximately 80%, whereas for values above 5 gC m^{-2} d^{-1}, the increase was approximately fivefold. Seasonally, summer exhibited nearly a twofold increase in r^2 , compared with an increase of about 10% in winter. Across plant functional types, evergreen broadleaf forests showed a 1.5-fold improvement, while deciduous needleleaf forests exhibited a sevenfold increase. Improvements in NEE were characterised by a fusiform distribution pattern, with larger r^2 increases under conditions of strong carbon uptake or release. This effect was particularly evident in croplands (with r^2 increased by approximately 2.5 times), where land management processes that are not realised in the baseline model but are implicitly captured in observational data. For LE, representing ecosystem evapotranspiration, improvements followed a pattern similar to RECO, with larger gains at higher flux values. Among plant functional types, wetlands showed the largest increase, exceeding 100%, where the high hydrological complexity are not fully represented in the baseline model. Seasonally, also summer exhibited more than a twofold increase in r^2 .

3.1.2. Carbon allocation emulation

Fig. 3(a) presents the validation results of UFLUX against the TRENDY models, with corresponding validation statistics summarised in Table 4.

The r^2 values for carbon fluxes and stocks were generally above 0.9, with particularly high agreement when emulating the mean of the 21 TRENDY models (approximately 0.95), compared to the 25th and 75th percentiles across these models. The differences between the estimated and modelled carbon fluxes and stocks were approximately 5% for most variables. Variations in r^2 across different uncertainty levels (mean,

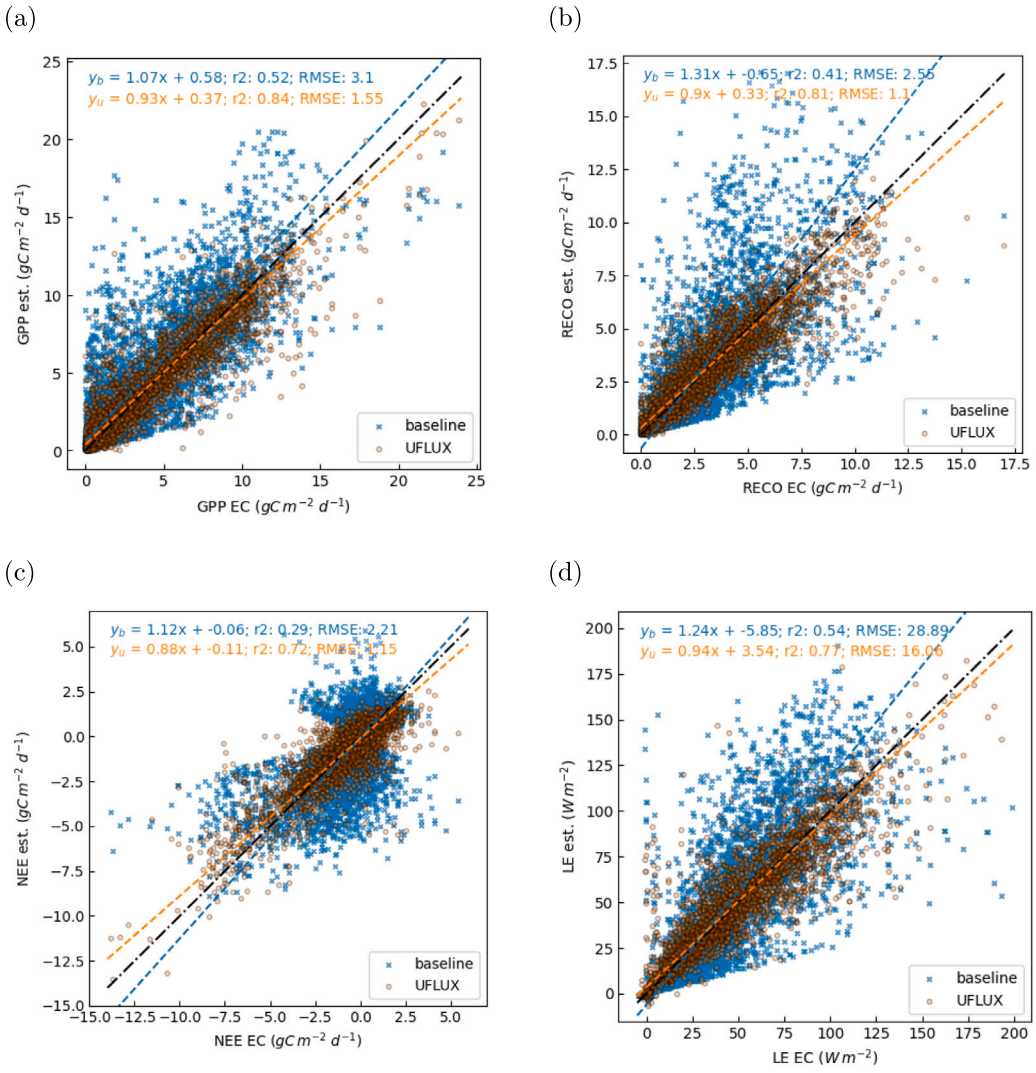

Fig. 2. Global validation of model estimates against the FLUXNET2015 EC (x-axis) dataset for four flux types. Blue dots represent semi-empirical baseline models, while orange dots correspond to UFLUX. The subscripts y_b and y_u denote estimates from the baseline and UFLUX, respectively. RMSE refers to the root mean squared error. The unit for carbon fluxes is gC m^{-2} d^{-1} and W m^{-2} for the water flux LE.

25th percentile, and 75th percentile) were within 5% for carbon stocks but reached nearly 10% for Ra and Rh, respectively.

The nRMSE ranged from 5 to 30% across carbon fluxes and stocks, with an average of approximately 15% in most cases. Mean estimates exhibited lower nRMSE values than the 25th and 75th percentiles, typically by about 5%. On average, Ra and Rh had the lowest nRMSE (approximately 11% across the three uncertainty levels), whereas cWood showed the highest nRMSE, around 25%.

Regarding the linear regression analysis against TRENDY model validation data, the regression slopes were close to unity but consistently slightly lower by approximately 5%. The variation in slopes across fluxes and stocks was around 2%.

Regarding spatial distribution, the r^2 of UFLUX in emulating TRENDY carbon dynamics indicated generally good performance, with most regions showing r^2 values around 0.8 or higher. Lower values were observed in tropical regions, such as the Amazon and central Africa, where the seasonality of ecosystem carbon dynamics is weaker than in temperate regions, with r^2 around 0.5 (Fig. 3a, top). For nRMSE (Fig. 3a, bottom), values were very low globally, typically well below 0.5. Tropical regions, including the Amazon and central Africa, also exhibited particularly low nRMSE, approaching zero. Across plant functional types, the average r^2 remained high, around 0.8 or above. Evergreen broadleaf forests showed slightly lower r^2 (near 0.8) and the lowest nRMSE of about 1%. Snow and ice regions, which are largely non-vegetated, had r^2 around 0.22 and higher nRMSE of approximately 20%. Managed agricultural croplands, typically patchy and less distinguishable at coarse global resolution, exhibited a relatively low r^2 of 0.7 and nRMSE near 10%. Seasonally, r^2 was highest during periods of active vegetation and elevated GPP, while lower values (about 0.5) were observed in cold winter and hot summer seasons, when climatic conditions are more extreme and less represented in the training dataset.

3.1.3. UFLUX cAGB versus CCI cAGB benchmark

Fig. 4(a) illustrates the relationship between the UFLUX-derived cAGB and the CCI-derived cAGB through a hexbin density scatter plot. The statistical analysis revealed a positive correlation (r^2 = 0.70) between the two datasets, with a RMSE of 19.11 MgC ha^{-1}. This level of agreement is particularly noteworthy given the inherent methodological differences between the model-based approach and the satellite-derived observations. The regression line (red) indicates a systematic tendency for UFLUX to slightly underestimate cAGB compared to CCI

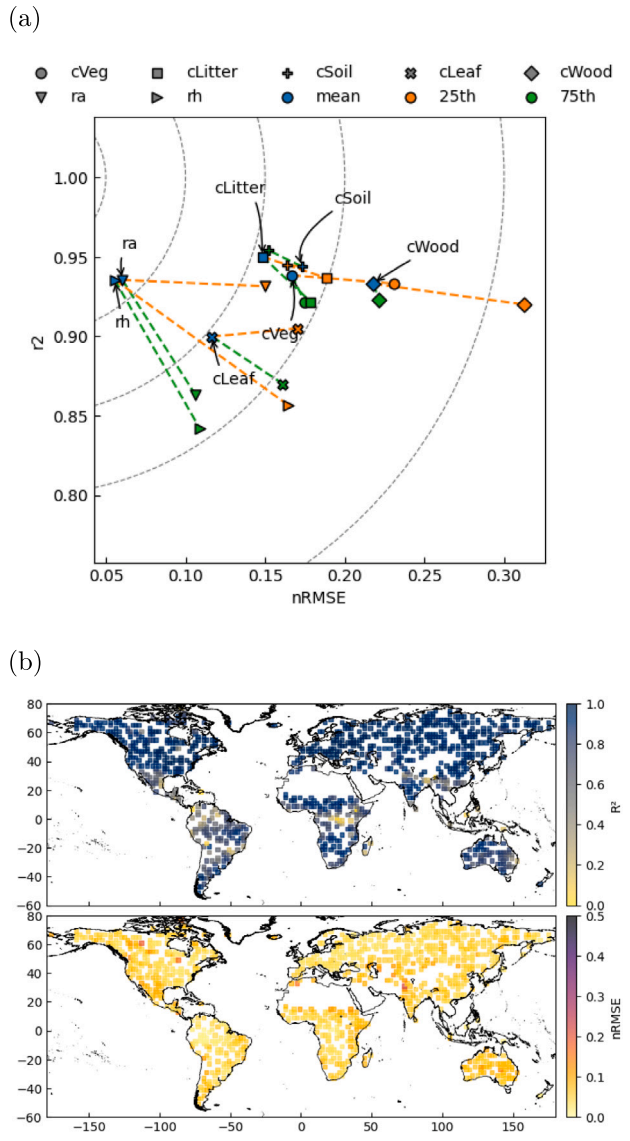

Fig. 3. (a) Validation of UFLUX for emulating TRENDY models. The *x*-axis represents the nRMSE, while the *y*-axis denotes the r^2 . Different marker shapes distinguish carbon flux and stock types, while colours indicate uncertainty: blue represents UFLUX performance for estimating the mean (of 21 TRENDY models), the 25th percentile is shown in orange, and the 75th percentile in green. The uncertainties for the same flux/stock are connected by dashed lines. (b) Spatial maps showing the validation r^2 and nRMSE averaged over multiple years for the emulated carbon dynamics.
Note: r^2 values range from 0 to 1, while nRMSE values are limited to 0–0.5 to improve visual clarity.

values at higher biomass ranges, as evidenced by its deviation from the 1:1 line (black dashed). However, the high density of points (orange–red hexagons) at lower cAGB values (0–50 MgC ha^{-1}) demonstrates overestimation on this range.

The spatial distribution of differences between UFLUX and CCI cAGB estimates is presented in Fig. 4(b). The global pattern reveals predominantly small differences (±20 MgC ha^{-1}) across large portions of the terrestrial surface, indicating overall spatial coherence between the two datasets. However, notable regional discrepancies are evident, particularly in tropical Africa where UFLUX tends to underestimate cAGB (blue regions) compared to the CCI benchmark. Conversely, consistent overestimation patterns (red regions) are observed in portions

Table 4
Global validation of UFLUX in emulating TRENDY models for carbon fluxes and stocks. RMSE represents the root mean squared error. The slope and intercept correspond to the linear regression of model estimates against EC fluxes. The mean represents the average EC flux. MAE denotes the mean absolute error. nRMSE is normalised RMSE. Carbon fluxes are measured in gC m^{-2} d^{-1} , while carbon stocks are in gC m^{-2} .

	r^2	RMSE	slope	intercept	mean	MAE	nRMSE
Ra	0.91	0.04	0.96	0.02	0.45	0.03	0.11
Rh	0.88	0.05	0.94	0.03	0.49	0.03	0.11
cLeaf	0.89	5.60	0.95	1.79	38.26	3.97	0.15
cWood	0.93	187.60	0.96	20.98	785.09	128.98	0.25
cVeg	0.93	192.51	0.97	22.29	1045.95	136.77	0.19
cLitter	0.94	49.15	0.97	7.23	290.32	31.96	0.17
cSoil	0.95	658.66	0.97	83.51	4067.64	403.08	0.16

of the western North American coastal forests, central Amazon basin, as well as in Southeast Asia.

