https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

Constraining biospheric carbon dioxide fluxes by combined top-down and bottom-up approaches

Samuel Upton[1,2], Markus Reichstein[1], Fabian Gans[1], Wouter Peters[2,3], Basil Kraft[1], and Ana Bastos[1]

[1]Department of Biogeochemical Integration, Max Plank Institute of Biogeochemistry
[2]Environmental Sciences Group, Wageningen University, Wageningen, The Netherlands
[3]University of Groningen, Centre for Isotope Research, Groningen, The Netherlands

Correspondence: Samuel Upton (supton@bgc-jena.mpg.de)

Abstract. While the growth rate of atmospheric CO_2 mole fractions can be measured with high accuracy, there are still large uncertainties in the attribution of this growth to diverse anthropogenic and natural sources and sinks. One major source of uncertainty is the net flux of carbon dioxide from the biosphere to the atmosphere, the Net Ecosystem Exchange (NEE). There are two major approaches to quantifying NEE: top-down approaches that typically use atmospheric inversions, and bottom-up estimates which rely on process-based or data-driven terrestrial biosphere models or inventories. Both approaches have known limitations. Atmospheric inversions produce estimates of NEE that are consistent with the atmospheric CO_2 growth rate at regional and global scales, but are highly uncertain at smaller scales. Bottom-up data-driven flux models match local observations of NEE, but have difficulty in accurately upscaling to a global estimate. We combine the two approaches, constraining a bottom-up data-driven flux model trained on meteorological, remotely-sensed, and eddy-covariance data with regional estimates of NEE derived from an ensemble of atmospheric inversions.

We link the two approaches using a region-specific sparse linear model for 18 regions consistent with the Regional Carbon Cycle Assessment and Processes-2 (RECCAP2) (Bastos et al., 2020; Tian et al., 2018), which allows us to quickly generate regional estimates of NEE based on the data-driven flux model by simulating only a small number of optimally representative pixels. These regional totals then become part of a machine-learning objective function that compares them with top-down regional estimates from an ensemble of atmospheric inverse models. By adding this additional constraint from the top-down objective term, we produce a new "dual-constraint" data-driven flux model that is informative across spatial scales, producing consistent estimates both of the local per-pixel flux and at regional and global scales.

The inferred global terrestrial carbon flux from land, excluding fires and riverine evasion across 2001-2017 is -3.14±1.75 PgC year^{-1} (±1 σ). This is a strong improvement over the -20.28±1.75 PgC year^{-1} from the exact same data-driven flux model trained without the additional regional top-down constraint (i.e., single constraint) when compared to current best estimates of the global carbon flux from land. The shift in the carbon flux from land estimated by the model with the additional atmospheric constraint occurs largely in tropical regions where the data-driven flux model is poorly constrained, or affected by biased observations of NEE derived from difficult micrometeorological conditions. In extratropical regions, the estimated NEE from dual and single constraint data-driven flux models are very similar, reflecting the denser observational networks of ecosystem fluxes and atmospheric CO_2. Our approach, training a data-driven flux model with multiple constraints at site-

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

level and continentally integrated scales, and different temporal resolutions, opens new avenues for data-driven flux models constrained by other observations of atmospheric carbon dioxide, making use of the wealth of available Earth observation data.

1 Introduction

Since 1957, the growth of carbon dioxide in the atmosphere can be directly measured with high accuracy (Keeling et al., 1989). However, linking atmospheric CO_2 changes to the diverse anthropogenic and natural sources and sinks is still prone to uncertainties. One of the main sources of uncertainty is the land biosphere, which has large uncertainties both in the human-related (land-use change) and the natural flux components (Friedlingstein et al., 2022). Improved understanding of driving processes and drivers of variations in the global carbon budget requires, among others, improved observational constraints on the net flux of carbon dioxide from the land surface to the atmosphere or Net Ecosystem Exchange (NEE), at local, regional, and global scales (Bastos et al., 2022; Ciais et al., 2022; Gaubert et al., 2019; Saeki and Patra, 2017; Thompson et al., 2016).

Two different approaches to constrain sources and sinks of carbon dioxide can be distinguished: top-down estimates, typically from atmospheric inversions of observed mole fractions, and bottom-up estimates, usually relying on flux observations, remote-sensing products, or process-based models. Both approaches have known strengths and limitations (Jung et al., 2020; Kondo et al., 2020).

Top-down estimates from atmospheric inversions infer the surface fluxes over land and ocean based on observations of atmospheric CO_2 mixing ratio based on a Bayesian inversion framework using prior estimates of the flux of carbon dioxide and an atmospheric transport model. Inversions provide terrestrial fluxes that are consistent with the global atmospheric CO_2 growth rate and show increasing agreement between different inversion systems at the scale of continental-sized regions (Gaubert et al., 2019). However, inversion-based estimates of NEE are conditioned by the uncertainties of their flux priors, their atmospheric transport model, the distribution of the CO_2 observational data (Peylin et al., 2013), and their estimates are increasingly uncertain for smaller regions (Kaminski and Heimann, 2001). Inverse model evaluations at the scale of individual ecosystem sites is rarely done, as the mismatch in spatial scales represented is simply too large.

The FLUXCOM project (Jung et al., 2020) produced a comprehensive comparison of data-driven bottom-up approaches for upscaling terrestrial biosphere carbon dioxide and water fluxes based on eddy-covariance measurements. The project produced two products, one using remotely-sensed variables (RS), and another using both meteorological and remotely-sensed variables (RS+METEO), as predictors in a data-driven flux model. FLUXCOM model ensemble capture the shape of the seasonality, and the sign of the annual anomalies of NEE compared with inversion-based estimates of NEE in the extratropics, but they have systematic biases in the tropics, leading to an overestimation of NEE (Jung et al., 2020). This is partly due to the distribution of eddy-covariance towers, dense in the northern extratropics and sparse in the tropics, which therefore is not representative of the global biome distribution. This leads to an under-constrained model at locations that are distant from eddy-covariance observations (Tramontana et al., 2016; Chu et al., 2017). The systematic tropical bias may also be caused by errors and data quality issues in the tropical eddy-covariance record, in particular at night-time (Fu et al., 2018; Jung et al., 2020).

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

Previous studies suggested that observations of atmospheric CO_2 could provide an additional constraint to bottom-up data-driven flux models (Jung et al., 2020; Anav et al., 2015; Beer et al., 2010). It is important to note that the carbon fluxes estimated from top-down and bottom up approaches are not exactly the same, as emphasized in detail in Ciais et al. (2022). The primary observational constraint of top-down methods is atmospheric CO_2 observations, hence atmospheric inversions are sensitive to nearly all carbon exchange with land ecosystems as well as with the inland water systems (lakes, rivers) in these ecosystems, illustrated in the blue and green boxes in Fig 2 of Ciais et al. (2022). In contrast, eddy-covariance flux measurements reflect carbon exchange across a smaller footprint, and their location is often selected to explicitly exclude fire fluxes and inland water systems. A meaningful comparison between fluxes derived from each thus includes accounting for both fires, and inland water fluxes, as also discussed in Friedlingstein et al. (2022). In this study, we similarly account for these inherent differences as described in Section 3.2. From hereon, when we refer to "NEE" in the text we refer to land-ecosystem derived carbon exchange, excluding inland water system contributions, and excluding the release of carbon by fires.

Atmospheric inversion estimates of NEE are most reliable at regional scales, and as such are most comparable to the regional integral of a data-driven flux model's output. It is infeasible to run a data-driven flux model regionally or globally for every step during the training process. Reading the necessary data to create an inference at every pixel is too time-intensive for efficient experimentation. Given this barrier, adding an atmospheric constraint to a data-driven approach therefore requires a computationally efficient "bridge model" which connects the data-driven flux model with the atmospheric data. The central hypothesis of this study is that individually trained regional sparse linear models can serve this function, and enable a complementary constraint based on (top-down) atmospheric inversions to the (bottom-up) data-driven flux model, improving the estimates of regional and global land surface CO_2 fluxes. [It does not become clear that the bridge model conceptually is trivial (summing up), but the challenge to overcome is computational] This methodology of multiple constraints across diverse data using a computational bridge may point the way towards accessing the large volume of direct observations of atmospheric CO_2. The improvement we see from the inclusion of the second constraint, described in this study, indicates that our computational bridge efficiently transmits information from the top-down data to the bottom-up data-driven flux model. A computational bridge which calculates or emulates atmospheric transport could provide the same benefits, linking a data-driven flux model to a point observation of atmospheric CO_2 mole fraction by estimating the near-field sources and strength of the flux that impacted the observation.

2 Data

2.1 Eddy-Covariance / Driver Data

The data-driven flux models in this study are trained using meteorological observations and NEE data collected by globally distributed eddy-covariance towers in the La Thuile synthesis dataset of the FLUXNET network (https://fluxnet.fluxdata.org/data/la-thuile-dataset/). Following Tramontana et al. (2016), driver variables are created using a set of remotely-sensed and meteorological data from the Moderate Resolution Imaging Spectroradiometer (MODIS) data (http://daac.ornl.gov/MODIS/), and the European Center for Medium-Range Weather Forecasting ERA5 atmospheric reanalysis dataset (https://www.ecmwf.

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

int/en/forecasts/dataset/ecmwf-reanalysis-v5). At tower locations the meteorological data is derived from the FLUXNET and ERA5 data, while the global dataset uses only ERA5 data. See Tramontana et al. (2016) for a full discussion of the handling of meteorological data. The FLUXCOM-RS+METEO V1 NEE ensemble data using ERA5 meteorological forcing from Jung et al. (2020) were used to generate the regional sparse linear models described below.

2.2 Atmospheric Inversion Data

For the atmospheric constraint we use the members from the ensemble of atmospheric inversions from the Global Climate Budget (GCB22) (Friedlingstein et al. (2022), https://doi.org/10.18160/7AH8-K1X4) that are based on surface observations and are available for the period of our analysis, 2001-2017. We use the Copernicus Atmospheric Monitoring Service (CAMS v21r1) (1979-2021) (Chevallier et al., 2005), CarbonTracker Europe (CTE v21r1) (2001-2021) (van der Laan-Luijkx et al., 2017), Jena CarboScope (sEXTocNEET v2022) (1957-2021) (Rödenbeck, 2005; Rödenbeck et al., 2018), UoE in situ (v6.1b) (2001-2021) (Feng et al., 2016; Palmer et al., 2019), and NISMON-CO_2 (v2022.1) (2001-2021) (Niwa et al. (2022), available at https://www.nies.go.jp/doi/10.17595/20201127.001-e.html). All inversion results were provided on a common 1°grid.

Table 1. Atmospheric inversions and references from the Global Climate Project 2022 (Friedlingstein et al. (2022), https://doi.org/10.18160/7AH8-K1X4). All inversions are adjusted for fossil fuels, cement production and carbonation, and lateral river flux. To make the inversions more comparable with data-driven flux model estimating NEE, all inversions have the CAMS Global Fire Assimilation System (GFAS) estimate of fires removed. Only atmospheric inversions based on surface observations are used for reasons of data availability over the period 2001-2017.

