Constraining a data-driven CO₂ flux model by ecosystem and atmospheric observations using atmospheric transport

Samuel Upton[1,4], Markus Reichstein[1], Wouter Peters[4,5], Santiago Botía[2], Jacob A. Nelson[1], Sophia Walther[1], Martin Jung[1], Fabian Gans[1], László Haszpra[6,7], and Ana Bastos[1,3]

[1]Department of Biogeochemical Integration, Max Plank Institute of Biogeochemistry, Jena Germany
[2]Department of Biogeochemical Signals, Max Plank Institute of Biogeochemistry, Jena Germany
[3]Institute for Earth System Science and Remote Sensing, Leipzig University
[4]Environmental Sciences Group, Wageningen University, Wageningen, The Netherlands
[5]University of Groningen, Centre for Isotope Research, Groningen, The Netherlands
[6]International Radiocarbon AMS Competence and Training Center, HUN-REN Institute for Nuclear Research, Debrecen, Hungary
[7]Institute of Earth Physics and Space Science, Sopron, Hungary

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

Abstract.

Global estimates of the terrestrial land-atmosphere flux of CO₂ (NEE) from data-driven models differ widely depending on their underlying data and methodology. Bottom-up models trained on eddy-covariance data are most informative at the ecosystem-level. Top-down models, such as atmospheric inversions, produce regional and global results consistent with the observed atmospheric growth rate, accurately capturing the interannual variability (IAV) of NEE. Both approaches have limitations estimating NEE across scales: Bottom-up models can miss large-scale dynamics of NEE when aggregated globally. Top-down approaches have difficulty relating the large-scale atmospheric signal to biophysical processes at smaller scales. To address these limitations, we create a model that uses a hybrid combination of direct observations and atmospheric dynamics to integrate ecosystem-level eddy-covariance data and atmospheric CO₂ mole fraction data into a single coherent ecosystem-level flux model.

Aggregated globally, our new model estimates an annual sink with a low bias, and consistent IAV when compared with independent estimates. The IAV of the estimated NEE is closer in magnitude to an ensemble of atmospheric inversions, and our model produces a higher temporal coefficient of correlation with these data than state-of-the-art bottom-up data-driven models. This improvement in IAV is achieved without direct access to the observed variability of the atmosphere: the model is trained using only one year of daytime observations from 3 tall-tower observatories. No atmospheric information is available to the model during the production of global NEE estimates. This shows the efficiency of our method in synthesizing top-down information into bottom-up mapping of flux-environment relationships.

1 Introduction

The global net flux of biogenic CO_2 between the land surface and the atmosphere, or net ecosystem exchange (NEE), is a
20 critical but uncertain term in our understanding of the global carbon budget and the climate system (Friedlingstein et al., 2014;
Bastos et al., 2022). Current data-driven approaches to modeling NEE can be largely grouped into two categories: top-down
approaches that infer NEE from observations of atmospheric CO_2 by prescribing fire and fossil fuel emissions. The second
category are bottom-up approaches that model NEE at the ecosystem level from local observations (Kondo et al., 2020). Both
approaches provide important information about the function of the biosphere, but have their limitations.