To further characterise these differences, relative-difference maps were analysed (Figure A.10). The relative differences provide complementary information to the absolute differences by highlighting proportional deviations, particularly in low-biomass regions where small absolute differences can translate into large relative values. Comparison across the UFLUX-emulated ensemble percentiles shows that the emulated 25th-percentile field is systematically lower than ESA CCI over most regions, whereas the emulated 75th-percentile field is systematically higher, with the emulated mean providing an intermediate reference. This ordering is expected and reflects the consistent positioning of the ESA CCI product relative to the spread of UFLUX-emulated TRENDY biomass estimates, indicating that the UFLUX mean captures a central tendency across models.

3.2. Spatiotemporal variability of carbon dynamics

3.2.1. Flux interannual variability

Fig. 5 illustrates the interannual variability of global flux aggregations from UFLUX and TRENDY models, with mass conversation evaluations represented as dark lines, calculated using the carbon-balance relationships to verify consistency among GPP, RECO, and NEE.

The global aggregations of GPP and NPP from UFLUX and TRENDY were of similar magnitude, though GPP from the sum of NPP and autotrophic respiration Ra was lower by 10 PgC yr^{-1} when verifying mass conservation. TRENDY GPP values ranged from below 100 to over 150 PgC yr^{-1}, with ensemble mean estimates around 130 PgC yr^{-1}. The GPP growth rate (with a *p*-value less than 0.05) was 0.48 PgC yr^{-1} for UFLUX and 0.41 PgC yr^{-1} for TRENDY, showing similar values between the two datasets. The difference between TRENDY and UFLUX was generally within 1 PgC yr^{-1}, though in certain years, such as 2010, UFLUX exhibited greater interannual variability, reaching up to 5 PgC yr^{-1}, and approximately 90% of GPP peaks were captured in the Earth observation signals during the experiment. For NPP, the mean values for both models were approximately 60 PgC yr^{-1}, with uncertainties ranging from 30–40 PgC yr^{-1} to 50 PgC yr^{-1}. The mean differences between UFLUX and TRENDY remained below 1 PgC yr^{-1} for most years, though UFLUX consistently showed higher interannual variability (0.03 vs 0.02 for TRENDY). The NPP increase rate (*p*-value < 0.05) was 0.23 PgC yr^{-1} for UFLUX and 0.2 PgC yr^{-1} for TRENDY. As described above, when accounting for mass conservation, the sum of emulated NPP and Ra was approximately 10 PgC yr^{-1} lower than that of TRENDY and UFLUX (dark dashed line in Fig. 5), and the sum's interannual variability closely followed that of UFLUX. NPP showed no consistent systematic differences, indicating that Ra remains the primary component of focus in this comparison.

The global NEP aggregation from UFLUX was approximately 10 PgC yr^{-1} higher than TRENDY, while RECO was nearly 20 PgC yr^{-1}

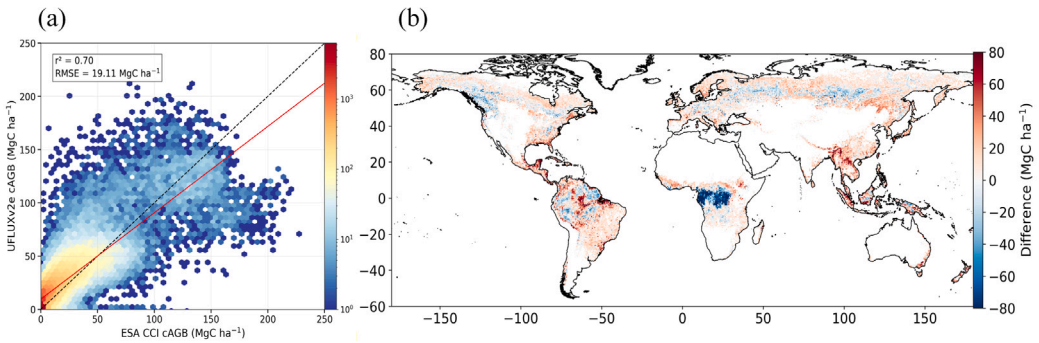

Fig. 4. Comparison of cAGB estimates between the ESA CCI and UFLUX products. (a) Scatter density plot showing the relationship between ESA CCI (x-axis) and UFLUX (y-axis) cAGB values (MgC ha^{-1}). The black dashed line represents the one-to-one relationship, and the red solid line shows the linear regression fit. (b) Spatial distribution of absolute differences in cAGB between UFLUX and ESA CCI (UFLUX – CCI, MgC ha^{-1}). Blue areas indicate regions where UFLUX estimates are lower than ESA CCI, while red areas show regions where UFLUX estimates exceed ESA CCI values.

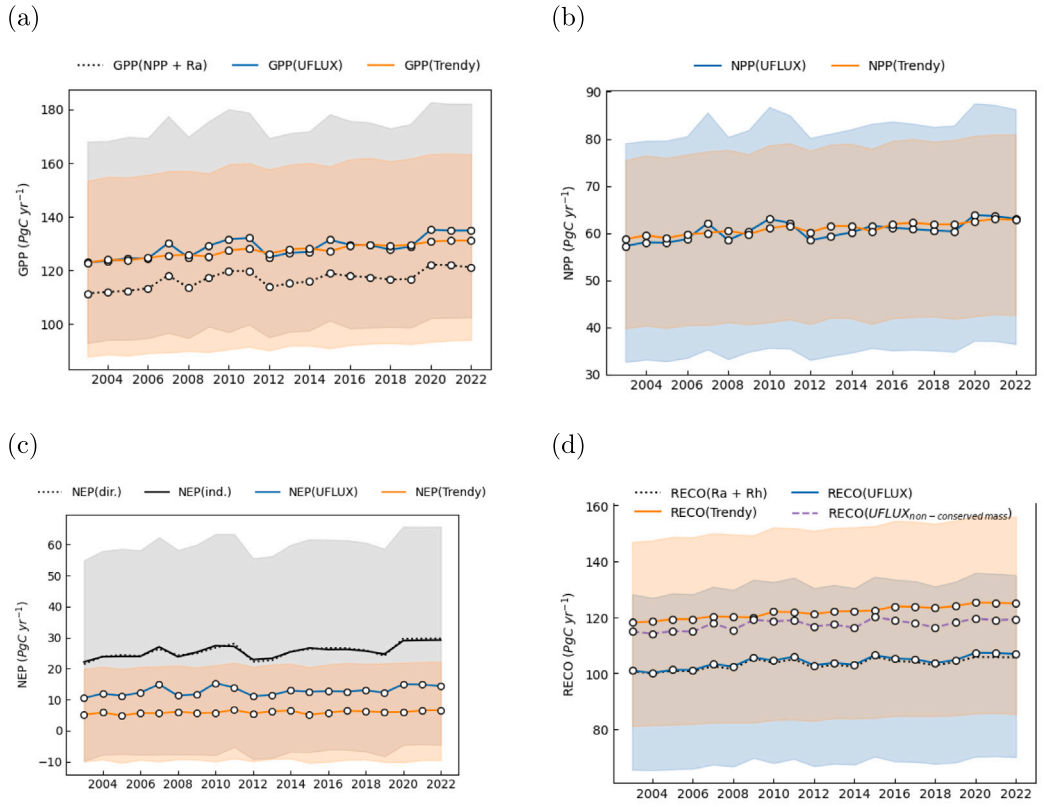

Fig. 5. Interannual variability of UFLUX (blue lines, with blue shading indicating uncertainty) compared to TRENDY models (orange lines, with orange shading indicating uncertainty) for four carbon fluxes: GPP, NPP ($GPP - Ra$), NEP ($NPP - Rh$), and RECO ($Ra + Rh$). The dark dashed lines with shaded areas represent uncertainty: for GPP and RECO, they indicate the summed fluxes from the UFLUX emulator to assess the mass conservation, while for NEP, they represent upscaled estimates from EC flux measurements. For RECO, an additional purple dashed line shows the machine-learning fit obtained without imposing mass-conservation constraints. The figure shows globally aggregated values with units of PgC yr^{-1}.

lower, with UFLUX again exhibiting greater interannual variability. NEP had a global mean of approximately 4 PgC yr^{-1} in TRENDY compared to nearly 15 PgC yr^{-1} in UFLUX. The NEP increase rate (p-value < 0.05) was 0.12 PgC yr^{-1} for UFLUX and 0.05 PgC yr^{-1} for TRENDY. The uncertainty range for TRENDY NEP spanned from −10 PgC yr^{-1} to 20 PgC yr^{-1}, encompassing the UFLUX mean estimate. RECO was approximately 120 PgC yr^{-1} in TRENDY (Rh was about 85% of Ra) and 100 PgC yr^{-1} in UFLUX, with respective uncertainty ranges of 80–140 PgC yr^{-1} and 60–120 PgC yr^{-1}, overlapping between models. The RECO increase rate (p-value < 0.05) was 0.3 PgC yr^{-1} for UFLUX and 0.37 PgC yr^{-1} for TRENDY. When mass conservation was switched off in UFLUX during modelling (purple dashed line in Fig. 5d), RECO values were much closer to those of TRENDY, with only about 5 PgC yr^{-1} lower. The interannual variability of RECO remained largely unchanged. This demonstrates UFLUX's ability to reproduce TRENDY carbon dynamics under unconstrained conditions. When mass balance was considered, NEP (as $NEP = NPP - Rh$) showed an approximate 10 PgC yr^{-1} bias, with NPP showing little difference from TRENDY (Fig. 5b), indicating that Rh is the main source of the NEP difference. Despite this bias, process-informed UFLUX predictions were closer to TRENDY NEP than purely data-driven methods (Zhu et al., 2024a; Zeng et al., 2020).

NEP derived directly from upscaled EC data (NEP (dir.), Fig. 5c) was approximately 25 PgC yr^{-1}, closely matching NEP inferred indirectly

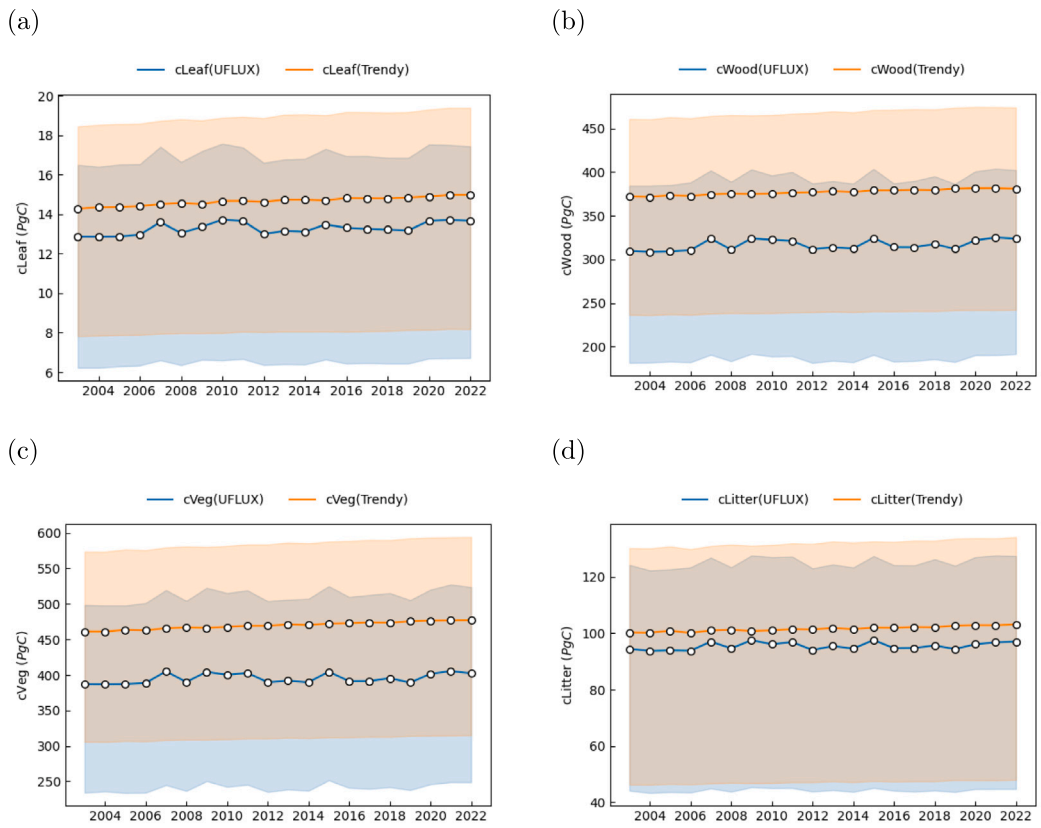

Fig. 6. Interannual variability of UFLUX (blue lines, with uncertainty indicated by blue shading) compared to TRENDY models (orange lines, with uncertainty indicated by orange shading) for four carbon stocks: cLeaf (carbon stored in the leaf pool of vegetation), cWood (carbon stored in the wood/stem pool of vegetation), cVeg (carbon stored in both aboveground and belowground vegetation biomass pools), and cLitter (carbon stored in plant litter, such as non-living biomass). The figure displays globally aggregated values in units of PgC.