Inversion	Version	Date range	Reference
CAMS	v21r1	1979-2021	Chevallier et al. (2005)
CarbonTracker Europe (CTE)	v21r1	2001-2021	van der Laan-Luijkx et al. (2017)
Jena Carboscope (sEXTocNEET)	v2022	1957-2021	Rödenbeck et al. (2003), Rödenbeck et al. (2018)
UoE in situ	v6.1b	2001-2021	Feng et al. (2016), Palmer et al. (2019)
NISMON-CO2	v2022.1	1990-2021	Niwa et al. (2022)

3 Methods

3.1 Driver Data

We use the driver variables from the FLUXCOM-RS+METEO with ERA5 forcing ensemble as in Jung et al. (2020): Enhanced Vegetation Index (EVI), fraction of Absorbed Photosynthetically Active Radiation (fAPAR), Daytime Land Surface Temperature (LST_{day}), Nighttime Land Surface Temperature (LST_{night}), the Medium Infrared Reflectance band (MIR), Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), extracted for each site and globally from

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

MODIS, and the ERA5 variables incoming global radiation (R_g), top of atmosphere potential radiation (R_{pot}), Water Availability Index (WAI), and Air Temperature (T_{air}). The full set of drivers were constructed following Jung et al. (2020) and Tramontana et al. (2016), see Appendix table A2 for the driver formulation. These variables are were computed globally and stored by Plant Functional Type (PFT) derived from the MODIS Land Cover Type Yearly L3 Global 500m dataset Collection 5 (Friedl et al., 2010), and are reconstructed here by the on the percentage of the component PFTs in each pixel following the approach in Jung et al. (2019). Two drivers, WAI and T_{air} are used at daily time step, while the other eight are constructed from the mean seasonal cycle (MSC) signal across the component variables. All are used at a 0.5°spatial resolution.

3.2 Atmospheric Inversion Adjustment

As discussed above, and in more detail in Ciais et al. (2022), these atmospheric data are not directly comparable with the NEE output of an data-driven flux model based on eddy-covariance data. FLUXNET observations of NEE include the difference between the gross primary productivity and the ecosystem respiration across a small footprint and implicitly include disturbance and management fluxes. Atmospheric inversions use observed CO_2 mole fraction with estimates of the non-biogenic fluxes to infer the exchange of carbon between the land and atmosphere create a product corresponding better to net biome productivity, or the total regional gain or loss of carbon from all processes, i.e. including the signal from fires and other disturbances, land use change and management, and river evasion (Ciais et al., 2022).

To account for this mismatch, this study uses the atmospheric inversion data from the GCB22, which are adjusted to rectify small differences in prescribed fossil fuel emissions and cement production and carbonation fluxes. The inversions are further adjusted for lateral riverine CO_2 transport to make their output more comparable with bottom-up models. See Friedlingstein et al. (2022) for a full discussion of inversion systems. To facilitate the comparison to, and combination with, EC-driven carbon exchange, we subtracted the fire flux from all inversion land fluxes at each grid cell. For this adjustment, we used the gridded fire fluxes from CAMS Global Fire Assimilation System (GFAS) (Di Giuseppe et al., 2018) without additional adjustment, which is consistent with the CarbonTracker Europe treatment of fire. The resulting fluxes thus represent land-ecosystem carbon exchange excluding inland waters and fires.

3.3 Modelling approach

We use two identical data-driven flux models, EC (eddy-covariance only), and EC-ATM (eddy-covariance plus regional NEE from atmospheric inversion) which both accept a set of drivers and predict daily NEE (Fig. 1). They consist of feed-forward neural networks, or a set of fully connected network layers, which we train using the standard gradient-based backpropagation algorithm (Kelley, 1960). To rule out effects of parameter initialization, we use the same initial state for the EC and EC-ATM networks. The fully-connected layers consist of nodes or 'neurons', which are exposed to the output of all neurons in the previous layer. Non-linearity is introduced by passing each node output through a non-linear activation function. Our network is a set of three fully-connected layers with the ReLU activation function (Agarap, 2019). We provide a detailed illustration of the model architecture in Appendix Figure B1.

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

The EC model is trained with a single term in the objective function, minimizing the difference between model inferences of NEE and eddy-covariance observations of NEE. The EC-ATM model has two terms in the objective function: the first is identical to the EC model. The second minimizes the difference between the estimates of the integral of per-region NEE and the atmospheric inversions regional NEE values. As a result, the comparison of EC and EC-ATM results reflects the specific effect of the atmospheric inverse constraint compared with the underlying EC model.

The EC and EC-ATM output provide a side-by-side view of the paired learning process and a check on the impact of the additional term in the cost-function. The hyperparameter set is available in the Appendix (table B1).

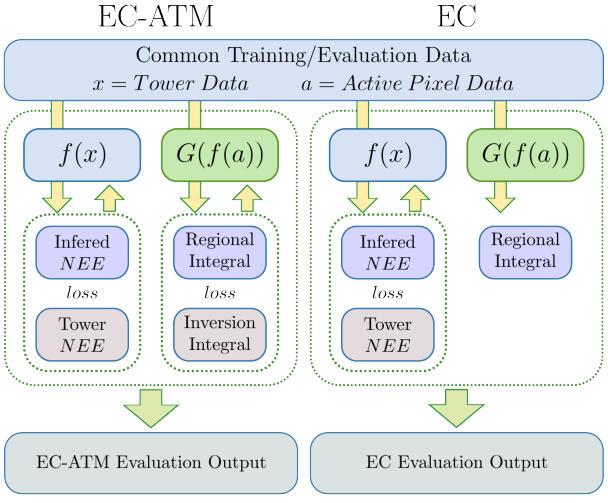

Figure 1. Experimental setup for parallel model training: Two data-driven flux models, EC and EC-ATM, are created for each experimental run with identical initial weights for the model's neural network $f(x)$. The two models are optimized for the same EC data x. The EC-ATM model additionally sends output of $f(x)$ to the set of regional sparse linear models $G(f(x))$, inferring the regional monthly integral of NEE. The inferred integrals are compared with the monthly regional integrals from the inversions, and this term is added to the objective function. Both the EC and EC-ATM return the output of $G(f(x))$ as an evaluation output during training.

Each data-driven flux model is trained using a 10-fold cross-validation scheme, splitting the eddy-covariance observations by site, holding out one fold per training cycle for validation. The folds are the same as those in Tramontana et al. (2016) for comparability. For the atmospheric inversion training data used by the EC-ATM model, two random years from the full 18 year set are held out for validation. There are insufficient years available for a fully independent set of test years outside the training and validation folds. We use the land sink, S_{land} for the CGB22 from the as an independent data to test the model at global scale. Comparisons with the atmospheric inversions in results below are the same inversion data that is used in training. The reported results are for the ensemble mean across the 10 folds. For both the EC and EC-ATM models, the ensemble members are the data-driven flux model with the weights from epoch with the best validation result for that fold. These members are used for a full forward run across the period 2001-2017.

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

3.4 EC-ATM constraint

The link between the EC-ATM model and the regional atmospheric inversion data are sparse linear models trained for each RECCAP2 region, created using Least Absolute Shrinkage and Selection Operator (Lasso) regression. The targets for the regression are created using the global NEE data cube created by the FLUXCOM intercomparison (Jung et al., 2020). These targets are the sum of NEE for all non-zero pixels $p \in P$ in the region for time steps t in all times T (Eq. 1). A subset of candidate pixels for a region $s \in S_r$ is selected from P at time t using stratified random sampling from a set of clusters generated using Dynamic Time Warping (DTW) (Tormene et al., 2008) as a metric of similarity, followed by spectral clustering. DTW is a method for comparing two time series which finds an minimum distance between them by allowing non-aligned time steps to be paired in the distance calculation as a similarity metric. The candidates from the stratified random sampling S_r are then used for Lasso regression which extends Ordinary Least Squares (OLS) regression (Eq. 2) by adding a term to the objective function. This term, a weighted L1 norm, or the mean of the absolute values of the weights of the linear model β times the weighting term α, has the effect of driving some weights to zero as α is increased. The result of this regression analysis is a sparse set of non-zero parameters which linearly combine to produce a high-fidelity estimate of the regional sum of NEE. The locations with non-zero coefficients are referred to as "contributing pixels". For each region a threshold of 0.95 R^2 is set, and the training loop iterates, reducing the α term, and increasing the number of contributing pixels in the region until the correlation threshold is reached using the training set. This result is tested using cross-validation.

$$y_t = \sum_{p=0}^{P} NEE_{tp} \tag{1}$$

$$\min \left\{ \sum_{t=0}^{T} (y_t - \sum_{s=0}^{S_r} NEE_{ts}\beta_s)^2 + \alpha \sum_{s=1}^{S_r} |\beta_s| \right\} \tag{2}$$

The robustness of the sparse linear model was tested by 1000 runs with random stratified sampling within the discovered classes. The output shows stable spatial locations of contributing pixels within each region. Heat maps of contributing pixels (Fig. B2) show that the Lasso approach repeatedly finds similar pixel locations, and will select spatial neighbors if the most advantageous pixels are not included in the randomization.

The stability of these spatial regimes indicate that there is a statistical link at the spatial resolution of the analysis between the contributing pixels and the regional integral of the EC model. It is plausible that this relationship might be scale-independent due to the spatial auto-correlation in the NEE fields. The contributing pixels cover the range of the PFTs in the region (fig B3. The EC and EC-ATM models receive no PFT information during training, and the included PFT breakdown is based on the majority class at 0.5°resolution, not the available PFT information at the eddy-covariance sites, which is specific to the footprint of the EC tower.