25 Data-driven bottom-up models are most commonly informed by eddy-covariance observations at the ecosystem scale, along
with remotely sensed variables (greenness and water indices, spectral data) (Nelson* and Walther* et al., 2024; Jung et al.,
2020; Kondo et al., 2020; Bodesheim et al., 2018). However, the spatial distribution of NEE and the total magnitude of NEE
at global scale remain highly uncertain as the available eddy-covariance data is sparse in important biomes across the globe,
like the tropical rain forests (Chu et al., 2017; Hayek et al., 2018; Fu et al., 2018; Jung et al., 2020). Some processes, such as
30 carbon emissions from fires, are not captured by these local observations. Eddy-covariance data are further potentially subject
to systematic errors overestimating the carbon uptake (Aubinet et al., 2005), particularly in the tropics due to complex CO_2
nighttime storage and horizontal advection in the canopy (Fu et al., 2018; Moncrieff et al., 1996). Historically, when these
eddy-covariance data are used to train a data-driven bottom-up flux model for global upscaling of NEE (e.g. FLUXCOM V1,
Jung et al. (2020)), the underlying issues of data availability and representation are propagated into the global result resulting in
35 an overestimation of the tropical carbon sink and by extension, the global carbon sink. Therefore data-driven global products
such as FLUXCOM V1 NEE, which are very dependent on the quality and completeness of the training data, are not fully
consistent with the magnitude of the growth rate of atmospheric CO_2, or the interannual variability (IAV) of the global NEE
signal (Jung et al., 2020; Kondo et al., 2015). However the length of the available eddy-covariance record has increased, and
the data collection and processing have improved (Pastorello et al., 2020). These longer, improved data, along with improved
40 handling and gap-filling of the complementary co-located remotely sensed variables (Walther* and Besnard* et al., 2022)
have allowed for FLUXCOM X-BASE to improve its estimates of the magnitude of regional and global NEE. Despite these
improvements, the eddy-covariance record may miss complex drivers of IAV, or they may be obscured by sensor issues (Jung
et al., 2024). The under-estimation of the global NEE IAV in FLUXCOM X-BASE remains unresolved (Nelson* and Walther*
et al., 2024).

45 Top-down approaches, most commonly atmospheric inversions, are trained using observations of CO_2 mole fraction from
tall-tower observatories and/or satellite retrievals of CO_2. Atmospheric inversions use a Bayesian inversion framework and
an atmospheric transport model to produce global or regional estimates of NEE which are consistent with the observed at-
mospheric signal (Chevallier et al., 2005; Peylin et al., 2013; Crisp et al., 2022; Ciais et al., 2022). These inversion systems
by design provide estimates of NEE that agree with the growth rate of CO_2 at the global scale and across broad latitudinal
50 bands (Peylin et al., 2013) despite differences in priors and representation of transport mechanisms. Atmospheric inversions

are formulated to capture the structure and dynamics of atmospheric CO_2, and can reproduce the observed IAV of NEE with very high accuracy (Rödenbeck et al., 2018).

However, despite improvements in the ability of regional inversions to provide spatially explicit flux estimates (Munassar et al., 2022), global top-down estimates of NEE lack the ability to spatially map the atmospheric signal to local biophysical conditions. Global inversions are more commonly used to understand the land surface at larger integrated scales, and are not directly comparable with eddy-covariance data (Kaminski and Heimann, 2001). Like bottom-up systems, the top-down approach is also limited by a lack of observations. The observational network does not provide sufficient tropical coverage for a robust top-down estimate of the tropical land flux (Palmer et al., 2019). In the southern hemisphere, the weaker gradients in CO_2, and shorter observational records also exacerbate issues of data availability and data quality (Peylin et al., 2013). Additionally, in tropical regions there are less independent data to validate atmospheric inversion estimates (Chevallier et al., 2019).

In previous work it was demonstrated that top-down information can be effectively combined with a bottom-up data-driven flux model, improving the regional and global performance by creating a dual constraint from eddy-covariance data and atmospheric inversion estimates of NEE (Upton et al., 2024). Although this dual-constrained model was able to infer regional and global NEE integrals with a much lower bias compared with other best estimates of NEE than the comparable FLUXCOM RS+METEO V1 results (Jung et al., 2020), the model has several important limitations; First, the atmospheric constraints were pre-calculated using large-scale mean NEE from an ensemble of inversions, rather than from atmospheric CO_2 itself. Second, the atmospheric constraint was connected to the local bottom-up data-driven flux model using a static, statistical model rather than an atmospheric transport model. This means that the additional atmospheric information is aggregated, and adds no additional information on the spatial distribution of NEE. Both issues limited the amount of information available to the model from the atmospheric constraint.