from upscaled GPP and RECO (NEP (ind.), Fig. 5c) as well as values reported in pure data-driven upscaling studies (Zhu et al., 2024a; Zeng et al., 2020), showing that the UFLUX captured flux upscaling patterns accurately as well. Both direct and indirect NEP estimates were around 10 cFluxAggUnit higher than UFLUX and fell outside TRENDY's uncertainty range. UFLUX RECO was consistent in both magnitude and interannual variability with the sum of emulated Ra and Rh (Fig. 5d). Feature importance analysis indicated that photosynthetic fluxes such as NPP were mainly influenced by satellite vegetation indices (approximately 20%), with additional contributions from climate drivers. In contrast, Ra and Rh were primarily driven by air temperature (66%) and soil temperature (20%), with limited variability attributed to other factors such as functional types.

3.2.2. Stock interannual variability

Fig. 6 presents the interannual variability of global carbon stock aggregates from UFLUX and TRENDY, with mass-conserved estimates highlighted in darker colours in Figure A.13. Overall, UFLUX and TRENDY showed similar temporal patterns, but UFLUX generally estimated lower stock values and displayed higher interannual variability across most pools.

For leaf carbon (cLeaf), TRENDY averaged just above 14 PgC (range: 8–19 PgC), while UFLUX averaged approximately 13 PgC (range: 6–17 PgC), representing a 5% difference. Both models showed similar positive trends of 0.03 PgC yr^{-1} ($p < 0.05$). The UFLUX estimates were slightly lower across most years, but the year-to-year variability closely followed TRENDY, indicating a strong temporal consistency between the two models for leaf carbon.

For wood and stem carbon (cWood), which together with cLeaf form aboveground biomass, TRENDY values averaged around 370 PgC (range: 240–460 PgC), whereas UFLUX averaged 320 PgC (range: 180–400 PgC), a difference of approximately 14%. The trends differed slightly, with UFLUX increasing at 0.47 PgC yr^{-1} and TRENDY at 0.54 PgC yr^{-1} ($p < 0.05$). Compared year by year, UFLUX consistently estimated lower cWood values than TRENDY, particularly in years of high interannual variability, suggesting that differences are most pronounced in aboveground woody biomass.

For total vegetation carbon (cVeg), including leaves, stems, and roots, TRENDY averaged slightly above 450 PgC (range: 300–600 PgC), while UFLUX averaged just below 400 PgC (range: 220–530 PgC), a 13% difference. TRENDY exhibited a slightly higher trend of 0.88 PgC yr^{-1} ($p < 0.05$) compared to 0.46 PgC yr^{-1} for UFLUX ($p = 0.09$). This demonstrates general temporal consistency with TRENDY, while trends in belowground carbon like in roots, show greater uncertainty. This is consistent with the known challenges of large-scale observations of underground biomass (Spawn et al., 2020).

For litter carbon (cLitter), TRENDY averaged 100 PgC (range: 45–130 PgC), and UFLUX averaged 96 PgC (range: 40–130 PgC), corresponding to a 5% difference. Trends were 0.15 PgC yr^{-1} for TRENDY ($p < 0.05$) and 0.08 PgC yr^{-1} for UFLUX ($p = 0.11$). The differences were smaller than in aboveground pools, and interannual variability patterns were largely similar.

As one of the first large-scale observationally derived estimates of carbon fluxes and stocks, UFLUX showed some systematic differences compared to TRENDY, but overall the discrepancies fall within the range of the 21 TRENDY models, with the multi-model mean serving as a reference rather than a definitive value. Comparison with independent observational data (Fig. 4) indicated that UFLUX estimates are generally consistent with leading studies, providing long-term, large-scale coverage. Differences between UFLUX and TRENDY were mostly

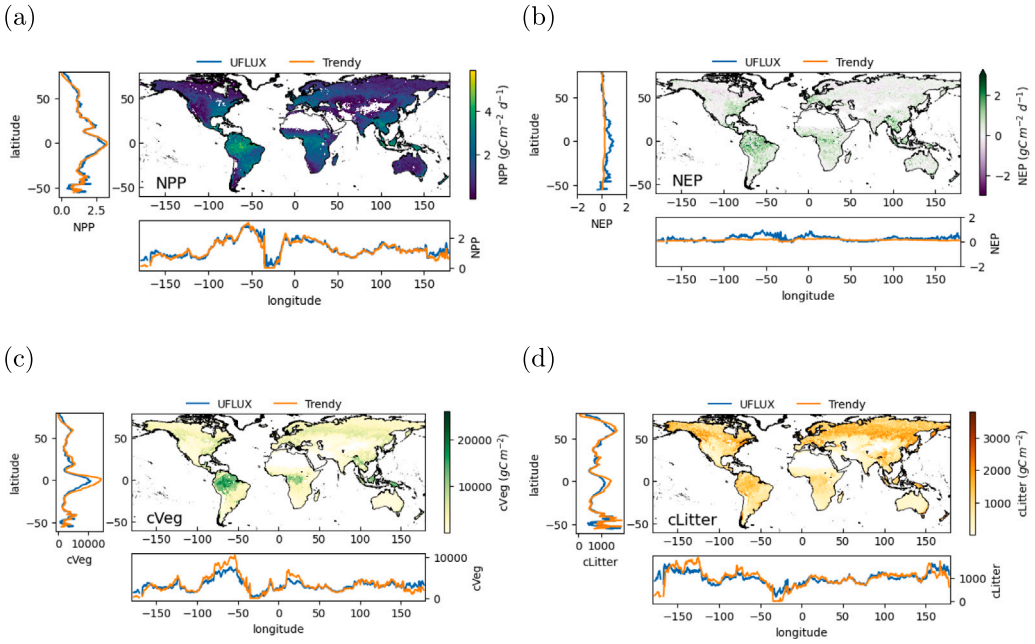

Fig. 7. The spatial distributions of fluxes (NPP in panel (a) and NEP in panel (b), unit: gC m^{-2} d^{-1}) and stocks (cVeg in panel (c) and cLitter in panel (d), unit: gC m^{-2}) for UFLUX, along with the latitudinal and longitudinal distributions for both UFLUX (blue) and TRENDY (orange), are presented next to the main frame of each subplot. These spatial distributions represent multi-year averages from 2003 to 2022.

20% or less, which is within the range of uncertainties reported in existing field and modelling studies, from 44% to over 200% (Huang et al., 2021).

When mass conservation is considered in allocating net flux NEP into carbon pools (Figure A.13), the accumulated NEP over one year approximates the total change in all pools, accounting for time lags. UFLUX results were comparable to TRENDY, with differences between pool changes and accumulated NEP of 0.4 PgC for UFLUX and 2.41 PgC for TRENDY, showing a slightly closer mass balance in UFLUX. Year-to-year changes in net carbon across pools were substantial in both models, reaching 50 PgC or more, considerably exceeding the NEP interannual variability reported in processed models.

3.2.3. Spatial distribution

Fig. 7 illustrates the spatial distributions of carbon fluxes (NPP and NEP) and carbon stocks (cVeg and cLitter) across the globe. Strong carbon uptake and storage were concentrated in the tropics, in line with the literature (Friedlingstein et al., 2024), notably the Amazon, Central Africa, and Southeast Asia, as well as in the northern mid-latitudes, including the mid-eastern United States, Europe, and eastern Asia. Both UFLUX and TRENDY captured these major spatial patterns, with latitudinal and longitudinal distributions generally aligned between the models.

For NPP, model differences were small, typically within approximately 0.1 gC m^{-2} d^{-1}, indicating strong agreement across most vegetated regions. NEP showed that most vegetated areas acted as carbon sinks, with exceptions in arid and semi-arid regions, such as the mid-western United States, southern Spain, Central Asia, and central Australia. Differences between UFLUX and TRENDY in NEP were around 2 gC m^{-2} d^{-1}, with the largest discrepancies occurring in the Amazon and Congo Basin, where UFLUX tended to simulate higher NEP than TRENDY.

Vegetation carbon stock (cVeg) exhibited similar spatial patterns, though TRENDY predicted higher values than UFLUX in the Amazon and Congo Basin by roughly one-third, with the difference of about 1.5 kgC m^{-2}. This contrast was less pronounced at higher latitudes, where model predictions were closely aligned. Litter carbon (cLitter) also showed higher TRENDY values than UFLUX, but differences in the tropics were smaller (about 0.2 kgC m^{-2}) than those for cVeg, while remaining comparable to discrepancies observed in temperate and boreal regions.

Across latitudes, the largest model differences in carbon stocks and fluxes occurred in regions of dense tropical forests, where NPP and NEP exhibited higher variability. At the global scale, UFLUX captured the overall spatial gradients in carbon fluxes and stocks, with localised differences highlighting regions of divergence between models. Despite these differences, the spatial patterns of fluxes and stocks were broadly consistent between UFLUX and TRENDY across the two-decade period.

3.3. Carbon dynamics in a changing climate

3.3.1. Continental carbon sequestration variations

Fig. 8 illustrates the statistical distributions of carbon fluxes (a) and stocks (b) across continents over a 20-year period (2003–2022).

Carbon fluxes (NEP, GPP, and RECO) exhibited distinct variations among continents. South America had the highest fluxes, followed by Africa and Asia, while Oceania exhibited the lowest values, with Europe and North America ranking in between. The fluxes in South America were approximately 20%–33% higher than those in Africa and Asia, which showed similar values, differing by less than 1 PgC yr^{-1}. Oceania's carbon fluxes were close to 0 PgC yr^{-1} across all three flux types. In Europe and North America, flux patterns varied by flux type: NEP was approximately 1–1.5 PgC yr^{-1}, about one-fourth of that in South America; GPP was around 17 PgC yr^{-1}, about half of South America's and two-thirds of Africa's and Asia's values; and RECO ranged from 15 to 18 PgC yr^{-1}, closely aligning with Africa and Asia (approximately 20 PgC yr^{-1}) but at two-thirds of South America's level. No significant decreasing trends were observed for any flux type in any continent. However, increasing trends were detected (p-values < 0.05) for NEP in Asia, as well as for GPP and RECO in Africa, Asia, and South America. The NEP trend in Asia was 0.03 PgC yr^{-1}, while the GPP trends in these three continents were nearly twice as high as those of RECO.