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

3.5 Model Training - EC

At each training step, the EC model f with parameters ω is run for a set of driver variables from eddy-covariance measurements x_{batch} (Eq. 3). The resulting inferences of NEE, $\hat{y}$, are used to calculate the first term in the objective function, the loss at the EC tower locations, ℓ_{EC}, which is the Mean Squared Error (MSE) between the $\hat{y}$, and the observed NEE at the tower locations in the training batch y_{batch} (Eq. 4).

$$\hat{y} = f(\omega, x_{\text{batch}}) \tag{3}$$

$$\ell_{EC} = MSE(y_{\text{batch}}, \hat{y}) \tag{4}$$

3.6 Model Training - EC-ATM

The EC-ATM model has two constraints. The first, ℓ_{EC}, is based on EC tower observations of NEE and driver variables measured or extracted at the EC tower sites. This term is identical to the objective function of the EC model. The second constraint ℓ_{ATM}, is calculated for each region r in all regions R, for all months m in a year M using the linear equation with region-specific parameters Θ_r and intercept b_r discovered with Lasso regression using the inferred flux estimate from $f(\omega)$ at the contributing pixel locations with data $x_{r,d}$ as inputs. The sparse linear model is run for for the days of the month $d \in D$, and averaged for a monthly estimate (Eq. 5). This regional monthly estimate, $\hat{m_{r,m}}$, is compared using the ℓ_3-norm with inversion-based estimates from inversions $a \in A$ of regional NEE, $a_{r,m}$, producing the monthly regional loss $\ell_{r,m}$ (Eq. 6). The ℓ_3-norm was chosen because it improved the Interannual Variability (IAV) of the final NEE estimates.

$$\hat{m_r} = \frac{1}{D}\sum_{d=0}^{D} f(\omega, x_{r,d}) \cdot \Theta_r + b_r \tag{5}$$

$$\ell_{r,m} = \sqrt[3]{\frac{1}{A}\sum_{a=0}^{A}(A_{a,r,m} - \hat{m}_{r,m})^3} \tag{6}$$

This $\ell_{r,m}$ term is normalized by the range of the regional NEE from atmospheric inversions for that month $range_{r,m}$ (Eq. 7). This reduces the weight of the loss where the atmospheric inversion ensemble has a higher uncertainty. The weighted losses are averaged for all months in M creating a regional loss term $\ell_{ATM,r}$. The area-weighted average of the regional losses creates an atmospheric loss term ℓ_{ATM} (Eq. 8).

$$\ell_{ATM,r} = \frac{1}{M}\sum_{m=0}^{M}\left(\frac{\ell_{r,m}}{1 + range_{r,m}}\right) \tag{7}$$

$$\ell_{ATM} = \sum_{r=0}^{R} \ell_{ATM,r} \times \frac{land\,area_r}{land\,area_{globe}} \tag{8}$$

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

The two terms of the objective function, ℓ_{EC} and ℓ_{ATM} are combined using an empirically learned weighting scheme (Kendall et al., 2018), which learns the appropriate relative weights for the set of losses. This method adds a parameter to the learned weights of the data-driven model that estimates the task-dependent, homoscedastic uncertainty for each of the different terms of the objective function, which is dependent on the inherent noise in the data, rather than the scale or quality of the inputs. This term is an estimate of the variance of each component loss of the objective function over all training steps. For the EC-ATM model these parameters, σ^2_{EC} and σ^2_{ATM}, are are added to the model training weights, and updated by the regular backpropagation step of the neural network training. The σ^2_{LOSS} parameter is used to create a weighting term w_{LOSS} (Eq. 9) and a regularization term s_{LOSS} (Eq. 9) for each component term of the objective function, $LOSS \in [EC, ATM]$. These are then combined to provide a learned estimate of the total loss, balanced by the learned uncertainty of the terms (Eq. 11).

The total loss of the EC-ATM model is then:

$$w_{[LOSS]} = \frac{1}{2\sigma^2_{[LOSS]}} \tag{9}$$

$$s_{[LOSS]} = \log\sqrt{\sigma^2_{[LOSS]}} \tag{10}$$

$$\mathcal{L}_{total} = (w_{EC} \times \ell_{EC}) + (w_{ATM} \times \ell_{ATM}) + s_{EC} + s_{ATM} \tag{11}$$

3.7 Post-Hoc Analysis

We additionally conduct a post-hoc test of the relationship between the regional sparse linear models and the calculated integrals by applying the regional sparse linear models over the contributing pixel locations within the output of our training runs. This helps diagnose differences between the loss values calculated against the regional linear model and the calculated integrals that occur during training.

4 Results

The addition of an atmospheric constraint based on modeled NEE from atmospheric inversions to a data-driven flux model (i.e. the EC-ATM model) leads to a global estimate of NEE that is closer to the current best estimates from GCB22 (Friedlingstein et al., 2022) than the model only based on eddy-covariance flux data (EC model). Figure 2a shows the global annual NEE for EC-ATM, to be much closer to the the atmospheric inversion ensemble mean and member totals, and the GCB22 estimate of the land flux than the EC and FLUXCOM-RS+METEO V1 (Jung et al., 2020) estimates. Figure 2b shows that the IAV of NEE estimated by EC-ATM ensemble largely captures the sign of the atmospheric inversion and GCB22 anomalies, but the IAV is still underestimated by both the EC-ATM and EC models, which is a persistent issue with the FLUXCOM approach (see Jung et al. (2020) for full discussion). The EC-ATM ensemble mean preserves the correlation with the scaled anomalies, producing very similar results to the FLUXCOM RS+METEO results 3.

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

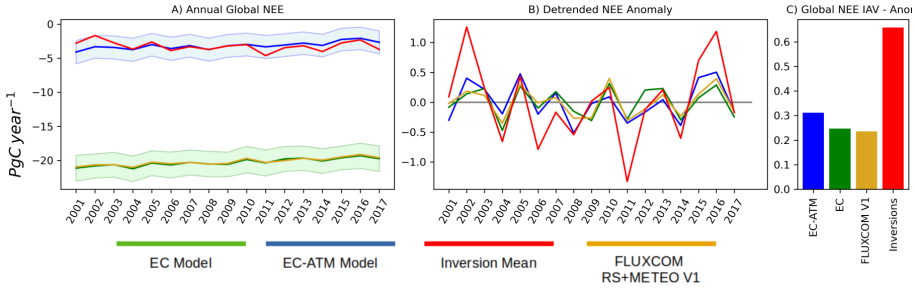

Figure 2. Panel (A) is the global annual NEE in PgC year^{-1}. (B) shows the detrended anomalies in PgC year^{-1}. (C) shows the IAV, or standard deviation, of detrended annual anomalies

Table 2. Results of monthly NEE aggregated by regions: Pearson's R and RMSE of the monthly time-series of regional/global integrals over the period 2001-2017 and the corresponding monthly NEE from atmospheric inversions, and the Pearson's R of the regional/global Mean Seasonal Cycle (MSC) of NEE, the RMSE of the MSC relative to the inversion mean and the model MSC. FLUXCOM refers to the RS+METEO V1 product (Jung et al., 2020). Country/region abbreviations are expanded in Appendix table A1

	Pearson's R – monthly integrals			**RMSE – monthly integrals**			**Pearson's R – MSC**			**RMSE – MSC**		
	EC-ATM	**EC**	**FLUXCOM**	**EC-ATM**	**EC**	**FLUXCOM**	**EC-ATM**	**EC**	**FLUXCOM**	**EC-ATM**	**EC**	**FLUXCOM**
GLOBAL	**0.991**	0.969	0.976	**0.176**	1.502	1.542	**0.996**	0.976	0.982	**0.125**	1.496	1.537
USA	**0.989**	0.971	0.974	**0.036**	0.090	0.105	**0.998**	0.979	0.982	**0.022**	0.086	0.101
CAN	**0.991**	0.949	0.974	**0.048**	0.083	0.049	**0.997**	0.958	0.981	**0.044**	0.080	0.045
CAM	**0.658**	0.557	0.607	**0.022**	0.061	0.071	**0.861**	0.685	0.751	**0.020**	0.060	0.070
NSA	**0.133**	-0.007	-0.023	**0.018**	0.148	0.127	**0.197**	-0.072	-0.080	**0.012**	0.147	0.126
BRA	**0.787**	-0.028	0.040	**0.055**	0.354	0.350	**0.893**	-0.051	0.024	**0.043**	0.352	0.348
SSA	**0.751**	0.388	0.387	**0.032**	0.176	0.173	**0.869**	0.431	0.431	**0.025**	0.175	0.172
EU	0.988	0.986	**0.989**	**0.034**	0.059	0.061	**0.997**	0.994	0.997	**0.026**	0.055	0.058
NAF	**0.929**	0.809	0.902	**0.031**	0.058	0.033	**0.959**	0.834	0.926	**0.027**	0.056	0.030
EQAF	**0.675**	0.186	0.253	**0.057**	0.251	0.247	**0.783**	0.173	0.248	**0.046**	0.249	0.245
SAF	**0.920**	0.600	0.665	**0.037**	0.102	0.117	**0.982**	0.628	0.698	**0.025**	0.099	0.114
RUS	**0.993**	0.931	0.970	**0.074**	0.225	0.138	**0.999**	0.939	0.977	**0.059**	0.220	0.130
CAS	**0.936**	0.835	0.655	**0.039**	0.069	0.054	**0.957**	0.856	0.698	**0.036**	0.067	0.052
MIDE	0.913	0.882	**0.919**	0.025	0.046	**0.020**	0.926	0.893	**0.931**	0.025	0.046	**0.019**
CHN	0.969	0.973	**0.973**	**0.035**	0.104	0.073	0.987	**0.990**	0.990	**0.025**	0.101	0.070
KAJ	**0.962**	0.937	0.939	**0.004**	0.009	0.012	**0.984**	0.960	0.959	**0.003**	0.008	0.011
SAS	**0.803**	0.184	0.607	**0.028**	0.111	0.046	**0.896**	0.208	0.680	**0.020**	0.109	0.042
SEAS	**0.273**	0.204	0.204	**0.050**	0.173	0.174	**0.436**	0.312	0.300	**0.042**	0.170	0.172
OCE	0.199	0.223	**0.240**	0.044	0.047	**0.036**	0.289	0.263	**0.310**	0.037	0.041	**0.030**

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

Table 3. Results of annual NEE aggregated by regions: Pearson's R of annual integral of inferred NEE with the annual integral of NEE from the inversion ensemble mean, and the RMSE of the annual regional/global NEE with the annual integral of NEE from the inversion ensemble mean in PgC year^{-1}

	Pearson's R – Annual			RMSE		
	EC-ATM	**EC**	**FLUXCOM**	**EC-ATM**	**EC**	**FLUXCOM**
GLOBAL	**0.280**	0.054	0.080	**0.7502**	16.8872	17.3236
USA	0.502	0.519	**0.543**	**0.2303**	0.8030	1.0836
CAN	**0.325**	-0.463	-0.236	0.3970	0.3747	**0.0627**
CAM	**-0.067**	-0.085	-0.149	**0.2236**	0.6598	0.7898
NSA	-0.360	**-0.140**	-0.253	**0.0969**	1.7467	1.4909
BRA	**0.358**	0.221	0.186	**0.1319**	3.8447	3.8912
SSA	-0.037	-0.094	**0.007**	**0.2021**	1.9080	1.8467
EU	-0.068	0.187	**0.191**	**0.2074**	0.6336	0.6805
NAF	0.557	0.616	**0.694**	**0.1900**	0.4362	0.2438
EQAF	**0.749**	0.725	0.740	**0.3876**	2.8844	2.8576
SAF	**0.780**	0.734	0.621	**0.1352**	0.5242	1.0036
RUS	**0.534**	-0.353	-0.409	0.6273	1.1558	**0.1809**
CAS	-0.379	-0.258	**-0.167**	0.1326	0.7133	**0.0759**
MIDE	0.254	0.405	**0.442**	0.2277	0.5117	**0.1585**
CHN	0.084	**0.297**	0.292	**0.1103**	1.0917	0.8059
KAJ	0.185	0.173	**0.216**	**0.0147**	0.0798	0.1167
SAS	**0.262**	0.066	0.039	**0.1384**	0.9604	0.3685
SEAS	-0.227	-0.263	**-0.221**	**0.4350**	2.0118	2.0409
OCE	0.499	0.599	**0.508**	0.4094	0.3703	**0.1037**

4.1 Mean Seasonal Cycle, Monthly results

The mean seasonal cycle (MSC) of the global NEE estimated by the EC-ATM model shows a clear adjustment towards the atmospheric inversion ensemble mean, as expected. The global correlation is very close to the FLUXCOM RS+METEO results, and the extratropical regions are similar, but the tropical regions show a meaningful improvement of the correlation (Table 2). The amplitude of the global MSC is also closer to the atmospheric inversion ensemble mean (3). Figure 3 also shows that the distribution of the annual flux is very close to the atmospheric inversion ensemble mean, although the EC-ATM model appears to underestimate the source during the Northern Hemisphere winter and conversely overestimate it in the Northern Hemisphere early spring. The EC-ATM model shows an improvement in the global RMSE of monthly NEE from 1.54 PgC/month for the EC ensemble mean to 0.13 PgC/month for the EC-ATM ensemble mean (Figure 2). This result confirms that additional atmospheric constraints are indeed reflected in the EC-ATM model at seasonal and continental scales, i.e., the scales where atmospheric data are most informative. We will focus next on the sub-continental scales, turning to RECCAP2 regions.