To address these limitations, we create a new model that uses a hybrid combination of direct observations and atmospheric dynamics to integrate ecosystem-level eddy-covariance data and atmospheric CO_2 mole fraction data into a single coherent ecosystem-level flux model. To create the computational link between ecosystem-level fluxes and atmospheric CO_2 mole fraction we use a physical atmospheric transport model, the Stochastic Time-Inverse Lagrangian Transport (STILT) model (Lin et al., 2003). In this way, both eddy-covariance and atmospheric observations are used to train and update the model at the ecosystem level, the result is a bottom-up model that encodes top-down information. This technique opens a large corpus of new observations for constraining the model. Due to varying transport patterns around each tower, the data-driven model is exposed to a large number of non-eddy-covariance locations, enriching the model's view of the underlying relation between driver variables and NEE. These changes, we hypothesize, allow the system to learn additional dynamics in global, and regional NEE.

2 Data

2.1 Ecosystem-level data

At the ecosystem level, we define Net Ecosystem Exchange (NEE) as the simple difference between Gross Primary Production

85 (GPP) and Total Ecosystem Respiration (TER). These terms are measured as the flux density $\mu mol CO_2\ m^{-2}\ s^{-1}$. The sign of NEE indicates the direction of the flux. When NEE is negative, GPP exceeds TER and the flux is a local sink of CO_2. If NEE is positive, TER exceeds GPP and the ecosystem is a local source. NEE represents the net carbon exchange of the vegetation and does not include disturbance fluxes, fires, or out-gassing from freshwater ecosystems.

The core ecosystem-level meteorological and NEE data was collected between 2001 and 2020 at 294 globally distributed

90 eddy-covariance (EC) towers (see appendix for a full list of sites). These data were then processed using the ONEFLUX processing pipeline (Pastorello et al., 2020). Variables from the eddy-covariance sites are air temperature, vapor pressure deficit, incoming shortwave radiation, the computed potential shortwave incoming radiation and the computed wind speed.

During training, at pixels co-located with eddy-covariance towers we use the surface reflectances (MCD43A4, Schaaf and Wang (2015a)) and land surface temperature (LST) from the MOderate Resolution Imaging Spectroradiometer (MODIS)

95 collection v006 (http://daac.ornl.gov/MODIS/). The surface reflectances are used to compute two vegetation indices, the enhanced vegetation index (EVI) (Huete et al., 2002), and the spectral reflectance of vegetation in the near-infrared (NIRv) (Badgley et al., 2019). A water index, the normalized difference water index (NDWI) with MODIS band 7 (Gao, 1996) was also included. MODIS data has a spatial resolution of 0.05 degrees, and a 1-day temporal resolution

When producing our atmospheric constraint, and when estimating global NEE, the drivers for our bottom-up model are the

100 same set of remotely-sensed and eddy-covariance quantities, extracted from remotely-sensed and meteorological reanalysis data. The RS data are derived from the 0.05 degree global MODIS product (MCD43C4, Schaaf and Wang (2015b)), and the global meteorological variables are derived from the European Center for Medium-Range Weather Forecasting (ERA5) atmospheric reanalysis data (https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5). ERA5 is provided at 0.25 degree spatial resolution, and hourly temporal resolution (Hersbach et al., 2020). Air temperature at 2m height, incoming shortwave

105 radiation, and vapor pressure deficit (computed from air temperature, relative humidity, and surface pressure) are used to create the global meteorological data.

All data are accessed through the FLUXCOM-X code base using the preprocessing, and gap-filling from Nelson* and Walther* et al. (2024) and the procedures of FluxnetEO data version 2 (Walther* and Besnard* et al., 2022), and the quality flagging described in Jung et al. (2023). Please see Nelson* and Walther* et al. (2024) for a full description of the FLUXCOM-

110 X data and processing environment.

2.2 Atmospheric data

The model in this study uses observations of atmospheric CO_2 mole fraction from tall-tower observatories measured at three sites: ATTO (Botia et al., 2022), Hegyhatsal (Haszpra, 2024) and Zottino (Tran et al., 2024a). These sites are chosen to include one tropical, one extra-tropical and one boreal domain. The mole fraction data used for the Amazon Tall Tower Observatory

115 in Brazil (ATTO), collected at 79 meters are the same used and fully described in Botia et al. (2022). Mole fraction data for the Hegyhátsál tower in Hungary (HUN), collected at 115 meters are provided by the Integrated Carbon Observatory System (ICOS). The observations for the Zotino Tall Tower Observatory in Siberia (ZOTTO), collected at 301 meters are provided by Tran et al. (2024b). The data for all three towers was provided hourly, for one calendar year (2019 for ZOTTO and Hegyhátsál, 2015 for ATTO). Only daytime observations from 13:00-17:00 (local time) were used. The selected times represent a daytime

120 planetary boundary layer (PBL) with assumed well-mixed convective conditions (Peters et al., 2010).