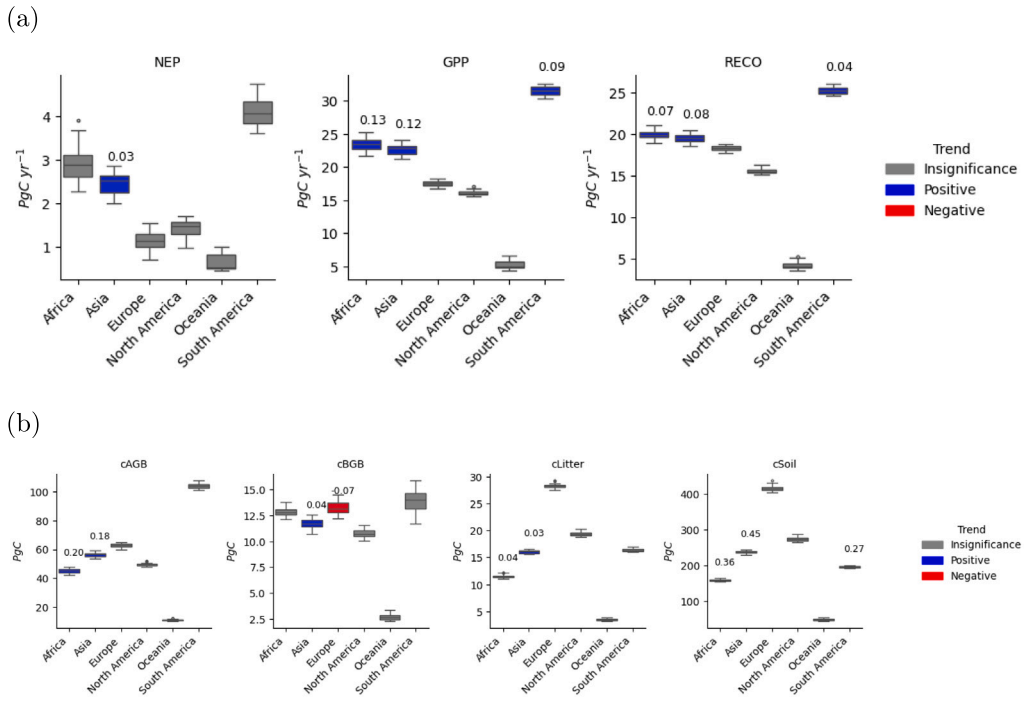

Fig. 8. Statistical distributions of carbon fluxes (a) and stocks (b) by continent (2003–2022). The boxplots represent the interannual variability, with wider boxes indicating greater year-to-year differences. Fluxes and stocks with significant positive trends (p-value < 0.05) are shown in blue, while significant negative trends are shown in red, with the numeric trend displayed above each corresponding box. Trends that are not statistically significant (p-value ≥ 0.05) are shown in grey. 'cAGB' represents the carbon stored in aboveground biomass (AGB), which is the sum of cLeaf and cWood. 'cBGB' refers to the carbon stored in belowground biomass (BGB), calculated as cVeg excluding cAGB.

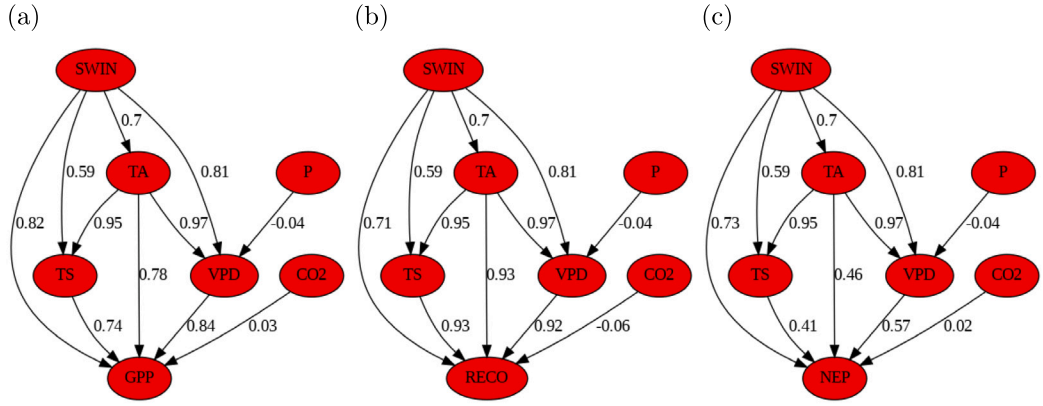

Fig. 9. Causal inferences of climate variables on GPP, RECO, and NEP in Europe. The values represent relative impact (unitless, ranging from 0 to 1). 'SWIN' denotes downward shortwave solar radiation, 'TA' is air temperature, 'P' is precipitation, 'TS' is soil temperature, 'VPD' is vapour pressure deficit, and 'CO2' represents atmospheric CO_2 concentration.

Carbon stocks (cAGB, cBGB, cLitter, and cSoil) also varied significantly across continents, though without a consistent distribution pattern. The only exception was Oceania, which consistently exhibited the lowest carbon stocks, approaching 0 PgC. For cAGB and cBGB, South America had the largest stocks: cAGB reached 100 PgC, nearly twice the amount observed in other continents, while cBGB in South America was approximately 14 PgC, comparable to values in other continents except Oceania, which had a slightly lower value (~12.5 PgC). In contrast, Europe exhibited the highest cLitter and cSoil stocks, exceeding those of other continents (except Oceania) by 30%–50%. North America ranked second in both cLitter (~20 PgC) and cSoil (~280 PgC), with values approximately 30% higher than in Africa, Asia, and South America. In terms of trends, Africa and Asia exhibited significant positive trends (p-values < 0.05) for cAGB (~0.19 gC m^{-2} d^{-1}), cLitter (~0.04 gC m^{-2} d^{-1}), and cSoil (~0.4 gC m^{-2} d^{-1}). Additionally, Asia showed a positive trend of 0.04 gC m^{-2} d^{-1} for cBGB, while Europe exhibited a negative trend of −0.07 gC m^{-2} d^{-1}. South America also displayed a significant positive trend of 0.27 gC m^{-2} d^{-1}.

3.3.2. Climatic causal inference

Fig. 9 presents the Bayesian causal inference analysis of climate variable impacts on carbon fluxes (GPP, RECO, and NEP). Europe was used as an example because it is the only continent showing a decrease in carbon dynamics (Fig. 8). Causal inferences for the other continents can be found in the Appendix (Figure A.14 for Africa, Figure A.15 for Asia, Figure A.16 for North America, Figure A.17 for South America, Figure A.18 for Oceania).

Solar radiation (SWIN), temperature (air temperature, TA; soil temperature, TS), and vapour pressure deficit (VPD) exhibited the largest

influences on carbon fluxes, whereas precipitation (P) and CO_2 concentration had the least impact on both fluxes and related climate variables. The causal effect of solar radiation on carbon fluxes was generally high (> 0.7), with a stronger influence on GPP (0.82) and a slightly lower impact on RECO (0.71). Temperature had the greatest causal effect on RECO (> 0.9), followed by GPP (0.76), while its impact on NEP was relatively minor (0.44). VPD strongly influenced GPP and RECO (0.84 and 0.92, respectively) but had a smaller effect on NEP (~0.57). In contrast, CO_2 concentration had minimal influence on all three fluxes, with an estimated impact of ~0.03 on GPP and NEP, and a slightly negative effect on RECO (−0.06).

Regarding the interdependencies among climate variables, solar radiation exhibited a large causal influence on temperature (~0.65) and VPD (0.81). Air temperature also significantly affected VPD (~0.97), whereas precipitation had a slight negative effect on VPD (−0.04).

4. Discussion

4.1. Effectiveness of UFLUX

4.1.1. Machine learning precision and baseline model optimisation

The process-informed UFLUX framework is highly effective for upscaling EC fluxes from local to global scales while advancing ecological understanding. The XGBoost machine learning algorithm underpinning SmartForest4D has demonstrated strong performance in capturing the non-linear relationships between flux residuals of the EC-baseline models and environmental drivers, particularly solar radiation, air temperature, and atmospheric dryness, quantified by VPD (Zhu et al., 2024a). Validation results indicate that UFLUX achieves robust flux estimation performance, with an r^2 exceeding 0.7 across four flux types (carbon and water fluxes) and reaching nearly 0.85 for GPP. The r^2 value reached 0.9 with deep learning models in our experiments (Zhu et al., 2024b), but an r^2 above 0.8 is still satisfactory, as deep learning may be infeasible for users with limited computing resources (Zhu et al., 2024b). Meanwhile, the associated errors remain relatively low, slightly above 1 gC m^{-2} d^{-1} for carbon fluxes and 16 W m^{-2} for water fluxes (Fig. 2), further suggesting the reliability of UFLUX across flux types. This performance is comparable to purely data-driven upscaling studies (Jung et al., 2020; Joiner and Yoshida, 2020). More importantly, UFLUX exhibited larger performance improvements in managed ecosystems such as croplands and grasslands relative to the baseline models (Section 3.1.1). This enhancement is particularly relevant because improved carbon flux estimates in these systems can provide more reliable information to inform land-management practices and policy decisions, including livestock grazing intensity and fertiliser application. These ecosystems account for approximately one third of annual terrestrial carbon fixation and occupy a comparable fraction of the global land surface, giving them a substantial role in climate-change mitigation; despite this importance, these ecosystems have historically received considerably less attention than forest systems in carbon-cycle studies (Cardenas et al., 2022).

The hybrid process-information nature of UFLUX can promote the advancement of ecological knowledge. The UFLUX surpasses the baseline models, such as the light-use efficiency model (Stocker et al., 2020), improving r^2 by over 50% and reducing error by nearly 100% (Fig. 2). Beyond predictive accuracy, the hybrid design of UFLUX also informs ecological insights from three mechanisms: correcting biases in baseline models using EC observations and upscaling these corrections globally; optionally constraining the deep-learning neural network implementation through process-informed loss functions; and learning carbon allocation parameters from process-based model ensembles (e.g. TRENDY) instead of carbon fluxes and stocks. For example, the bias correction approach highlights the necessity of optimising the semi-empirical baseline models for RECO and LE, as the Lloyd–Taylor and Priestley–Taylor models are mainly temperature driven and do not take into account the canopy resistance (Jones et al., 2024), particularly for RECO, which is a significant source of uncertainty in quantifying net carbon sequestration (Fig. 5). There is therefore an urgent need for more advanced parameterisations of ecosystem respiration that explicitly account for variability across plant functional types and climatic scenarios. In addition, the development of scalable, and where possible remotely sensed, proxies is required to improve large-scale estimates of ecosystem respiration. Such advances would enhance our understanding of respiration processes and, more broadly, ecosystem carbon dynamics under increasingly variable future climate conditions (Liu et al., 2026; Qiu et al., 2026). For instance, incorporating advanced process-based models, such as the two-pool substrate model for RECO (Jones et al., 2024), or quantifying the semi-empirical model coefficients specific to different plant functional types and environmental conditions, potentially using explainable machine learning, could be viable options. The UFLUX performance could also be enhanced by improving the climate drivers, as the current ERA5-Land data (0.1° resolution) shows variable correlation and nRMSE compared to EC observations across plant functional types (Figure A.19). Moreover, our previous field-level study at 10–30 m spatial resolution indicated that the spatial variability of estimated carbon fluxes is primarily governed by satellite-derived vegetation proxy imagery (Zhu et al., 2023b). In contrast, air and soil temperature account for more than 80% of the explanatory information for respiration flux modelling (Section 3.2.1). These findings suggest that improvements in the spatial resolution and accuracy of climate forcing data would substantially benefit flux estimation. Further details on the global validation, inter-product comparisons with existing upscaling products, and methodological advancements of process-informed UFLUX for GPP can be found in Dong et al. (2025).

4.1.2. Carbon stock emulation and consistency

The effectiveness of UFLUX extends beyond flux estimation to the representation of carbon allocation processes, as indicated by its strong validation performance in emulating carbon allocation parameters from the TRENDY ensemble and independent observational datasets. The r^2 for emulating TRENDY carbon dynamics exceeded 0.8, often reaching around 0.95 (Fig. 3 and Table 4), indicating a strong fit and suggesting that UFLUX effectively captured the variance in carbon dynamics. This is further supported by the nRMSE, which ranged from 5% to 30%, reflecting a reasonable level of agreement. Additionally, the UFLUX emulator effectively captured the uncertainty associated with carbon dynamics. From a spatial perspective, UFLUX exhibited generally consistent performance across regions. In tropical areas, such as the Amazon, r^2 values were slightly lower compared to other regions (Fig. 3b); however, the corresponding error metric, nRMSE, was also smaller, suggesting that overall model performance remained strong. This pattern may reflect the characteristics of tropical ecosystems, such as evergreen broadleaf forests, which display relatively low seasonality (Chen et al., 2020). Consequently, the slightly reduced r^2 does not imply poor performance, but rather a limitation of r^2 in capturing variability in near-constant flux regimes.