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

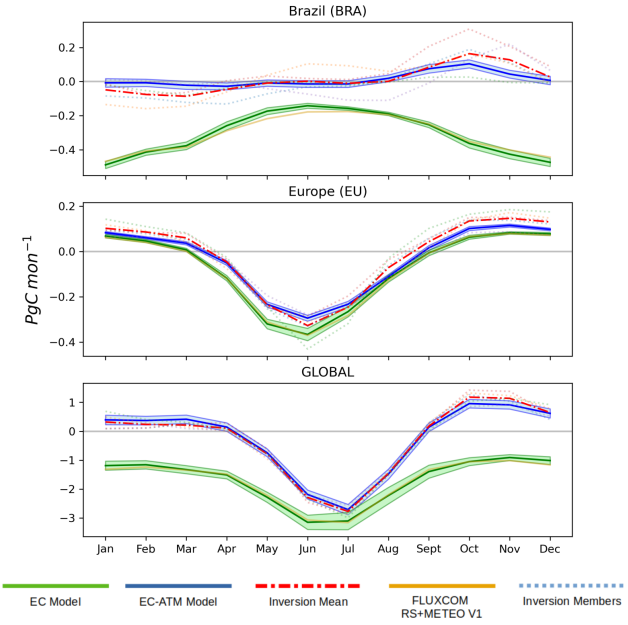

Figure 3. Mean seasonal cycle of ensemble mean of monthly NEE (PgC mon^{-1}) for a representative tropical region (Brazil BRA), extratropical region (Europe EU) and the globe, for years 2001-2017. The solid line is the ensemble mean, and the shaded region is the mean $\pm$ the ensemble standard deviation.

4.2 RECCAP2 Regions

When the estimates from EC-ATM and EC are compared for the RECCAP2 regions, the additional atmospheric constraint shifts the regional NEE integrals from EC-ATM towards the inversion regional estimates in tropical regions, but the EC and EC-ATM MSC results are very similar in extratropical regions, where the EC model is better constrained by the denser eddy-covariance measurement network. Figure 3 shows two selected regions, (1) Brazil which is representative of tropical regions with sparse EC observations and potential systematic bias in the EC data, and (2) Europe which is representative of extratropical regions with a dense EC network. The results in Europe are very similar for the EC, EC-ATM and FLUXCOM, with all three models producing a MSC very close to the ensemble mean of the atmospheric inversions. In Brazil, the EC-ATM ensemble mean shows closer MSC seasonality and magnitude to the mean of the atmospheric inversions. At a monthly time-step, the correlation of the EC-ATM with the atmospheric inversions is largely similar with the FLUXCOM RS+METEO V1, only out-performing it in some tropical regions. EC-ATM shows a persistent improvement in the RMSE, showing a reduction of the difference to the inversions across almost all regions (Table 2).

The shift towards the inversion regional values can be formalized with the normalized Nash Sutcliffe model efficiency (nNSE) metric (fig 4). Normalized Nash Sutcliffe model efficiency is a rescaling of standard Nash Sutcliffe metric which assesses the predictive skill of a model in regard to a reference set of values. The normalization transforms the range of the metric from $(-\infty, 1)$ to $(0, 1)$. An nNSE of 0 represents no skill, 0.5 is where the model predicts the reference better than the

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

mean, and 1.0 represents perfect skill. The score is calculated for the EC and EC-ATM regional monthly integrals compared to the inversion mean monthly integrals as the reference value. Figure 4 shows that in almost all regions the EC-ATM model is better able to predict of the atmospheric inversions than the EC model.

From these different metrics, we demonstrate a successful integration of the new constraint into the EC-ATM model which, at the regional level, can improve the inference of NEE. We turn now to the spatial distribution of inferred NEE.

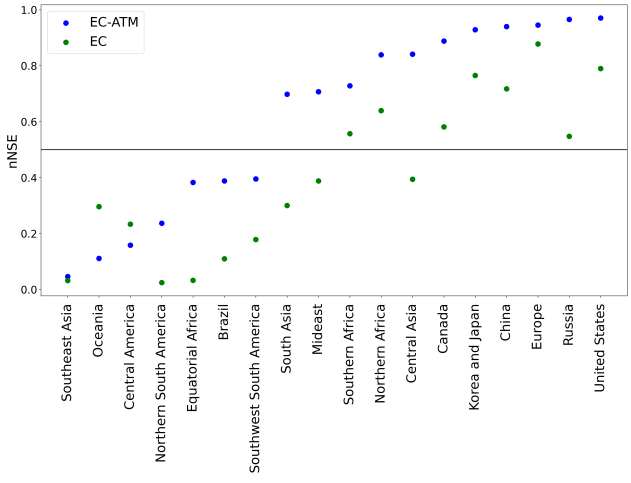

Figure 4. Normalized Nash-Sutcliffe model efficiency over all regions ordered by EC-ATM performance.

4.3 Spatial Analysis

The spatial patterns of mean annual NEE (Fig. 5) shows that the EC-ATM estimates deviates from the EC model, from a strong annual sink to a mix of sources and sinks in the tropical regions. For example, in the Amazon region, the EC model estimates a strong sink and the mean of the atmosphere inversions a weak and rather homogeneous source, while the EC-ATM model infers a strong pattern of sinks and sources across the Amazon and Cerrado regions. The strong source and sink bands in the Amazon basin may indicate some instability of the model, or amplification of the limited signal coming from the low density of eddy-covariance sites. The EC-ATM ensemble mean shows an annual source in western tropical Africa, where the EC and atmospheric inversion members both show a sink. It must be noted that also among the inverse models themselves, tropical fluxes are highly uncertain and the location of sources and sinks varies strongly (Friedlingstein et al., 2022; Gaubert et al., 2019; Palmer et al., 2019)). Reconciliation with the EC-based constraints is not likely to resolve these spatial differences but rather to include a bit of both, as exemplified by the patterns in the figure.

Moreover, we note that the EC-ATM and EC outcomes display much weaker year-to-year variability in NEE compared to the inversions, as reflected in the standard deviation of each grid box. This difference highlights the difficulty that both the EC and EC-ATM models have in capturing the magnitude of the interannual variability, as also shown in Jung et al. (2020). We will discuss this further in Section 5.2.

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

At a monthly level the distribution of the mean monthly flux from years 2001 to 2017 shows the reduced flux in the tropics throughout the year (Fig 6, blue regions in column 3). The EC-ATM model also shows a stronger extratropical sink for C3-crop areas in the USA, Europe and north-central Asia.

The spatial distribution of NEE estimated by the EC-ATM model is largely consistent with the EC distribution, while reducing the tropical sink. It shows some irregularities where there is insufficient information in either the eddy-covariance data or atmospheric inversions to robustly localize the NEE. Next, we discuss the strength of the relationship between the spatial distribution of NEE between the different data sources and the resulting data-driven flux models.

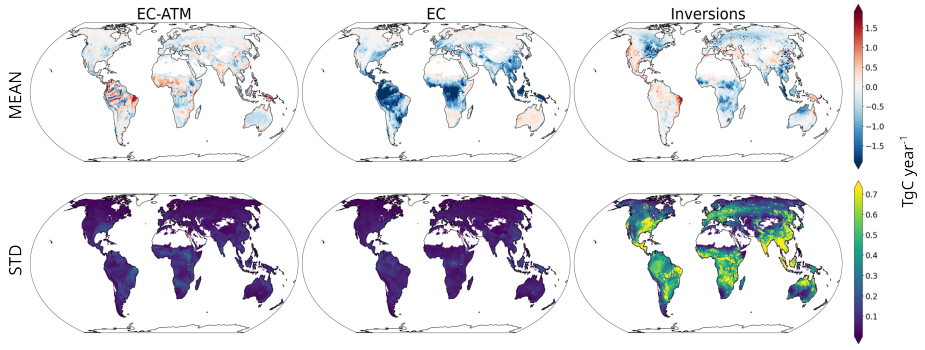

Figure 5. Mean (top panel) and standard deviation (bottom-panel) of the ensemble mean annual NEE from the EC-ATM and EC models, compared to the ensemble mean of the atmospheric inversions. The top row shows the per-pixel mean of the annual NEE. The bottom row shows the per-pixel standard deviation of the annual NEE in TgC year$^{-1.}$

Figure 7 shows that both the EC-ATM and EC models correlate well with the inversion mean in the extratropics where both products are better constrained by observations, while the EC ensemble mean has a negative correlation with the inversion mean in the tropics. The EC-ATM shows a corresponding negative correlation with FLXUCOM RS+METEO V1 in tropical forests. We note, however, that there is low confidence in atmospheric inversions' estimates at the pixel level. Nevertheless, there is an overall reduction in correlation between the EC-ATM ensemble and the EC ensemble with the FLUXCOM RS+METEO V1 spatial distribution (Fig 7, Difference C-D) with a only partially corresponding increase in correlation with the inversion mean (A-B). This demonstrates that the model is not simply learning a bias correction, but a new pattern of land flux that incorporates some but not all of the information in the atmospheric inversions.