2.3 Reference data

For global and regional comparison we use the land flux corrected for fossil fuel and cement production from the ensemble of N=14 atmospheric inversions from GCB23 (https://doi.org/10.18160/4M52-VCRU, Luijkx et al. (2024), last accessed 4.3.2024). The atmospheric inversions are provided as monthly data with a 1°spatial resolution. See Friedlingstein et al. (2023)

125 for a full description of the ensemble members. These fluxes are adjusted for fire by removing the fire emissions from Global Fire Emissions Database, Version 4.1 (GFEDv4) (last access 13.2.24) (RANDERSON et al., 2017).

For comparison with the current state-of-the-art bottom-up NEE model, we use the global NEE product from the latest version of FLUXCOM X-BASE (Nelson* and Walther* et al. (2024), last access January 2025). The FLUXCOM X-BASE NEE model (hereafter X-BASE) is trained using the same data processing and data quality flagging, including gap-filling.

130 X-BASE uses a different data-driven algorithm, XGBoost (Chen and Guestrin, 2016), instead of a neural network. Comparison with X-BASE allows us to understand the performance of our system relative to a highly optimized bottom-up model. So, despite differences in structure, we consider X-BASE to be the most appropriate bottom-up model for comparison.

For regional analysis we use the set of 11 large regions from the TransCom 3 intercomparison project (TransCom) project (Baker et al., 2006). See appendix figure A1 for the coverage and names of the 11 land regions used in this study.

135 ## 3 Methods

In this work, we start with a neural network model which uses eddy-covariance, meteorological, and remotely sensed data to estimate NEE at the ecosystem level. We create a hybrid objective function which allows direct constraint by eddy-covariance (EC) observations and hybrid constraint from direct mixing ratio observations of CO_2 using the Stochastic Time-Inverted Lagrangian Transport (STILT) atmospheric transport model as the computational link from the EC-model to the top-down

140 observations. We therefore refer to this new data-driven system with its dual constraints as EC-STILT. Its components and their recombination in a machine-learning framework are described below.

3.1 Ecosystem-level model

The ecosystem-level model takes as input observations of meteorological drivers (from eddy-covariance or reanalysis data) and remotely-sensed drivers and predicts NEE in $\mu mol CO_2\ m^{-2}\ s^{-1}$ (Fig. 1, red lines) at an hourly tempo. As a machine-learning