The effectiveness of UFLUX is further highlighted by its comparison to the independent ESA CCI-AGB benchmark dataset. Previous studies have reported limited consistency between satellite-based biomass products and TRENDY model estimates of terrestrial carbon stocks and sinks, reflecting differences in data sources and modelling assumptions (Randerson et al., 2025). In this context, the strong correlation (r^2 = 0.70) and moderate RMSE (19.11 MgC ha^{-1}) between UFLUX and ESA CCI cAGB estimates indicate a substantial level of agreement between two fundamentally different approaches. This agreement is particularly notable as UFLUX estimates are based on modelled carbon pools (cLeaf and cWood), whereas the CCI data consist of direct remote sensing measurements calibrated with ground reference data. However, traditional machine learning approaches, though strong in predictive accuracy, do not inherently contribute to ecological knowledge or enhance our understanding of terrestrial ecosystem responses to climate

change. In contrast, process-informed machine learning achieves a balance between accurate flux upscaling and ecological interpretability.

The strong performance of UFLUX can be attributed not only to its effectiveness but also to the relatively smooth and coarse resolution of TRENDY simulations. Compared to direct EC measurements, TRENDY model estimates tend to be smoother in both spatial (0.5° and temporal dimensions, lacking the high-frequency variability ("spiky" values) often present in observational data which was reported to provide more accurate estimates on ecosystem carbon dynamics (Randerson et al., 2025). This smoothness likely contributed to better validation statistics, as machine learning models perform better when estimating targets that align closely with the training dataset. Consequently, UFLUX exhibited stronger statistical validation performance when emulating TRENDY compared to EC data.

The differences in global aggregates between UFLUX and the TRENDY ensemble means, for example, a ~10 PgC difference in NEP (Fig. 5c), do not necessarily indicate a substantial discrepancy. This is because the ensemble mean of the 21 TRENDY models represents only the average value and does not constitute the "true" state of the system. UFLUX estimates generally fall within the uncertainty ranges of the TRENDY ensemble, whereas previous purely data-driven upscaling products often fell outside these bounds (Zeng et al., 2020) — a pattern reproduced by the purely data-driven version of UFLUX (Fig. 5c). This suggests that UFLUX achieves a closer alignment with leading process-based models, representing a notable convergence between two fundamentally different modelling approaches: observation-based, data-driven methods and knowledge-based, process-driven frameworks. While there remains a long path for UFLUX to improve and substantial knowledge gaps to address, these results demonstrate a promising and significant starting point.

4.1.3. Conservation of mass in carbon dynamics estimation

The modelling of carbon dynamics should adhere to the law of mass conservation to realistically represent ecosystem carbon cycling, a task that remains challenging. There are three aspects particularly: (1) consistency between upscaled EC NEE and the difference between upscaled GPP and RECO; (2) balance in carbon allocation, where e.g., RECO equals the sum of Ra and Rh; and (3) alignment between net carbon stock changes and net flux accumulations over the same period.

A key advantage of UFLUX is its ability to maintain mass conservation. The two NEE upscaling pathways (in aspect one above) using UFLUX were in close agreement (Figure A.12 and Fig. 5c), with an r^2 of approximately 0.7 and a linear regression slope close to one. In the case of UFLUX, it demonstrated good mass balance for NEE (Figure A.12), which is critical for subsequent carbon allocation steps. Maintaining accurate NEE/NEP fluxes is essential, as they provide the net flux from which carbon is distributed among ecosystem pools, forming the foundation for a reliable understanding of whole-ecosystem carbon dynamics. This is because EC towers measure only the net flux NEE, with GPP and RECO modelled through partitioning (Aubinet et al., 2012). Partitioning should adhere to mass conservation principles (Reichstein et al., 2005), and the same requirement applies to upscaled GPP, RECO, and NEE. However, performing partitioning after upscaling on a pixel-by-pixel basis is challenging, highlighting the importance of the NEE upscaling pathway. A key methodological question thereby arises: is it preferable to upscale NEE directly using machine learning with satellite proxies and climate drivers, or to upscale GPP and RECO separately and then calculate NEE? Upscaling NEE directly appears to offer greater accuracy, as uncertainties and errors inherent in the initial partitioning process can propagate if GPP and RECO are upscaled first. Conversely, the indirect approach of upscaling GPP and RECO separately can ensure strict mass conservation, since independently upscaled GPP and RECO do not always sum precisely to NEE.

In the context of global carbon flux allocation emulation, UFLUX demonstrated the ability to maintain mass conservation across key flux components when emulating TRENDY. For example, the sum of emulated autotrophic Ra and heterotrophic respiration Rh closely matched the EC-upscaled total ecosystem respiration RECO (Fig. 5d), ensuring consistency between NEP, GPP, and RECO in the global emulation products. This mass-conserving behaviour was critical: without explicitly enforcing mass balance, UFLUX would have produced RECO values similar to TRENDY (purple dashed line, Fig. 5d), yet this could have corresponded to a mass imbalance of approximately 20 PgC yr^{-1}. Examination of individual components revealed that while NPP estimates were good and well captured by satellite-based proxies and climate drivers, respiration estimates were more uncertain, resulting in 10 PgC yr^{-1} of the RECO imbalance attributable to Ra (Fig. 5a). This is because machine learning relied heavily on temperature information when predicting respiration, and their estimates were largely insensitive to plant functional type. Specifically, another 10 PgC yr^{-1} of the RECO imbalance could be attributed to Rh, given that mass conservation between RECO, Ra, and Rh was maintained (Fig. 5d). Considering previous upscaling products (Zeng et al., 2020), the hybrid UFLUX framework, and process-based TRENDY, GPP estimates from EC upscaling and TRENDY were generally in good agreement, as were NPP estimates between TRENDY and UFLUX, and RECO estimates from EC and UFLUX. Nevertheless, NEP derived from EC upscaling remained approximately 20 PgC yr^{-1} higher than TRENDY, primarily reflecting uncertainties in respiration. By accounting for these imbalances, UFLUX reduced the NEP discrepancy to around 10 PgC yr^{-1}, representing a marked improvement over purely data-driven upscaling approaches and demonstrating its ability to incorporate mechanistic constraints into global carbon flux emulation.

Another important aspect of carbon cycle modelling is the ability to conserve mass when linking annual net fluxes NEP to changes in carbon stocks across years. Ideally, the accumulation of NEP over a year should correspond closely to interannual changes in carbon pools, although some tolerance is expected due to delays in carbon allocation and accumulation. In this context, UFLUX also demonstrated robust mass-conserving behaviour. Both UFLUX and TRENDY exhibited comparable performance in relating NEP to carbon stock changes (Figure A.13), but UFLUX showed a notable improvement. The multi-year average difference, or imbalance, between NEP and carbon stock changes (i.e., 2.4 PgC yr^{-1} for TRENDY, 0.4 PgC yr^{-1} for UFLUX), representing approximately a six-fold reduction. When considering the net flux magnitude, with TRENDY's annual NEP around 6 PgC yr^{-1} and UFLUX's nearly 13 PgC yr^{-1}, the relative imbalance rate in UFLUX remained substantially lower, highlighting its improved mass-conserving performance. This improved performance in UFLUX is not merely a technical detail; it has meaningful implications for ecosystem and climate studies. Ensuring accurate mass conservation in flux accumulation is critical, as it allows long-term fluxes to translate consistently into carbon stock dynamics, with direct implications for vegetation growth, mortality, and ecosystem responses to future climate variability. By reducing mass imbalances, UFLUX provides a more consistent framework for integrating flux-based observations into projections of ecosystem carbon storage, enhancing both predictive reliability and ecological interpretability.

4.2. Spatiotemporal variability of carbon dynamics

4.2.1. Improving respiration modelling for spatial consistency across methods

The UFLUX exhibited strong agreement with TRENDY models regarding the spatial distribution of carbon fluxes and stocks; however, notable differences exist in respiration, particularly in tropical regions. These differences were likely due to the simplified respiration semi-empirical models used in both process-based models and EC partitioning methods, which do not account for the complexity of respiration dynamics on a case-by-case basis — such as the nighttime decline in plant respiration (Reichstein et al., 2005; Lasslop et al., 2010;

Jones et al., 2024). The spatial distribution of NPP between UFLUX and TRENDY is nearly identical, a similarity further confirmed by their longitudinal and latitudinal patterns (Fig. 7a). This strong spatial agreement extends to other carbon stocks and fluxes (Fig. 7), though with some discrepancies, particularly in the Amazon region. These differences in carbon stocks can be attributed to variations in net ecosystem production NEP allocation. The strong agreement in NPP and the relatively weaker agreement in NEP suggest that the primary discrepancy lies in Rh, which subsequently affects carbon stock estimations. Rh is difficult to model due to its dependence on complex, variable factors like microbial activity, soil conditions, and substrate availability, as well as challenges in measuring it directly (Davidson and Janssens, 2006; Todd-Brown et al., 2013).

While Rh exhibited as a main source uncertainty for carbon allocation modelling, autotrophic respiration Ra could also play a critical role in shaping carbon flux dynamics and warrants careful consideration. When Ra and Rh are considered together, their combined contribution determines total ecosystem respiration RECO and strongly influences NEP and carbon stock estimates. Differences in Ra can arise from variations in plant physiology, phenology, and allocation strategies, as well as from simplifications in process-based or semi-empirical models that do not fully capture the dynamic response of plant respiration to environmental conditions. Similar to Rh, Ra is challenging to measure directly, particularly across diverse ecosystems and temporal scales, which introduces additional uncertainty in global carbon flux assessments (Reichstein et al., 2005; Lasslop et al., 2010). Improving the representation of Ra is therefore essential for enhancing model fidelity, as inaccuracies in Ra propagate through NEP estimates and subsequently affect carbon stock calculations. By refining both Ra and Rh, models like UFLUX can better reconcile flux-based observations with process-based expectations, ultimately supporting more robust estimates of ecosystem carbon dynamics and their spatial variability.

4.2.2. Temporal global carbon aggregates diverge due to respiration modelling differences

Beyond spatial patterns, UFLUX and TRENDY also exhibited strong agreement in interannual variability, particularly in long-term trends, which are critical for understanding and projecting carbon cycle dynamics. The observed differences in respiration likely stem from discrepancies between EC measurements and process-based model estimates in terms of magnitude. Notably, UFLUX and other data-driven, observation-based upscaling methods tend to estimate carbon fluxes with greater interannual variability than process-based models (Figs. 5 and 6). The high interannual variability in UFLUX aligns with previous studies (Zeng et al., 2020; Joiner and Yoshida, 2020; Zhu et al., 2024a,b), which have linked such variability to fluctuations in satellite-based vegetation proxies and climate variables (Hersbach et al., 2020).

In terms of GPP, Fig. 5a and b show that UFLUX-based GPP, derived from upscaling baseline model correction with FLUXNET2015 EC data, and NPP emulated from TRENDY, were in strong agreement with TRENDY in both magnitude and trend. However, the sum of UFLUX-based NPP and Ra, which should theoretically equal GPP, is lower than the EC-derived GPP, indicating a discrepancy of approximately 10 PgC yr^{-1} in Ra as UFLUX NPP aligned with process-based models. Additionally, as previously discussed, differences in Rh contribute further to discrepancies, as seen in Fig. 5d, where the difference in total ecosystem respiration RECO is approximately 20 PgC yr^{-1}. This suggests that Rh accounts for another 10 PgC yr^{-1} of the discrepancy. These differences explain the variations observed between EC-based, UFLUX, and TRENDY NEP estimates (Fig. 5c).

The NEP estimates derived from the pure data-driven version of UFLUX were in strong agreement with other EC upscaling studies (Zeng et al., 2020), yet they were approximately 20 PgC yr^{-1} higher than TRENDY estimates. This suggests a systematic discrepancy of around 20 PgC yr^{-1} between observationally constrained EC-based estimates and process-based model simulations of global land carbon sequestration. Given that mass conservation was well-maintained when upscaling GPP, RECO, and NEP (Figure A.12), this discrepancy is likely attributable to an incomplete understanding of ecosystem respiration processes in process-based models.