4.4 Eddy-covariance site comparison

The comparison of EC and EC-ATM modelled NEE at the eddy-covariance sites (Fig. 8) shows that the EC-ATM and EC models have a similar RMSE with observed NEE globally. The both the EC and EC-ATM model RMSE performances of 1.349 and 1.321 gC m^{-2} day^{-1} respectively are similar to the median results of RMSE for NEE using the setup in Tramontana et al. (2016) of 1.298 gC m^{-2} day^{-1}. The optimization of the model hyperparameters for EC-ATM performance, may lead to a slight underperfomance by the EC model. The RMSE for the EC and EC-ATM ensembles is very similar globally and by

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

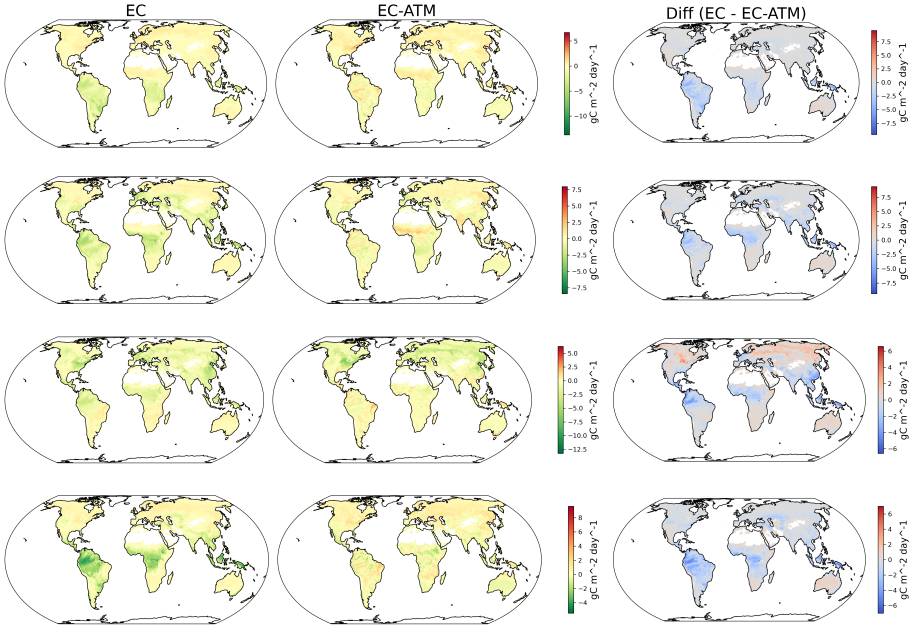

Figure 6. Monthly mean fluxes for 2001-2017. The left column shows the EC-ATM results, the middle column shows EC results and the right column shows the difference between them (EC - EC-ATM).

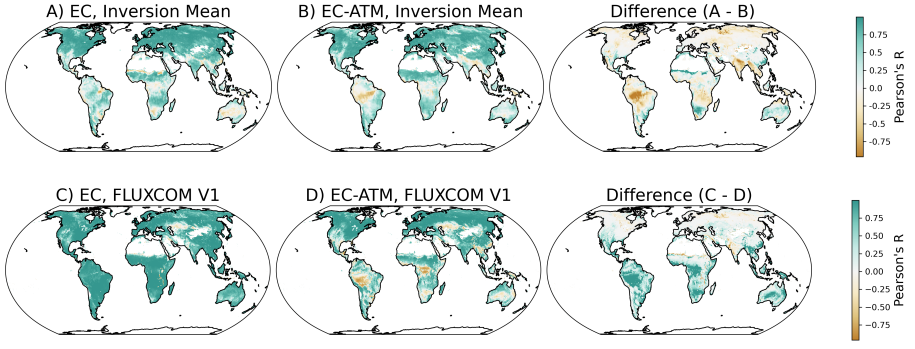

Figure 7. Pearson's R for pairwise spatial distributions. The correlation is calculated using the time-series of each pixel.

majority PFT. A breakdown of tower performance by majority PFT is available in the appendix C2. This may indicate that for the majority of sites the additional information acts as a complementary constraint, very slightly improving the EC-ATM model response at the site level. However, in some cases the model learns different responses in some tropical eddy-covariance towers (see Fig 9). This shows where the atmospheric constraint adds information that allows the EC-ATM model to better adapt to potential biases in the tower data from its additional exposure to non-tower pixels. In figure 9, the EC and EC-ATM model results at the tower are plotted against the observed NEE. At BR-Ji2 (left) the two data-driven flux models have very

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

similar responses, which show a reduced sink compared to the tower observations. At BR-Ma2 (right) the EC-ATM ensemble has learned an overall reduced sink compared with the EC ensemble.

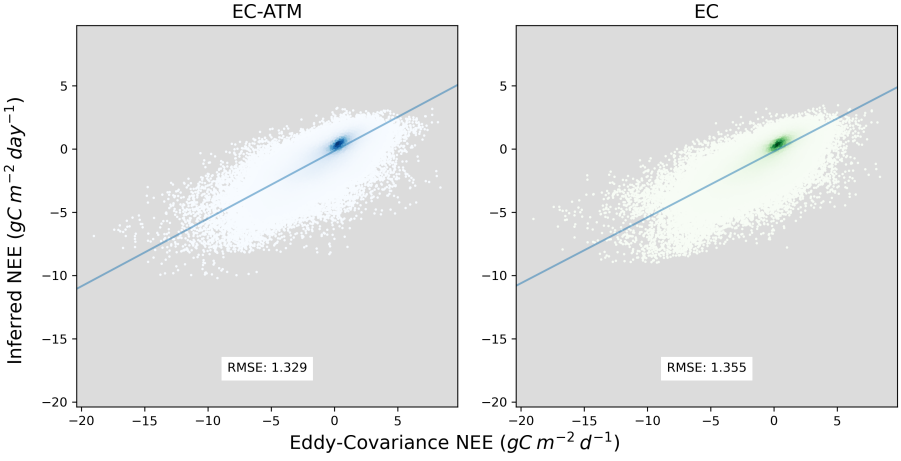

Figure 8. Comparison of EC tower inferences with EC and EC-ATM model output across the whole training set. The Y axes are the NEE EC-ATM and NEE EC in gC m^{-2} day^{-1}, the X axes are the eddy-covariance observations in gC m^{-2} day^{-1}.

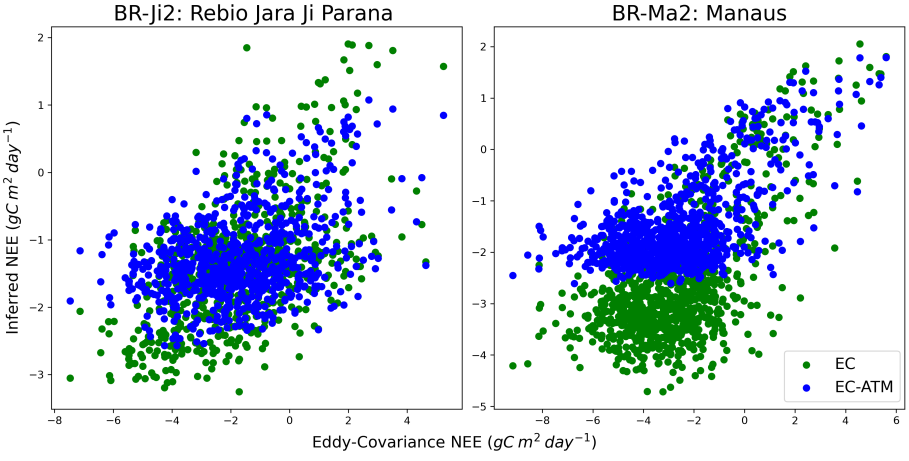

Figure 9. Comparison of EC and EC-ATM model output at two Brazilian eddy-covariance sites. X axes are the model estimates and the Y axes are the tower observations.

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

5 Discussion

5.1 Constraint versus bias-correction

The addition of atmospheric information to a bottom-up data-driven flux model has several effects. It acts as a constraint, adding new information about the response in NEE from the drivers. The EC-ATM model results show changes in the seasonal pattern and magnitude of the MSC across regions (see Fig C1), indicating that the atmospheric information in EC-ATM is modifying the data-driven flux model's (EC) response away from the eddy-covariance measurements. This indicates that the additional information does not create a simple global or regional bias-correction term. This is also evident in the correction at some tower locations (Fig 9). The difference between the observed and inferred NEE by the EC and EC-ATM models at specific towers shows that the atmospheric information can also act locally during training.

5.2 Interannual Variability (IAV)

This study shows very limited improvement in the IAV magnitude estimated by EC-ATM compared with the FLUXCOM RS+METEO V1 results, it is still well below the IAV of the atmospheric inversions. This may be attributed to the optimization of the neural network, the formulation of the driver variables, which are common to the FLUXCOM RS+METEO V1 results, and to missing information in the training set, for example due to under-representation of semi-arid tropical region (Ahlström et al., 2015; Poulter et al., 2014) in the FLUXNET network. In this discussion of IAV, we use the standard deviation of the detrended yearly integrals. Detrending the data removes the influence of land-use change (LUC) (Friedlingstein et al., 2022) which is not represented in the eddy-covariance data.

The absolute value of the IAV component of the signal is small compared to the MSC, so the data-driven flux model may tend to optimize for the MSC, which will be the major component of the objective function term, while ignoring the IAV. Experiments (not shown) with the architecture of the data-driven flux model and the formulation of the loss were not successful in producing an IAV that is consistent with the atmospheric inversions. The structure of the model was varied, both in depth (number of hidden layers) and width (number of neurons per layer). Increasing network complexity on both axes tended to reduce the estimated IAV, while marginally increasing the accuracy of the MSC. Experiments were conducted on the objective function both structurally, and in its formulation. Structurally, the number of regions included in the atmospheric term at each step, and the number of months included at each step were varied. The best results, as indicated in the methods, were the maximum; all regions run for a full year at each training step. Additionally we experimented adding a third global term in the objective function, formulated as the integral of all regions over a year with compared with the yearly integral of the ensemble of atmospheric inversions. This term should directly inform the model about its performance with regard to the IAV, but our experiments did not produce a better IAV estimate with this term included. We also experimented with the formulation of the objective term, with a small increase in IAV resulting from the use of the ℓ_3-norm compared with a MSE (ℓ_2-norm), or mean absolute error.

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

The driver variables are largely constructed from the MSC of the underlying remotely-sensed or meteorological data directly (Table A1), or with data derived from the MSC, such as the minimum or range of the MSC for the driver variable. This limits the amount of information available for the EC or EC-ATM model to learn about the IAV component of the signal.

The eddy-covariance sites represented in the training set do not sufficiently represent important biomes that are important for the global IAV of NEE, particularly semi-tropical arid regions (Poulter et al., 2014). These regions have been shown to control the IAV of NEE while the mean of global NEE is largely controlled by tropical forests (Ahlström et al., 2015). This missing influence may also degrade the ability of the data-driven flux model to fully capture the IAV.

5.3 Outlook and Challenges

This study demonstrates both the impact of regional atmospheric information on the training of a data-driven flux model, and the framework for using a bridge model to link a data-driven flux model to additional constraints. In this case, these are regional integrals of NEE from atmospheric inversions, but these can be, in principle, multiple data streams at different scales (temporal, spatial), or in different formats (grid, point). Incorporating these different data streams would require different model formulations, potentially including neural network architecture and objective functions, as well as data-driven, or physics-based bridge models to create the link the data-driven flux-model to these new data. Our aim here is to demonstrate that adding an atmospheric "top-down" constraint can positively impact the evolution of a "bottom-up" data-driven flux model during training, leading to meaningful improvement in the data-driven flux model outputs.