145 system, the ecosystem-level model can be described as a feed-forward neural network trained using standard backpropagation
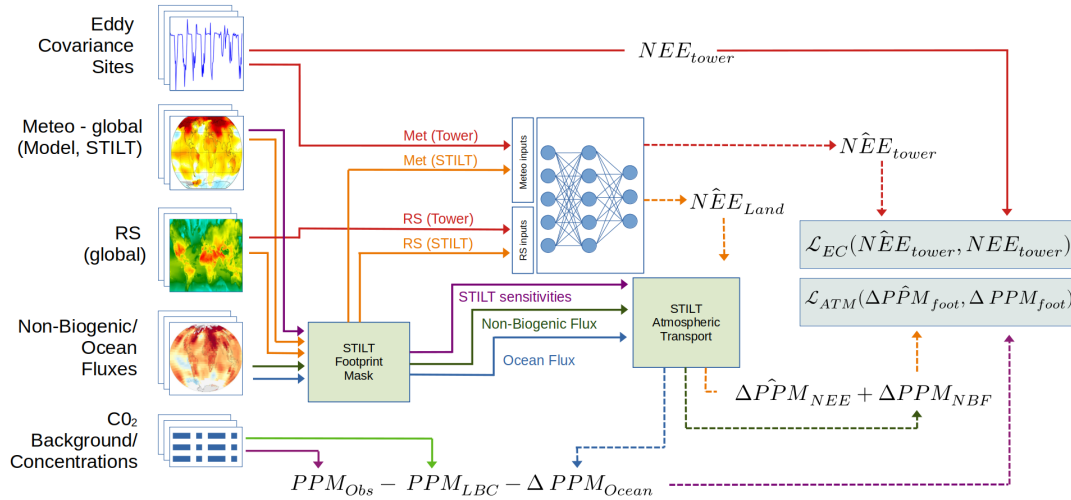
Figure 1. Data flow for atmospheric constraint calculation. An objective term from eddy-covariance (red lines) is created from tower observations of driver variables, co-located remotely sensed data, and tower observations of NEE. An objective term from atmospheric observations (model - orange lines, ancillary data - other colors) uses the STILT model to transport modeled flux densities into concentration space, and corrected for non-local and non-biogenic CO_2: The lateral boundary condition (PPM_{LBC}), or background CO_2 from the well-mixed atmosphere, and any ocean flux contributions (ΔPPM_{Ocean}) are removed from the tower observation (PPM_{Obs}) to create the observed value for the contribution from the footprint (ΔPPM_{foot}). Non-biogenic fluxes (ΔPPM_{NBF}) from fire and fossil fuels are added to the predicted NEE from the footprint ($\Delta P\hat{P}M_{NEE}$) to produce the predicted contribution from the footprint ($\Delta P\hat{P}M_{foot}$). Dashed lines indicate terms that are created during the process of training. See sections 3.3 and 3.4 for a full description of terms.

techniques (Kelley, 1960). The network is a set of fully-connected layers which consist of nodes or "neurons". The neurons are exposed to the output of all neurons in the previous layer. Between each layer of neurons, the output is passed through a non-linear activation function. Additionally, the output of the activation function is passed through a normalization function, which scales the output to a mean of 0 and a standard deviation of 1. Our network is a set of three fully-connected layers with

150   the Rectified Linear Unit (ReLU) activation function (Agarap, 2019) and batch normalization. Each layer in a neural network is a complex, non-linear embedding of the previous inputs into a new latent space. This series of embedding functions is the learning space of the network. These latent spaces allow the model to learn complex non-linear relationships between driver variables. Because these relationships are discovered in these abstract non-geographic spaces, they can represent complex space-for-time and time-for-space substitutions.

155   ## 3.2   STILT Transport

The computational link from this ecosystem-level model to the atmospheric constraint (Fig. 1, purple lines) is built around the STILT atmospheric transport model (Lin et al., 2003). STILT is a Lagrangian particle dispersion model (LPDM) which uses meteorological forcing data to compute the transport of sources and sinks to the time and location of an observation. The