Based on the above, observational perspectives, consistent with process-based models, indicate that global gross carbon uptake GPP is increasing at a rate exceeding 0.4 PgC yr^{-1} under current climate conditions. This increase surpasses the growth rate of ecosystem respiration RECO, resulting in a modest net increase in carbon uptake, though with variations across different estimation methods. Consequently, carbon stocks are also increasing (Figs. 5 and 6). The magnitude of carbon stocks across continents aligns with existing knowledge (Friedlingstein et al., 2024), with substantial portions of terrestrial carbon stocks exhibiting an upward trend (Fig. 8), suggesting that terrestrial ecosystems are adapting to climate change. However, an exception is observed in Europe, where belowground biomass shows a concerning declining trend (Fig. 8b).

4.2.3. Ecosystem responses and adaptation to future climate change

UFLUX illuminates global ecosystem responses to climate change, revealing both adaptive strengths and critical vulnerabilities. The global rise in GPP to 135 PgC yr^{-1} in 2022, increasing at 0.4 PgC yr^{-1}, signals robust photosynthetic enhancement. Bayesian causal inference attributes this to rising temperatures (relative impacts of 0.8 across continents), which accelerate metabolic processes, and CO_2 fertilisation (0.03), which boosts carbon assimilation (Sitch et al., 2005; Shekhar et al., 2024). Yet, the significant divergence in NEP estimates — 16 PgC yr^{-1} (UFLUX), 6 PgC yr^{-1} (process-based), 26 PgC yr^{-1} (EC upscaling) — underscores a critical weakness: oversimplified respiration modelling. Temperature-driven respiration increases (0.89–0.92 impact) erode net carbon gains, suggesting models fail to capture complex autotrophic and heterotrophic dynamics, such as nighttime respiration declines.

Regionally, Asia's NEP growth (0.03 PgC yr^{-1}) reflects a favourable balance, with GPP surpassing respiration. Positive drying (VPD 0.83) and CO_2 fertilisation effects (+0.08) support this resilience. In contrast, South America and Africa, despite similar GPP increases (0.13 and 0.09 PgC yr^{-1}, respectively), show constrained NEP due to elevated respiration and drying (VPD 0.12–0.14 positive contribution), especially in the Amazon and savannas. North America's stable NEP balances CO_2 fertilisation against respiration, while Oceania's negligible uptake highlights arid constraints. Europe's declining belowground biomass (−0.07 PgC yr^{-1}) is alarming, as respiration losses (−0.06) outweigh CO_2 benefits (+0.03), possibly exacerbated by soil moisture deficits.

Globally, rising biomass stocks (300 PgC aboveground, 100 PgC belowground, +0.5 PgC yr^{-1}) indicate adaptation, but stable soil/litter pools suggest limited sequestration potential. High respiration in tropics risks tipping points (Lenton et al., 2024). UFLUX's 70% mass conservation validates findings, yet TRENDY's coarse resolution may obscure variability. Enhanced respiration models and expanded EC networks are essential to refine predictions and support tailored climate strategies.

4.3. Limitations and prospective advancements

UFLUX has significantly reduced uncertainty in estimating global carbon dynamics (Figure A.11) and has demonstrated strong agreement with TRENDY models. However, this agreement is not yet sufficient for fully constraining the carbon cycle. While GPP and net primary production NPP align well between UFLUX and TRENDY, substantial uncertainties remain in respiration estimates, with discrepancies of approximately 20 PgC yr^{-1} (around 20%). The overestimation of modelled respiration, partly due to the nighttime decline in plant respiration, is poorly represented in current models and aligns with our findings (Jones et al., 2024). Addressing these differences requires an improved representation of both vegetation and microbial respiration

in models to enhance consistency between observational data and process-based simulations. Besides, improving the representation of climate drivers at fine spatial resolution would also be beneficial.

Another critical limitation concerns the precise quantification of global net ecosystem carbon uptake, which is essential for informing climate mitigation strategies, particularly nature-based solutions (Griscom et al., 2017). Resolving this issue necessitates a more accurate assessment of ecosystem respiration. One potential approach is expanding the spatial coverage of EC sites in a cost-effective manner, particularly in underrepresented regions (Hill et al., 2017; Zhu et al., 2024a, 2023b).

Furthermore, while UFLUX incorporates land management practices such as livestock grazing, as represented in the TRENDY models (Friedlingstein et al., 2024), it does not yet account for other key human activities, including afforestation, deforestation, and agricultural practices such as fertiliser and pesticide application (He et al., 2025). Given that human land-use activities are major drivers of climate change, integrating these factors into UFLUX is essential for advancing sustainable land management strategies (Beillouin et al., 2022). Future efforts should also aim to validate and refine the representation of livestock grazing impacts within the framework. Given UFLUX's ability to apply learned carbon allocation to both EC-based flux measurements and upscaled estimates, a natural next step is to extend its application to the field scale. This would provide a comprehensive view of ecosystem carbon dynamics, supporting sustainable land management and agricultural practices. To this end, we are developing decision-support tools for farmers and land managers in Southwest England as a pilot study. UFLUX is designed to be simple and computationally efficient, ensuring that these tools remain low-cost and accessible for users worldwide while delivering actionable insights for local management decisions.

To further improve the accuracy of carbon dynamics representation — particularly in understanding respiration processes and net carbon uptake — UFLUX should be extended to better emulate internal carbon fluxes between pools. A promising direction involves incorporating more detailed carbon pool dynamics, such as those captured by the Data Assimilation Linked Ecosystem Carbon (DALEC) model (Williams et al., 2005). This could be achieved by optimising model parameters using data assimilation techniques, such as the CARbon DAta MOdel FraMework (CARDAMOM) (Bloom et al., 2016), thereby improving predictive accuracy and enhancing our understanding of global carbon cycling. In addition, many field measurements, such as biomass inventories, are collected intermittently rather than continuously like EC towers. To address this, we are developing data assimilation tools that integrate these discrete ground observations to calibrate the UFLUX model, improving our understanding of ecosystem carbon dynamics and their responses to novel climate events and land management scenarios.

5. Conclusion

This study used a process-informed machine learning approach, UFLUX, to bridge the gap between process-based modelling and observational EC flux measurements in understanding terrestrial ecosystem carbon dynamics. The UFLUX enhances our understanding of global carbon dynamics by providing an observational perspective that integrates EC measurements with large-scale carbon cycle modelling. Unlike traditional methods that focus primarily on fluxes, UFLUX is the first framework to enable EC-based estimations of carbon stocks, offering a novel approach to assessing carbon accumulation and allocation processes. The goal of this study is not to replicate the carbon allocation of process-based models (e.g., TRENDY), but rather to learn the knowledge embedded in allocation parameters to enhance EC tower measurements and upscaled data. This approach could provide land users and managers with a comprehensive view of their land's carbon dynamics, supporting informed management decisions in a cost-effective and user-friendly manner. By incorporating three types of mass conservation constraints and machine learning techniques, UFLUX allows for a more comprehensive evaluation of ecosystem carbon budgets, improving the consistency between flux observations and stock changes. While good agreement was achieved in gross carbon uptake (ca. 130 PgC yr^{-1} with an increase rate of 0.46 PgC yr^{-1}), divergence remains in net carbon sequestration, with a discrepancy of about 10 PgC yr^{-1}. The primary source of this uncertainty lies in the understanding of plant autotrophic and soil heterotrophic respiration. Given the competing impacts of climate change, such as increasing carbon uptake and rising ecosystem respiration RECO, improved parameterisations for respiration processes are needed. Although carbon stocks are currently increasing, indicating that ecosystems are adapting to climate change, caution is warranted, as the balance could shift in the future.

CRediT authorship contribution statement

Songyan Zhu: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Writing – original draft, Writing – review & editing. **Wenquan Dong:** Data curation, Formal analysis, Methodology, Resources, Validation, Visualization, Writing – original draft. **Vasileios Myrgiotis:** Formal analysis, Writing – review & editing. **Jian Xu:** Formal analysis, Writing – review & editing. **Pablo Reyes-Muñoz:** Formal analysis, Writing – review & editing. **Mana Gharun:** Formal analysis, Writing – review & editing. **Rui Ma:** Formal analysis, Writing – review & editing. **Man Chen:** Formal analysis, Writing – review & editing. **Jadu Dash:** Formal analysis, Writing – review & editing.

Declaration of competing interest

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

Acknowledgment

The authors thank the TRENDY modeling team for providing the model outputs.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.agrformet.2026.111197.

Data availability

All data and code used in this study are archived at Zenodo and are publicly available at: https://doi.org/10.5281/zenodo.16202323.

Model comparison used the TRENDY ensemble data described in Friedlingstein et al. (2023), available at: https://doi.org/10.5194/essd-15-5301-2023.

References

Araza, A., De Bruin, S., Herold, M., Quegan, S., Labriere, N., Rodriguez-Veiga, P., Avitabile, V., Santoro, M., Mitchard, E.T., Ryan, C.M., et al., 2022. A comprehensive framework for assessing the accuracy and uncertainty of global above-ground biomass maps. Remote Sens. Environ. 272, 112917.

Aubinet, M., Vesala, T., Papale, D., 2012. Eddy covariance: a practical guide to measurement and data analysis. Springer Science & Business Media.

Badgley, G., Anderegg, L.D., Berry, J.A., Field, C.B., 2019. Terrestrial gross primary production: Using NIRV to scale from site to globe. Global Change Biol. 25 (11), 3731–3740.

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

Baldocchi, D.D., 2020. How eddy covariance flux measurements have contributed to our understanding of global change biology. Global Change Biol. 26 (1), 242–260.

Bao, S., Wutzler, T., Koirala, S., Cuntz, M., Ibrom, A., Besnard, S., Walther, S., Šigut, L., Moreno, A., Weber, U., et al., 2022. Environment-sensitivity functions for gross primary productivity in light use efficiency models. Agricult. Forest. Meterol. 312, 108708.

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci. Data 5 (1), 1–12.

Beillouin, D., Cardinael, R., Berre, D., Boyer, A., Corbeels, M., Fallot, A., Feder, F., Demenois, J., 2022. A global overview of studies about land management, land-use change, and climate change effects on soil organic carbon. Global Change Biol. 28 (4), 1690–1702.

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

Bugnot, A., Mayer-Pinto, M., Airoldi, L., Heery, E., Johnston, E., Critchley, L., Strain, E., Morris, R., Loke, L., Bishop, M., et al., 2021. Current and projected global extent of marine built structures. Nat. Sustain. 4 (1), 33–41.

Cardenas, L., Olde, L., Loick, N., Griffith, B., Hill, T., Evans, J., Cowan, N., Segura, C., Sint, H., Harris, P., et al., 2022. CO2 fluxes from three different temperate grazed pastures using eddy covariance measurements. Sci. Total Environ. 831, 154819.

Chapin III, F.S., Matson, P.A., Vitousek, P., 2011. Principles of Terrestrial Ecosystem Ecology. Springer Science & Business Media.

Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining. pp. 785–794.

Chen, X., Maignan, F., Viovy, N., Bastos, A., Goll, D., Wu, J., Liu, L., Yue, C., Peng, S., Yuan, W., et al., 2020. Novel representation of leaf phenology improves simulation of amazonian evergreen forest photosynthesis in a land surface model. J. Adv. Model. Earth Syst. 12 (1), e2018MS001565.

Davidson, E.A., Janssens, I.A., 2006. Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. Nature 440 (7081), 165–173.

Delire, C., Séférian, R., Decharme, B., Alkama, R., Calvet, J.-C., Carrer, D., Gibelin, A.-L., Joetzjer, E., Morel, X., Rocher, M., et al., 2020. The global land carbon cycle simulated with ISBA-CTRIP: Improvements over the last decade. J. Adv. Model. Earth Syst. 12 (9), e2019MS001886.

Dong, W., Zhu, S., Xu, J., Ryan, C.M., Chen, M., Zeng, J., Yu, H., Cao, C., Shi, J., 2025. UFLUX v2. 0: A process-informed machine learning framework for efficient and explainable modelling of terrestrial carbon uptake. IEEE Geosci. Remote. Sens. Lett..