While demonstrating the ability of top-down constraint to partially resolve the mismatch between the atmospheric inversions and the bottom up data-driven estimates of NEE, These results also shows the potential for a confounding effect from the training process. The EC-ATM model is a learned statistical response between the drivers and the training data. There are mismatches between the EC-ATM inference of NEE and the atmospheric data used for the top-down constraint. The atmospheric data still implicitly includes disturbance and trade fluxes, along with other flux components that are not accounted for in our model. This means that in reducing the the MSC error, these flux components are implicitly incorporated into the EC-ATM inference, although the model lacks the necessary process information, since these are not included in the drivers. And, because of the statistical nature of the training process, the data-driven model is not an analog for a process-based model where terms can be more easily backed out. Training using our double constraint should become less confounded as more additional spatially-explicit flux components become available. However, despite these data mismatches, training a data-driven flux model using a dual constraint does create a useful estimate of the NEE at multiple scales.

6 Conclusions

This study demonstrates that a bottom-up data-driven flux model derived from eddy-covariance data can be successfully constrained by top-down observations of atmospheric carbon dioxide. The global and regional integrals of NEE preserve the strengths of top-down estimates, while the spatial distribution remains closer to the better spatially-resolved bottom-up es-

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

timates. The global annual NEE in the EC-ATM model with the atmospheric constraint is closer to the global annual NEE estimated by independent datasets in the Global Carbon Budget 2022 Friedlingstein et al. (2022).

This multi-scale approach may allow us to leverage the large volume of atmospheric carbon dioxide mole fraction observations for constraining a data-driven flux model. Using an atmospheric transport model as the bridge, the data-driven flux model can generate an inference of the mole fraction at a location by estimating the flux at the set of locations that contributed to that observation. Because this set of contributing locations would vary with the local meteorology, covering a range of biomes, the data-driven flux model would see a more diverse training set. This may improve the performance of the data-driven flux model by learning from a more representative distribution of the driver variables across the land surface. In the future, this logic could also be extended to other datasets, for example satellite retrievals of xCO_2, pairing the deep archive of EC observations with novel 'flux towers in the sky' (Schimel et al., 2019).

Data availability. Working on a DOI

Appendix A: Data

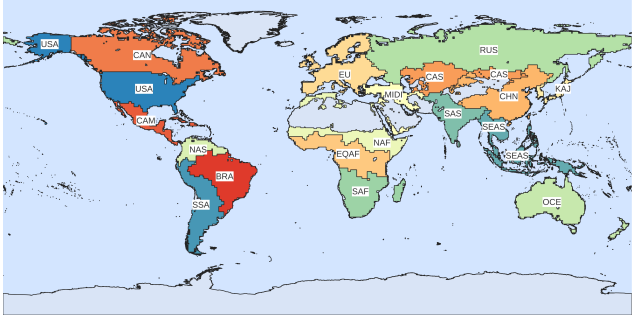

Figure A1. RECCAP2 Regions

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

Table A1. RECCAP2 region ID, names and abbreviations

ID	Region	Abbreviation
0	United States	USA
1	Canada	CAN
2	Central America	CAM
3	Northern South America	NSA
4	Brazil	BRA
5	Southwest South America	SSA
6	Europe	EU
7	Northern Africa	NAF
8	Equatorial Africa	EQAF
9	Southern Africa	SAF
10	Russia	RUS
11	Central Asia	CAS
12	Mideast	MIDE
13	China	CHN
14	Korea and Japan	KAJ
15	South Asia	SAS
16	Southeast Asia	SEAS
17	Oceania	OCE

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

Table A2. Driver variables used for the data-driven EC and EC-ATM models and the calculation of the drivers from the base variables above. The global dataset uses only the MODIS and ERA5 data, while the data used at the eddy-covariance sites also uses meteorological observations from the tower instruments. See Tramontana et al. (2016) for a full discussion.

Name	**Variable**	**MSC calculation period**	(Source)
WAI2	Water Availability Index		Calculated from a water balance model derived from MODIS variables (see Tramontana et al. (2016) supplement S3 for full model description)
MSC_EVIRpot	Mean Season Cycle(EVI $\times$ R_g)	2001–2012	MODIS (EVI), ERA5 (R_g)
MSC_FparLST	Mean Season Cycle(fAPAR $\times$ LST_{day})	2001–2012	MODIS
MIN_MSC_NDWI	Min(Mean Season Cycle(NDWI))	2001–2012	MODIS
AMP_Band4	Amplitude(band 4 reflectance)		MODIS
MSC_LST_Night	Mean Season Cycle(LST_{night})	2001–2012	MODIS
Rg_VIMSC	Mean Season Cycle(NDVI) $\times$ R_{pot}	2001–2012	MODIS (NDVI), ERA5 (R_g)
AMP_MSC_NDVI	Amplitude(Mean Season Cycle(NDVI))	2001–2012	MODIS
Tair	Air Temperature	2001–2012	ERA5
AMP_MSC_WAI	Amplitude(Mean Season Cycle(WAI))	2001–2012	ERA5

All variables: Daily values 2000-2017, 0.5°spatial resolution

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

Table B1. Hyperparameters for reported EC and EC-ATM model runs

Input shape	10
Latent dimension	32
Learning rate	0.003
EC batch	10000

Appendix B: Technical Implementation

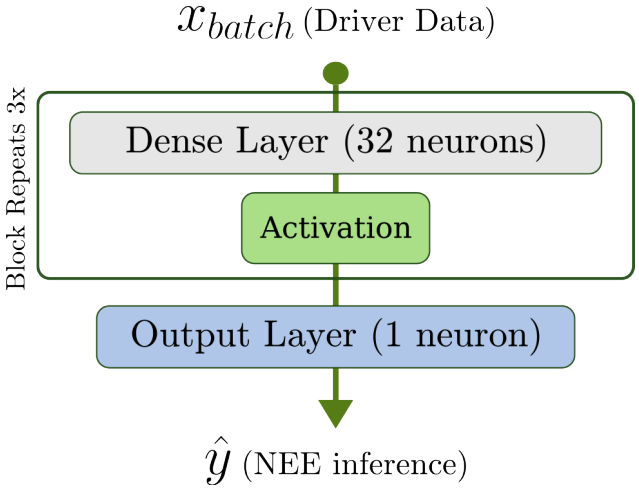

Figure B1. Model Architecture, the model is a feed-forward neural network, or a set of fully connected network layers. The fully-connected layers consist of nodes or 'neurons', which are exposed to the output of all neurons in the previous layer. Non-linearity is introduced by passing each node output through a non-linear activation function. Our network is a set of three fully-connected layers with the ReLU activation function (Agarap, 2019).

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

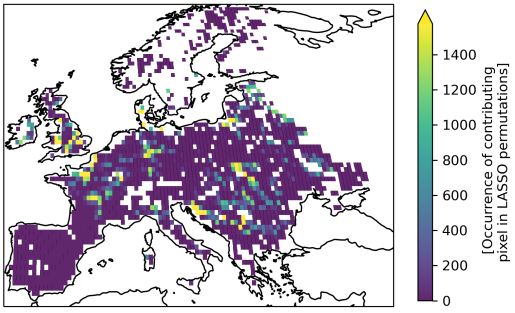

Figure B2. Robustness of contribution pixel selection: A Heat map of pixel inclusion in the sparse linear model using Lasso regression. Values represent the log-scaled number of pixel inclusions in the non-zero set of parameters across 500 regressions using a randomized subset of the data. Pixels that are most often included provide a more important constraint to the calculation of a regionally summed NEE, minimizing Eq. 2

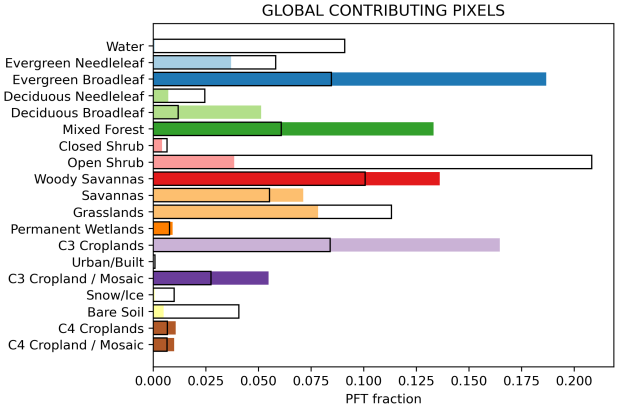

Figure B3. The representation of PFTs across all contributing pixels in all regions. All PFTs are the majority type per pixel. This image shows the relative number of times a certain PFT is included in the optimal set of contributing pixels which construct a regional integral of NEE, when selecting from all global land pixels. The black outlines show the proportion of that majority PFT type globally. A per-region analysis of PFT inclusion is available in appendix B4

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

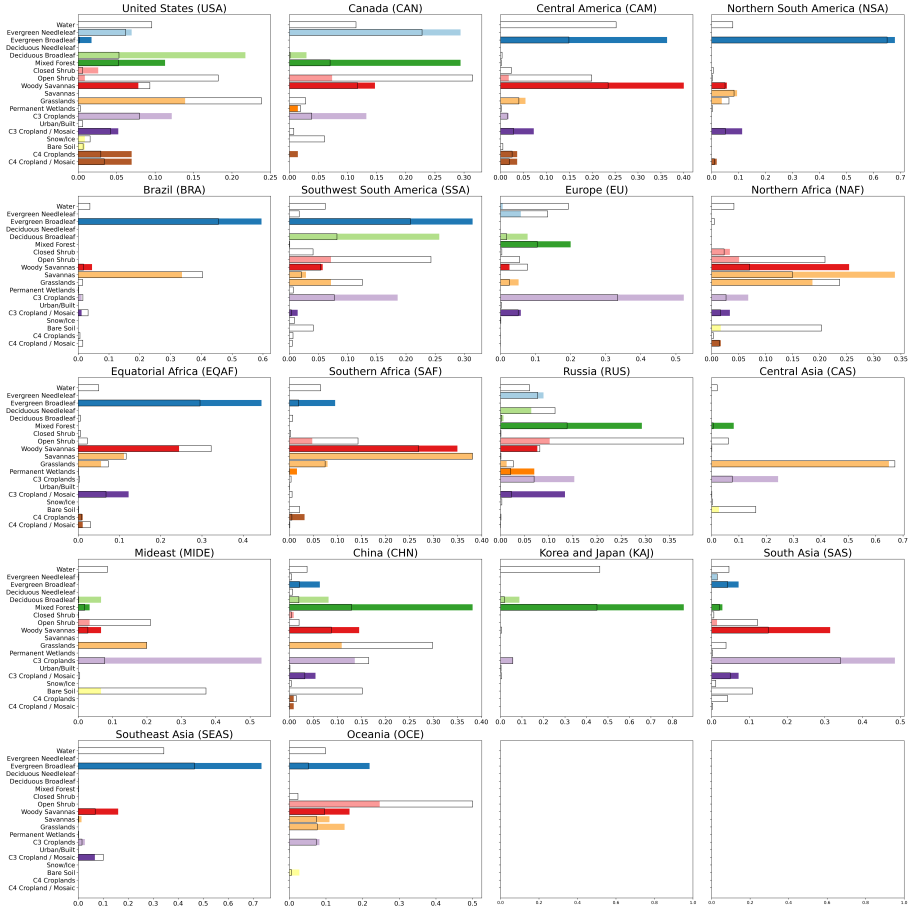

Figure B4. Regional composition of PFT in contributing pixels

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

Appendix C: Results

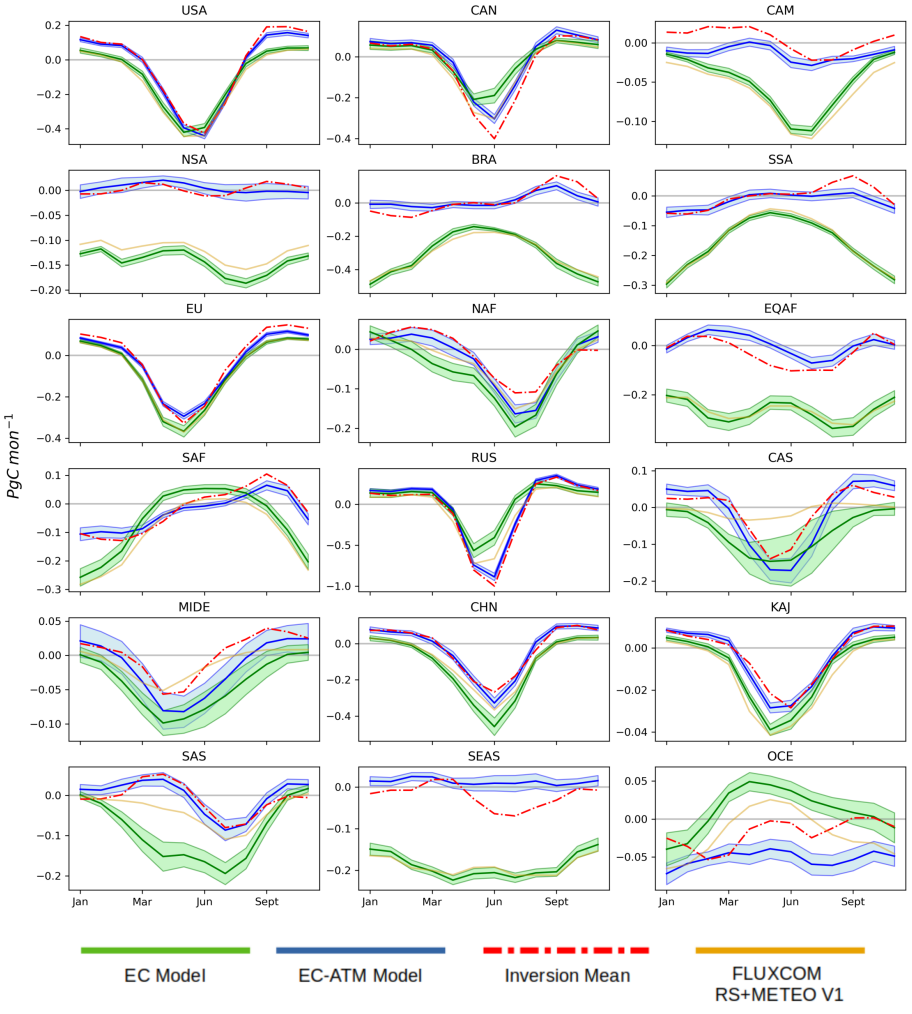

Figure C1. MSC of the ensemble mean of all regions in PgC mon^{-1}. The solid line is the ensemble mean, and the shaded region is the mean ± the ensemble standard deviation.

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

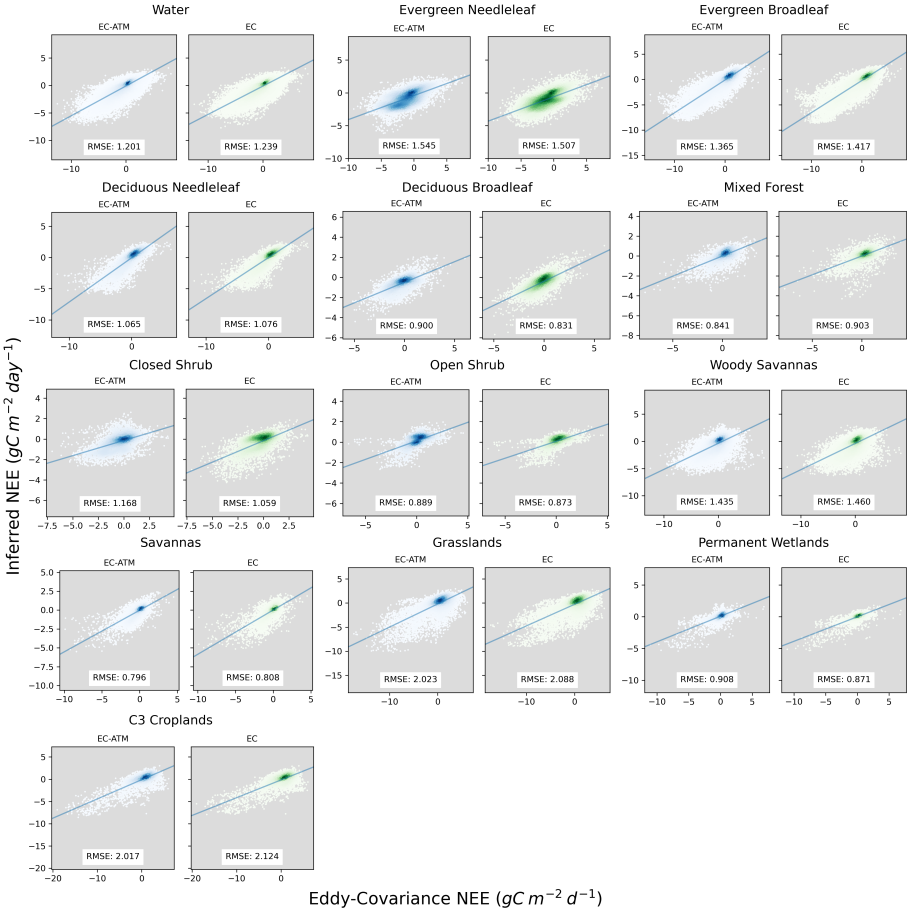

Figure C2. Scatter-plots of eddy-covariance NEE (x axis) and inferred NEE (y axis) by PFT and model (ATM-EC, EC).

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

Author contributions. SU, AB, MR, FG designed the study. SU performed the analysis, and drafted the manuscript. AB, MR, WP provided analysis and support. BK provided expertise with the machine learning framework. All authors revised and edited the text.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We would like to thank Martin Jung, Jakob A. Nelson, Sophia Walther, and the FLUXCOM team for their structural support, feedback and discussion. The Authors would like to thank the producers of the Inversion data included in this study: Ingrid Luijkx and Wouter Peters (CTE), Frederic Chevallier and the Copernicus Atmosphere Monitoring Service (CAMS), Christian Roedenbeck (Jena Carboscope sEXTocNEET), Yosuke Niwa (NISMON-CO2), and Liang Feng and Paul Palmer (UoE). We would like to thank the team behind the European Research Council (ERC) Synergy Grant 'Understanding and modeling the Earth System with Machine Learning (USMILE)' for their feedback and discussion.

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

References

Agarap, A. F.: Deep Learning using Rectified Linear Units (ReLU), https://doi.org/10.48550/arXiv.1803.08375, arXiv:1803.08375 [cs, stat], 2019.

Ahlström, A., Raupach, M. R., Schurgers, G., Smith, B., Arneth, A., Jung, M., Reichstein, M., Canadell, J. G., Friedlingstein, P., Jain, A. K., Kato, E., Poulter, B., Sitch, S., Stocker, B. D., Viovy, N., Wang, Y. P., Wiltshire, A., Zaehle, S., and Zeng, N.: The dominant role of semi-arid ecosystems in the trend and variability of the land CO_2 sink, Science, 348, 895–899, https://doi.org/10.1126/science.aaa1668, _eprint: https://www.science.org/doi/pdf/10.1126/science.aaa1668, 2015.

Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., Murray-Tortarolo, G., Papale, D., Parazoo, N. C., Peylin, P., Piao, S., Sitch, S., Viovy, N., Wiltshire, A., and Zhao, M.: Spatiotemporal patterns of terrestrial gross primary production: A review, Reviews of Geophysics, 53, 785–818, https://doi.org/10.1002/2015RG000483, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2015RG000483, 2015.

Bastos, A., O'Sullivan, M., Ciais, P., Makowski, D., Sitch, S., Friedlingstein, P., Chevallier, F., Rödenbeck, C., Pongratz, J., Luijkx, I. T., Patra, P. K., Peylin, P., Canadell, J. G., Lauerwald, R., Li, W., Smith, N. E., Peters, W., Goll, D. S., Jain, A., Kato, E., Lienert, S., Lombardozzi, D. L., Haverd, V., Nabel, J. E. M. S., Poulter, B., Tian, H., Walker, A. P., and Zaehle, S.: Sources of Uncertainty in Regional and Global Terrestrial CO2 Exchange Estimates, Global Biogeochemical Cycles, 34, e2019GB006 393, https://doi.org/10.1029/2019GB006393, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019GB006393, 2020.

Bastos, A., Ciais, P., Sitch, S., Aragão, L. E. O. C., Chevallier, F., Fawcett, D., Rosan, T. M., Saunois, M., Günther, D., Perugini, L., Robert, C., Deng, Z., Pongratz, J., Ganzenmüller, R., Fuchs, R., Winkler, K., Zaehle, S., and Albergel, C.: On the use of Earth Observation to support estimates of national greenhouse gas emissions and sinks for the Global stocktake process: lessons learned from ESA-CCI RECCAP2, Carbon Balance and Management, 17, 15, https://doi.org/10.1186/s13021-022-00214-w, 2022.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M. A., Baldocchi, D., Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K. W., Roupsard, O., Veenendaal, E., Viovy, N., Williams, C., Woodward, F. I., and Papale, D.: Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate, Science, 329, 834–838, https://doi.org/10.1126/science.1184984, publisher: American Association for the Advancement of Science, 2010.

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F.-M., Chédin, A., and Ciais, P.: Inferring CO2 sources and sinks from satellite observations: Method and application to TOVS data, Journal of Geophysical Research: Atmospheres, 110, https://doi.org/10.1029/2005JD006390, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2005JD006390, 2005.

Chu, H., Baldocchi, D. D., John, R., Wolf, S., and Reichstein, M.: Fluxes all of the time? A primer on the temporal representativeness of FLUXNET, Journal of Geophysical Research: Biogeosciences, 122, 289–307, https://doi.org/10.1002/2016JG003576, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2016JG003576, 2017.

Ciais, P., Bastos, A., Chevallier, F., Lauerwald, R., Poulter, B., Canadell, J. G., Hugelius, G., Jackson, R. B., Jain, A., Jones, M., Kondo, M., Luijkx, I. T., Patra, P. K., Peters, W., Pongratz, J., Petrescu, A. M. R., Piao, S., Qiu, C., Von Randow, C., Regnier, P., Saunois, M., Scholes, R., Shvidenko, A., Tian, H., Yang, H., Wang, X., and Zheng, B.: Definitions and methods to estimate regional land carbon fluxes for the second phase of the REgional Carbon Cycle Assessment and Processes Project (RECCAP-2), Geoscientific Model Development, 15, 1289–1316, https://doi.org/10.5194/gmd-15-1289-2022, publisher: Copernicus GmbH, 2022.