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., et al., 2007. The shuttle radar topography mission. Rev. Geophys. 45 (2).

Fisher, J.B., Tu, K.P., Baldocchi, D.D., 2008. Global estimates of the land–atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites. Remote Sens. Environ. 112 (3), 901–919.

Friedl, M., Sulla-Menashe, D., 2019. MCD12q1 MODIS/Terra+ aqua land cover type yearly L3 global 500m sin grid V006. NASA EOSDIS Land Processes Distributed Active Archive Center.

Friedlingstein, P., O'Sullivan, M., Jones, M.W., Andrew, R.M., Bakker, D.C.E., Hauck, J., Landschützer, P., Le Quéré, C., Luijkx, I.T., Peters, G.P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J.G., Ciais, P., Jackson, R.B., Alin, S.R., Anthoni, P., Barbero, L., Bates, N.R., Becker, M., Bellouin, N., Decharme, B., Bopp, L., Brasika, I.B.M., Cadule, P., Chamberlain, M.A., Chandra, N., Chau, T.-T.-T., Chevallier, F., Chini, L.P., Cronin, M., Dou, X., Enyo, K., Evans, W., Falk, S., Feely, R.A., Feng, L., Ford, D.J., Gasser, T., Ghattas, J., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Heinke, J., Houghton, R.A., Hurtt, G.C., Iida, Y., Ilyina, T., Jacobson, A.R., Jain, A., Jarníková, T., Jersild, A., Jiang, F., Jin, Z., Joos, F., Kato, E., Keeling, R.F., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J.I., Körtzinger, A., Lan, X., Lefèvre, N., Li, H., Liu, J., Liu, Z., Ma, L., Marland, G., Mayot, N., McGuire, P.C., McKinley, G.A., Meyer, G., Morgan, E.J., Munro, D.R., Nakaoka, S.-I., Niwa, Y., O'Brien, K.M., Olsen, A., Omar, A.M., Ono, T., Paulsen, M., Pierrot, D., Pocock, K., Poulter, B., Powis, C.M., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T.M., Schwinger, J., Séférian, R., Smallman, T.L., Smith, S.M., Sospedra-Alfonso, R., Sun, Q., Sutton, A.J., Sweeney, C., Takao, S., Tans, P.P., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., van der Werf, G.R., van Ooijen, E., Wanninkhof, R., Watanabe, M., Wimart-Rousseau, C., Yang, D., Yang, X., Yuan, W., Yue, X., Zaehle, S., Zeng, J., Zheng, B., 2023. Global carbon budget 2023. Earth Syst. Sci. Data 15 (12), 5301–5369. http://dx.doi.org/10.5194/essd-15-5301-2023, URL https://essd.copernicus.org/articles/15/5301/2023/.

Friedlingstein, P., O'Sullivan, M., Jones, M.W., Andrew, R.M., Hauck, J., Landschützer, P., Le Quéré, C., Li, H., Luijkx, I.T., Olsen, A., et al., 2024. Global carbon budget 2024. Earth Syst. Sci. Data Discuss. 2024, 1–133.

Geisser, S., 1975. The predictive sample reuse method with applications. J. Amer. Statist. Assoc. 70 (350), 320–328.

Gowik, U., Westhoff, P., 2011. The path from C3 to C4 photosynthesis. Plant Physiol. 155 (1), 56–63.

Graven, H.D., Warren, H., Gibbs, H.K., Khatiwala, S., Koven, C., Lester, J., Levin, I., Spawn-Lee, S.A., Wieder, W., 2024. Bomb radiocarbon evidence for strong global carbon uptake and turnover in terrestrial vegetation. Science 384 (6702), 1335–1339.

Grinsztajn, L., Oyallon, E., Varoquaux, G., 2022. Why do tree-based models still outperform deep learning on typical tabular data? Adv. Neural Inf. Process. Syst. 35, 507–520.

Griscom, B.W., Adams, J., Ellis, P.W., Houghton, R.A., Lomax, G., Miteva, D.A., Schlesinger, W.H., Shoch, D., Siikamäki, J.V., Smith, P., et al., 2017. Natural climate solutions. Proc. Natl. Acad. Sci. 114 (44), 11645–11650.

Haverd, V., Smith, B., Nieradzik, L., Briggs, P.R., Woodgate, W., Trudinger, C.M., Canadell, J.G., Cuntz, M., 2018. A new version of the CABLE land surface model (subversion revision r4601) incorporating land use and land cover change, woody vegetation demography, and a novel optimisation-based approach to plant coordination of photosynthesis. Geosci. Model. Dev. 11 (7), 2995–3026.

He, J., Li, L., Xu, J., Zheng, C., 2020. Relu deep neural networks and linear finite elements. J. Comput. Math. 38 (3), 502–527.

He, P., Ren, J., Ye, N., Chao, J., Zeng, Y., Li, J., Zhang, Q., Xiao, J., Zhu, S., Ma, J., 2025. Unanticipated cooling gains from afforestation in southwestern China. Environ. Res. Lett. 20 (12), 124013.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 770–778.

Heinke, J., Rolinski, S., Müller, C., 2022. Modeling the role of livestock grazing in c and n cycling in grasslands with lpjml5. 0-grazing. Geosci. Model. Dev. Discuss. 2022, 1–31.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146 (730), 1999–2049.

Heskel, M.A., O'sullivan, O.S., Reich, P.B., Tjoelker, M.G., Weerasinghe, L.K., Penillard, A., Egerton, J.J., Creek, D., Bloomfield, K.J., Xiang, J., et al., 2016. Convergence in the temperature response of leaf respiration across biomes and plant functional types. Proc. Natl. Acad. Sci. 113 (14), 3832–3837.

Hill, T., Chocholek, M., Clement, R., 2017. The case for increasing the statistical power of eddy covariance ecosystem studies: why, where and how? Global Change Biol. 23 (6), 2154–2165.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.

Huang, Y., Ciais, P., Santoro, M., Makowski, D., Chave, J., Schepaschenko, D., Abramoff, R.Z., Goll, D.S., Yang, H., Chen, Y., et al., 2021. A global map of root biomass across the world's forests. Earth Syst. Sci. Data 13 (9), 4263–4274.

Imambi, S., Prakash, K.B., Kanagachidambaresan, G., 2021. Pytorch. In: Programming with TensorFlow: Solution for Edge Computing Applications. Springer, pp. 87–104.

IPCC, 2006. 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Volume 4: Agriculture, Forestry and Other Land Use. Institute for Global Environmental Strategies (IGES), Hayama, Japan.

Joiner, J., Yoshida, Y., 2020. Satellite-based reflectances capture large fraction of variability in global gross primary production (GPP) at weekly time scales. Agricult. Forest. Meterol. 291, 108092.

Jones, S., Mercado, L.M., Bruhn, D., Raoult, N., Cox, P.M., 2024. Night-time decline in plant respiration is consistent with substrate depletion. Commun. Earth Environ. 5 (1), 148.

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P., Besnard, S., Bodesheim, P., Carvalhais, N., et al., 2020. Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach. Biogeosciences 17 (5), 1343–1365.

Kato, E., Kinoshita, T., Ito, A., Kawamiya, M., Yamagata, Y., 2013. Evaluation of spatially explicit emission scenario of land-use change and biomass burning using a process-based biogeochemical model. J. Land Use Sci. 8 (1), 104–122.

Kirschbaum, M., Eamus, D., Gifford, R., Roxburgh, S., Sands, P., 2001. Definitions of some ecological terms commonly used in carbon accounting. Coop. Res. Cent. Carbon Account. Canberra 2–5.

Koot, P., Mendoza-Lugo, M.A., Paprotny, D., Morales-Nápoles, O., Ragno, E., Worm, D.T., 2023. Pybanshee version (1.0): A python implementation of the MATLAB toolbox BANSHEE for non-parametric Bayesian networks with updated features. SoftwareX 21, 101279.

Lamahewage, S.H.G., Witharana, C., Riemann, R., Fahey, R., Worthley, T., 2025. Aboveground biomass estimation using multimodal remote sensing observations and machine learning in mixed temperate forest. Sci. Rep. 15 (1), 31120.

Lasslop, G., Reichstein, M., Papale, D., Richardson, A.D., Arneth, A., Barr, A., Stoy, P., Wohlfahrt, G., 2010. Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation. Global Change Biol. 16 (1), 187–208.

Lawrence, D.M., Fisher, R.A., Koven, C.D., Oleson, K.W., Swenson, S.C., Bonan, G., Collier, N., Ghimire, B., Van Kampenhout, L., Kennedy, D., et al., 2019. The community land model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. J. Adv. Model. Earth Syst. 11 (12), 4245–4287.

Lenton, T.M., Abrams, J.F., Bartsch, A., Bathiany, S., Boulton, C.A., Buxton, J.E., Conversi, A., Cunliffe, A.M., Hebden, S., Lavergne, T., et al., 2024. Remotely sensing potential climate change tipping points across scales. Nat. Commun. 15 (1), 343.