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

Di Giuseppe, F., Rémy, S., Pappenberger, F., and Wetterhall, F.: Using the Fire Weather Index (FWI) to improve the estimation of fire emissions from fire radiative power (FRP) observations, Atmospheric Chemistry and Physics, 18, 5359–5370, https://doi.org/10.5194/acp-18-5359-2018, publisher: Copernicus GmbH, 2018.

Feng, L., Palmer, P. I., Parker, R. J., Deutscher, N. M., Feist, D. G., Kivi, R., Morino, I., and Sussmann, R.: Estimates of European uptake of CO_2 inferred from GOSAT X_{CO_{2}} retrievals: sensitivity to measurement bias inside and outside Europe, Atmospheric Chemistry and Physics, 16, 1289–1302, https://doi.org/10.5194/acp-16-1289-2016, publisher: Copernicus GmbH, 2016.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., and Huang, X.: MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets, Remote Sensing of Environment, 114, 168–182, https://doi.org/10.1016/j.rse.2009.08.016, 2010.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Bakker, D. C. E., Hauck, J., Le Quéré, C., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., Bates, N. R., Becker, M., Bellouin, N., Bopp, L., Chau, T. T. T., Chevallier, F., Chini, L. P., Cronin, M., Currie, K. I., Decharme, B., Djeutchouang, L. M., Dou, X., Evans, W., Feely, R. A., Feng, L., Gasser, T., Gilfillan, D., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, O., Harris, I., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Luijkx, I. T., Jain, A., Jones, S. D., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lienert, S., Liu, J., Marland, G., McGuire, P. C., Melton, J. R., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Niwa, Y., Ono, T., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T. M., Schwinger, J., Schwingshackl, C., Séférian, R., Sutton, A. J., Sweeney, C., Tanhua, T., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F., van der Werf, G. R., Vuichard, N., Wada, C., Wanninkhof, R., Watson, A. J., Willis, D., Wiltshire, A. J., Yuan, W., Yue, C., Yue, X., Zaehle, S., and Zeng, J.: Global Carbon Budget 2021, Earth System Science Data, 14, 1917–2005, https://doi.org/10.5194/essd-14-1917-2022, publisher: Copernicus GmbH, 2022.

Fu, Z., Gerken, T., Bromley, G., Araújo, A., Bonal, D., Burban, B., Ficklin, D., Fuentes, J. D., Goulden, M., Hirano, T., Kosugi, Y., Liddell, M., Nicolini, G., Niu, S., Roupsard, O., Stefani, P., Mi, C., Tofte, Z., Xiao, J., Valentini, R., Wolf, S., and Stoy, P. C.: The surface-atmosphere exchange of carbon dioxide in tropical rainforests: Sensitivity to environmental drivers and flux measurement methodology, Agricultural and Forest Meteorology, 263, 292–307, https://doi.org/10.1016/j.agrformet.2018.09.001, 2018.

Gaubert, B., Stephens, B. B., Basu, S., Chevallier, F., Deng, F., Kort, E. A., Patra, P. K., Peters, W., Rödenbeck, C., Saeki, T., Schimel, D., Van der Laan-Luijkx, I., Wofsy, S., and Yin, Y.: Global atmospheric CO2 inverse models converging on neutral tropical land exchange, but disagreeing on fossil fuel and atmospheric growth rate, Biogeosciences, 16, 117–134, https://doi.org/10.5194/bg-16-117-2019, publisher: Copernicus GmbH, 2019.

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G., and Reichstein, M.: The FLUXCOM ensemble of global land-atmosphere energy fluxes, Scientific Data, 6, 74, https://doi.org/10.1038/s41597-019-0076-8, bandiera_abtest: a Cc_license_type: cc_publicdomain Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Climate sciences;Environmental sciences;Hydrology Subject_term_id: climate-sciences;environmental-sciences;hydrology, 2019.

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P., Besnard, S., Bodesheim, P., Carvalhais, N., Chevallier, F., Gans, F., Goll, D. S., Haverd, V., Köhler, P., Ichii, K., Jain, A. K., Liu, J., Lombardozzi, D., Nabel, J. E. M. S., Nelson, J. A., O'Sullivan, M., Pallandt, M., Papale, D., Peters, W., Pongratz, J., Rödenbeck, C., Sitch, S., Tramontana, G., Walker, A., Weber, U., and Reichstein, M.: Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach, Biogeosciences, 17, 1343–1365, https://doi.org/10.5194/bg-17-1343-2020, publisher: Copernicus GmbH, 2020.

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

Kaminski, T. and Heimann, M.: Inverse Modeling of Atmospheric Carbon Dioxide Fluxes, Science, https://doi.org/10.1126/science.294.5541.259a, publisher: American Association for the Advancement of Science, 2001.

Keeling, C. D., Bacastow, R. B., Carter, A. F., Piper, S. C., Whorf, T. P., Heimann, M., Mook, W. G., and Roeloffzen, H.: A three-dimensional model of atmospheric CO2 transport based on observed winds: 1. Analysis of observational data, in: Aspects of Climate Variability in the Pacific and the Western Americas, pp. 165–236, American Geophysical Union (AGU), https://doi.org/10.1029/GM055p0165, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/GM055p0165, 1989.

Kelley, H. J.: Gradient theory of optimal flight paths, Ars Journal, 30, 947–954, 1960.

Kendall, A., Gal, Y., and Cipolla, R.: Multi-task learning using uncertainty to weigh losses for scene geometry and semantics, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7482–7491, 2018.

Kondo, M., Patra, P. K., Sitch, S., Friedlingstein, P., Poulter, B., Chevallier, F., Ciais, P., Canadell, J. G., Bastos, A., Lauerwald, R., Calle, L., Ichii, K., Anthoni, P., Arneth, A., Haverd, V., Jain, A. K., Kato, E., Kautz, M., Law, R. M., Lienert, S., Lombardozzi, D., Maki, T., Nakamura, T., Peylin, P., Rödenbeck, C., Zhuravlev, R., Saeki, T., Tian, H., Zhu, D., and Ziehn, T.: State of the science in reconciling top-down and bottom-up approaches for terrestrial CO2 budget, Global Change Biology, 26, 1068–1084, https://doi.org/10.1111/gcb.14917, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.14917, 2020.

Niwa, Y., Ishijima, K., Ito, A., and Iida, Y.: Toward a long-term atmospheric CO2 inversion for elucidating natural carbon fluxes: technical notes of NISMON-CO2 v2021.1, Progress in Earth and Planetary Science, 9, 42, https://doi.org/10.1186/s40645-022-00502-6, 2022.

Palmer, P. I., Feng, L., Baker, D., Chevallier, F., Bösch, H., and Somkuti, P.: Net carbon emissions from African biosphere dominate pan-tropical atmospheric CO2 signal, Nature Communications, 10, 3344, https://doi.org/10.1038/s41467-019-11097-w, number: 1 Publisher: Nature Publishing Group, 2019.

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K., Peters, W., Rayner, P. J., Rödenbeck, C., van der Laan-Luijkx, I. T., and Zhang, X.: Global atmospheric carbon budget: results from an ensemble of atmospheric CO_2 inversions, Biogeosciences, 10, 6699–6720, https://doi.org/10.5194/bg-10-6699-2013, publisher: Copernicus GmbH, 2013.

Poulter, B., Frank, D., Ciais, P., Myneni, R. B., Andela, N., Bi, J., Broquet, G., Canadell, J. G., Chevallier, F., Liu, Y. Y., Running, S. W., Sitch, S., and van der Werf, G. R.: Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle, Nature, 509, 600–603, https://doi.org/10.1038/nature13376, number: 7502 Publisher: Nature Publishing Group, 2014.

Rödenbeck, C.: Estimating CO2 sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport, Technical reports, 2005, 2005.

Rödenbeck, C., Houweling, S., Gloor, M., and Heimann, M.: CO2 flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, Atmospheric Chemistry and Physics, 3, 1919–1964, https://doi.org/10.5194/acp-3-1919-2003, publisher: Copernicus GmbH, 2003.

Rödenbeck, C., Zaehle, S., Keeling, R., and Heimann, M.: How does the terrestrial carbon exchange respond to inter-annual climatic variations? A quantification based on atmospheric CO_2 data, Biogeosciences, 15, 2481–2498, https://doi.org/10.5194/bg-15-2481-2018, publisher: Copernicus GmbH, 2018.

Saeki, T. and Patra, P. K.: Implications of overestimated anthropogenic CO2 emissions on East Asian and global land CO2 flux inversion, Geoscience Letters, 4, 9, https://doi.org/10.1186/s40562-017-0074-7, 2017.

Schimel, D., Schneider, F. D., and JPL Carbon and Ecosystem Participants: Flux towers in the sky: global ecology from space, New Phytologist, 224, 570–584, https://doi.org/10.1111/nph.15934, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/nph.15934, 2019.

https://doi.org/10.5194/egusphere-2023-805
Preprint. Discussion started: 5 June 2023

Thompson, R. L., Patra, P. K., Chevallier, F., Maksyutov, S., Law, R. M., Ziehn, T., van der Laan-Luijkx, I. T., Peters, W., Ganshin, A., Zhuravlev, R., Maki, T., Nakamura, T., Shirai, T., Ishizawa, M., Saeki, T., Machida, T., Poulter, B., Canadell, J. G., and Ciais, P.: Top–down assessment of the Asian carbon budget since the mid 1990s, Nature Communications, 7, 10 724, https://doi.org/10.1038/ncomms10724, number: 1 Publisher: Nature Publishing Group, 2016.

Tian, H., Yang, J., Lu, C., Xu, R., Canadell, J. G., Jackson, R. B., Arneth, A., Chang, J., Chen, G., Ciais, P., Gerber, S., Ito, A., Huang, Y., Joos, F., Lienert, S., Messina, P., Olin, S., Pan, S., Peng, C., Saikawa, E., Thompson, R. L., Vuichard, N., Winiwarter, W., Zaehle, S., Zhang, B., Zhang, K., and Zhu, Q.: The Global N2O Model Intercomparison Project, Bulletin of the American Meteorological Society, 99, 1231–1251, https://doi.org/10.1175/BAMS-D-17-0212.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society, 2018.

Tormene, P., Giorgino, T., Quaglini, S., and Stefanelli, M.: Matching Incomplete Time Series with Dynamic Time Warping: An Algorithm and an Application to Post-Stroke Rehabilitation, Artificial Intelligence in Medicine, 45, 11–34, https://doi.org/10.1016/j.artmed.2008.11.007, 2008.

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M. A., Cescatti, A., Kiely, G., Merbold, L., Serrano-Ortiz, P., Sickert, S., Wolf, S., and Papale, D.: Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms, Biogeosciences, 13, 4291–4313, https://doi.org/10.5194/bg-13-4291-2016, publisher: Copernicus GmbH, 2016.

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