Li, F., Ding, P., Mealli, F., 2023. Bayesian causal inference: a critical review. Phil. Trans. R. Soc. A 381 (2247), 20220153.
Lienert, S., Joos, F., 2018. A Bayesian ensemble data assimilation to constrain model parameters and land-use carbon emissions. Biogeosciences 15 (9), 2909–2930.
Liu, J., Li, M., Pham, Q.B., Liu, T., Ren, Y., Shalamzari, M.J., He, P., 2026. From historical to future scenarios: A deep learning and dynamical detection framework for compound drought and heatwave events. Expert Syst. Appl. 131105.
Lloyd, J., Taylor, J., 1994. On the temperature dependence of soil respiration. Funct. Ecol. 315–323.
Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. 30.
Luo, X., Zhou, H., Satriawan, T.W., Tian, J., Zhao, R., Keenan, T.F., Griffith, D.M., Sitch, S., Smith, N.G., Still, C.J., 2024. Mapping the global distribution of C4 vegetation using observations and optimality theory. Nat. Commun. 15 (1), 1219.
Ma, L., Hurtt, G., Ott, L., Sahajpal, R., Fisk, J., Lamb, R., Tang, H., Flanagan, S., Chini, L., Chatterjee, A., et al., 2022. Global evaluation of the ecosystem demography model (ED v3. 0). Geosci. Model. Dev. 15 (5), 1971–1994.
Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., et al., 2019. Developments in the MPI-m earth system model version 1.2 (MPI-esm1. 2) and its response to increasing CO2. J. Adv. Model. Earth Syst. 11 (4), 998–1038.
Melton, J.R., Arora, V.K., Wisernig-Cojoc, E., Seiler, C., Fortier, M., Chan, E., Teckentrup, L., 2020. CLASSIC v1. 0: the open-source community successor to the Canadian land surface scheme (CLASS) and the Canadian terrestrial ecosystem model (CTEM)–part 1: Model framework and site-level performance. Geosci. Model. Dev. 13 (6), 2825–2850.
Nelson, J.A., Walther, S., Gans, F., Kraft, B., Weber, U., Novick, K., Buchmann, N., Migliavacca, M., Wohlfahrt, G., Šigut, L., et al., 2024. X-BASE: the first terrestrial carbon and water flux products from an extended data-driven scaling framework, FLUXCOM-x. Biogeosciences 21 (22), 5079–5115.
Organisation for Economic Co-operation and Development (OECD), 2024. Progress in national climate policy efforts remains insufficient to achieve 2030 targets. (Accessed: 12 April 2025).
Palviainen, M., Laurén, A., Pumpanen, J., Bergeron, Y., Bond-Lamberty, B., Larjavaara, M., Kashian, D., Köster, K., Prokushkin, A., Chen, H., et al., 2020. Decadal-scale recovery of carbon stocks after wildfires throughout the boreal forests. Glob. Biogeochem. Cycles 34 (8), e2020GB006612.
Pan, J., 2021. Can machines learn with hard constraints? Nat. Comput. Sci. 1 (4), 244.
Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., et al., 2020. The FLUXNET2015 dataset and the oneflux processing pipeline for eddy covariance data. Sci. Data 7 (1), 225.
Poulter, B., Frank, D., Hodson, E., Zimmermann, N., 2011. Impacts of land cover and climate data selection on understanding terrestrial carbon dynamics and the CO 2 airborne fraction. Biogeosciences 8 (8), 2027–2036.
Praveen, M., 2021. Torch. Autograd. Gradcheck support for tensor like objects in pytorch. California State University, Sacramento.
Prentice, I.C., Dong, N., Gleason, S.M., Maire, V., Wright, I.J., 2014. Balancing the costs of carbon gain and water transport: testing a new theoretical framework for plant functional ecology. Ecol. Lett. 17 (1), 82–91.
Priestley, C.H.B., Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Mon. Weather Rev. 100 (2), 81–92.
Qiu, K., Zhang, Q., Xie, Y., Shi, M., Wang, C., Dong, T., Ma, J., He, P., 2026. Desertification expansion significantly suppresses photosynthetic peak capacity of arid ecosystems at the global scale. Int. J. Appl. Earth Obs. Geoinf. 146, 105042.
Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J. Comput. Phys. 378, 686–707.
Randerson, J.T., Li, Y., Fu, W., Primeau, F., Kim, J.E., Mu, M., Hoffman, F.M., Trugman, A.T., Yang, L., Wu, C., et al., 2025. The weak land carbon sink hypothesis. Sci. Adv. 11 (37), eadr5489.
Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, F., 2019. Deep learning and process understanding for data-driven earth system science. Nature 566 (7743), 195–204.
Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., et al., 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Global Change Biol. 11 (9), 1424–1439.
Santoro, M., Cartus, O., 2024. ESA biomass climate change initiative (biomass_cci): Global datasets of forest above-ground biomass for the years 2010, 2015, 2016, 2017, 2018, 2019, 2020 and 2021, v5.01. http://dx.doi.org/10.5285/bf535053562141c6bb7ad831f5998d77.
Schaaf, C., Strahler, A., Chopping, M., Gao, F., Hall, D., Jin, Y., Liang, S., Nightingale, J., Román, M., Roy, D., et al., 2021. MODIS MCD43 product user guide V005. University of Massachusetts Boston, Boston: Boston, MA, USA.
Shekhar, A., Buchmann, N., Humphrey, V., Gharun, M., 2024. More than three-fold increase in compound soil and air dryness across europe by the end of 21st century. Weather. Clim. Extrem. 44, 100666.
Shu, S., Jain, A.K., Koven, C.D., Mishra, U., 2020. Estimation of permafrost SOC stock and turnover time using a land surface model with vertical heterogeneity of permafrost soils. Glob. Biogeochem. Cycles 34 (11), e2020GB006585.
Sitch, S., Brovkin, V., von Bloh, W., van Vuuren, D., Eickhout, B., Ganopolski, A., 2005. Impacts of future land cover changes on atmospheric CO2 and climate. Glob. Biogeochem. Cycles 19 (2).
Sitch, S., O'sullivan, M., Robertson, E., Friedlingstein, P., Albergel, C., Anthoni, P., Arneth, A., Arora, V.K., Bastos, A., Bastrikov, V., et al., 2024. Trends and drivers of terrestrial sources and sinks of carbon dioxide: An overview of the TRENDY project. Glob. Biogeochem. Cycles 38 (7), e2024GB008102.
Smallman, T.L., Milodowski, D.T., Neto, E.S., Koren, G., Ometto, J., Williams, M., 2021. Parameter uncertainty dominates c cycle forecast errors over most of Brazil for the 21st century. Earth Syst. Dyn. Discuss. 2021, 1–52.
Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., Zaehle, S., 2014. Implications of incorporating n cycling and n limitations on primary production in an individual-based dynamic vegetation model. Biogeosciences 11 (7), 2027–2054.
Spawn, S.A., Sullivan, C.C., Lark, T.J., Gibbs, H.K., 2020. Harmonized global maps of above and belowground biomass carbon density in the year 2010. Sci. Data 7 (1), 112.
Stocker, B.D., Wang, H., Smith, N.G., Harrison, S.P., Keenan, T.F., Sandoval, D., Davis, T., Prentice, I.C., 2020. P-model v1. 0: An optimality-based light use efficiency model for simulating ecosystem gross primary production. Geosci. Model. Dev. 13 (3), 1545–1581.
Sulla-Menashe, D., Friedl, M.A., 2018. User guide to collection 6 MODIS land cover (MCD12Q1 and MCD12C1) product. USGS: Rest. VA, USA 1, 18.
Todd-Brown, K.E., Randerson, J.T., Post, W.M., Hoffman, F.M., Tarnocai, C., Schuur, E.A., Allison, S.D., 2013. Causes of variation in soil carbon simulations from CMIP5 earth system models and comparison with observations. Biogeosciences 10 (3), 1717–1736.
Tramontana, G., Ichii, K., Camps-Valls, G., Tomelleri, E., Papale, D., 2015. Uncertainty analysis of gross primary production upscaling using random forests, remote sensing and eddy covariance data. Remote Sens. Environ. 168, 360–373.
Tramontana, G., Jung, M., Schwalm, C.R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M.A., Cescatti, A., Kiely, G., et al., 2016. Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms. Biogeosciences 13 (14), 4291–4313.
Ueyama, M., Ichii, K., Iwata, H., Euskirchen, E.S., Zona, D., Rocha, A.V., Harazono, Y., Iwama, C., Nakai, T., Oechel, W.C., 2013. Upscaling terrestrial carbon dioxide fluxes in alaska with satellite remote sensing and support vector regression. J. Geophys. Res.: Biogeosciences 118 (3), 1266–1281.
Vuichard, N., Messina, P., Luyssaert, S., Guenet, B., Zaehle, S., Ghattas, J., Bastrikov, V., Peylin, P., 2019. Accounting for carbon and nitrogen interactions in the global terrestrial ecosystem model ORCHIDEE (trunk version, rev 4999): Multi-scale evaluation of gross primary production. Geosci. Model. Dev. 12 (11), 4751–4779.
Walker, A.P., Quaife, T., Van Bodegom, P.M., De Kauwe, M.G., Keenan, T.F., Joiner, J., Lomas, M.R., MacBean, N., Xu, C., Yang, X., et al., 2017. The impact of alternative trait-scaling hypotheses for the maximum photosynthetic carboxylation rate (vcmax) on global gross primary production. New Phytol. 215 (4), 1370–1386.
Wang, H., Prentice, I., Davis, T., 2014. Biophsyical constraints on gross primary production by the terrestrial biosphere. Biogeosciences 11 (20), 5987–6001.
Williams, M., Schwarz, P.A., Law, B.E., Irvine, J., Kurpius, M.R., 2005. An improved analysis of forest carbon dynamics using data assimilation. Global Change Biol. 11 (1), 89–105.
Wiltshire, A.J., Burke, E.J., Chadburn, S.E., Jones, C.D., Cox, P.M., Davies-Barnard, T., Friedlingstein, P., Harper, A.B., Liddicoat, S., Sitch, S., et al., 2021. JULES-CN: a coupled terrestrial carbon–nitrogen scheme (JULES vn5. 1). Geosci. Model. Dev. 14 (4), 2161–2186.
Xia, X., Ren, P., Wang, X., Liu, D., Chen, X., Dan, L., He, B., He, H., Ju, W., Liang, M., et al., 2024. The carbon budget of China: 1980–2021. Sci. Bull. 69 (1), 114–124.
Xiao, J., Chen, J., Davis, K.J., Reichstein, M., 2012. Advances in upscaling of eddy covariance measurements of carbon and water fluxes. J. Geophys. Res.: Biogeosciences 117 (G1).
Xu, B., Pan, Y., Plante, A.F., Johnson, A., Cole, J., Birdsey, R., 2016. Decadal change of forest biomass carbon stocks and tree demography in the delaware river basin. Forest Ecol. Manag. 374, 1–10.
Yang, X., Thornton, P., Ricciuto, D., Wang, Y., Hoffman, F., 2023. Global evaluation of terrestrial biogeochemistry in the energy exascale earth system model (E3SM) and the role of the phosphorus cycle in the historical terrestrial carbon balance. Biogeosciences 20 (14), 2813–2836.
You, Y., Tian, H., Pan, S., Shi, H., Bian, Z., Gurgel, A., Huang, Y., Kicklighter, D., Liang, X.-Z., Lu, C., et al., 2022. Incorporating dynamic crop growth processes and management practices into a terrestrial biosphere model for simulating crop production in the United States: Toward a unified modeling framework. Agricult. Forest. Meterol. 325, 109144.
Yuan, W., Liu, S., Zhou, G., Zhou, G., Tieszen, L.L., Baldocchi, D., Bernhofer, C., Gholz, H., Goldstein, A.H., Goulden, M.L., et al., 2007. Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. Agricult. Forest. Meterol. 143 (3–4), 189–207.

Yue, X., Unger, N., 2015. The yale interactive terrestrial biosphere model version 1.0: description, evaluation and implementation into NASA giss ModelE2. Geosci. Model. Dev. 8 (8), 2399–2417.

Zaehle, S., Ciais, P., Friend, A.D., Prieur, V., 2011. Carbon benefits of anthropogenic reactive nitrogen offset by nitrous oxide emissions. Nat. Geosci. 4 (9), 601–605.

Zeng, J., Matsunaga, T., Tan, Z.H., Saigusa, N., Shirai, T., Tang, Y., Peng, S., Fukuda, Y., 2020. Global terrestrial carbon fluxes of 1999–2019 estimated by upscaling eddy covariance data with a random forest. Sci. Data 7 (1), 313.

Zhou, Z.-H., Feng, J., 2019. Deep forest. Natl. Sci. Rev. 6 (1), 74–86.

Zhu, S., Clement, R., McCalmont, J., Davies, C.A., Hill, T., 2022a. Stable gap-filling for longer eddy covariance data gaps: A globally validated machine-learning approach for carbon dioxide, water, and energy fluxes. Agricult. Forest. Meterol. 314, 108777.

Zhu, S., Dong, W., Xu, J., Mitchard, E., Xie, Q., Myrgiotis, V., Zheng, Z., Chen, M., Ryan, C.M., Dash, J., 2025a. Towards consistent carbon monitoring: The role of satellite resolution and process-informed machine learning in terrestrial carbon cycle modelling. Available At SSRN 5352244.

Zhu, S., McCalmont, J., Cardenas, L.M., Cunliffe, A.M., Olde, L., Signori-Müller, C., Litvak, M.E., Hill, T., 2023a. Gap-filling carbon dioxide, water, energy, and methane fluxes in challenging ecosystems: Comparing between methods, drivers, and gap-lengths. Agricult. Forest. Meterol. 332, 109365.

Zhu, S., Olde, L., Lewis, K., Quaife, T., Cardenas, L., Loick, N., Xu, J., Hill, T., 2023b. Eddy covariance fluxes over managed ecosystems extrapolated to field scales at fine spatial resolutions. Agricult. Forest. Meterol. 342, 109675.

Zhu, S., Quaife, T., Hill, T., 2024a. Uniform upscaling techniques for eddy covariance fluxes (UFLUX). Int. J. Remote Sens. 45 (5), 1450–1476.

Zhu, S., Xu, J., Zeng, J., Bao, S., Chen, Y., Shi, S., Zheng, Z., Dong, W., Wang, Y., Shi, J., 2025b. The UFLUX ensemble of multiple-scale carbon, water, and energy fluxes. Sci. Data.

Zhu, S., Xu, J., Zeng, J., Feng, X., Wang, Y., Bao, S., Shi, J., 2023c. Explainable machine learning confirms the global terrestrial CO_2 fertilization effect from space. IEEE Geosci. Remote. Sens. Lett. 20, 1–5.

Zhu, S., Xu, J., Zeng, J., He, P., Wang, Y., Bao, S., Ma, J., Shi, J., 2024b. UFLUX-GPP: A cost-effective framework for quantifying daily terrestrial ecosystem carbon uptake using satellite data. IEEE Trans. Geosci. Remote Sens..

Zhu, S., Xu, J., Zhu, H., Zeng, J., Wang, Y., Zeng, Q., Zhang, D., Liu, X., Yang, S., 2022b. Investigating impacts of ambient air pollution on the terrestrial gross primary productivity (GPP) from remote sensing. IEEE Geosci. Remote. Sens. Lett. 19, 1–5.