STILT footprint represents this atmospheric transport as a set of sensitivities of the final observation to fluxes at different times
and locations. This footprint can be multiplied with a set of local estimates of NEE, which transports them and converts them
into a simulated mole fraction. EC-STILT uses a set of pre-computed STILT footprints at the three tall tower locations, for
midday observations, for one year (Fig. 1). For every selected atmospheric mole fraction, for each tower, over the year, the
STILT model is run for the specific tower domain, hourly for 10 days prior to the observation time. STILT domains (Fig. 2)
are selected to capture the relevant near field over the time of analysis based on regional meteorological conditions (Lin et al.,
2003).
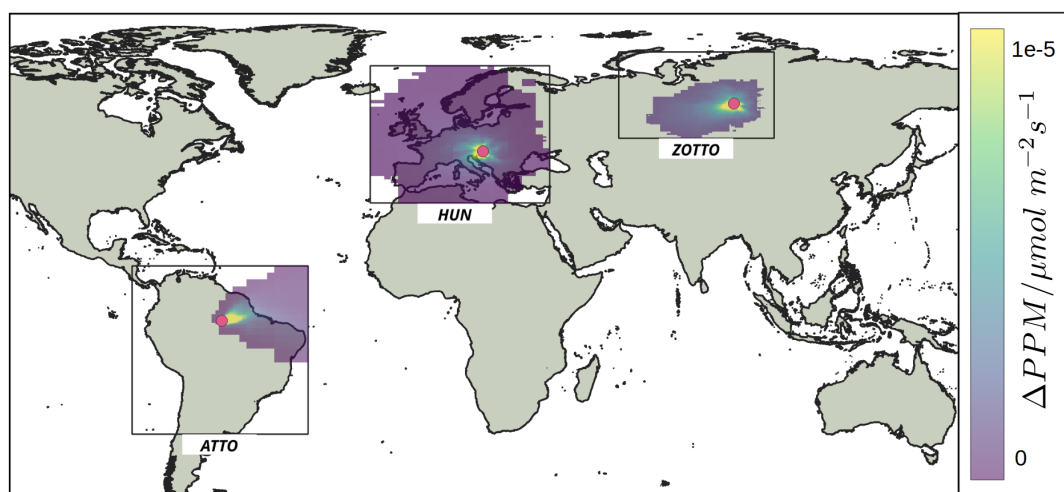
Figure 2. Locations of tall towers used by the atmospheric constraint (red dots). The boxes represent the near-field domain where
STILT is computed, and the grid within each domain is the mean annual sensitivities of the computed STILT footprints. The colorbar,
in $\Delta PPM/\mu mol\ m^{-2}s^{-1}$, represents the mean sensitivity of the tower observation to the region over the full year.

For the atmospheric constraint, we must isolate the biogenic contribution (ΔPPM_{NEE}) of an hourly observation of CO_2
(PPM_{Obs}). Conceptually, the biogenic contribution is the residual of the observation and the non-biogenic terms: the lateral
boundary condition (LBC), non-biogenic fluxes (NBF) and Ocean contributions. This is presented in equation 1, with spatially
explicit terms on the right.

$$PPM_{Obs} - PPM_{LBC} - \Delta PPM_{Ocean} = \Delta PPM_{NEE} + \Delta PPM_{NBF} \tag{1}$$

The NBF contribution (ΔPPM_{NBF}) consists of fossil fuel fluxes and fire emissions. For fossil fuel emissions, we use fossil
fuel/energy signal across sectors, using GridFED (version 2022.2, last access - 15.2.24) (Jones et al., 2021), provided at a 0.1
deg resolution. Fire emissions are provided by the Global Fire Emissions Database, Version 4.1 (GFEDv4) (last access 13.2.24)
(RANDERSON et al., 2017) provided at a 0.25 deg. For the GFEDv4 data, temporal profiles were used to add monthly, daily,
and diurnal cycles to the fire signal. The GridFED data are interpolated to the 0.25 degree spatial resolution of the analysis.

The lateral boundary condition of the region (PPM_{LBC}) is precomputed using The Jena CarboScope (s04ocv4.3) (Rödenbeck et al., 2003), following Botia et al. (2022). Optimized results from the CarbonTracker Data Assimilation Shell (CTDAS) (van der Laan-Luijkx et al., 2017) (CTE2020) are used for the ocean flux and its conversion to mole fractions ((ΔPPM_{Ocean}) with STILT.

3.3 Objective function - Eddy covariance

During training, EC-STILT generates an objective function with two terms, or losses: the ecosystem loss term $\mathcal{L}_{EC}$, and the atmospheric loss term $\mathcal{L}_{ATM}$. To generate the ecosystem loss term $\mathcal{L}_{EC}$ from eddy-covariance, the EC-model $\mathcal{F}()$ is run using a batch of data randomly selected by time and site, x_{batch} collected at eddy-covariance tower sites (Eq. 2, along with remotely sensed (RS) data at pixels that are co-located with the eddy-covariance tower, producing an estimate of NEE at the ecosystem level (eq. 2). The ecosystem loss term is computed as the mean squared error (MSE) between this estimate and the observed NEE from the eddy-covariance observations (Eq. 3).

$$N\hat{E}E = \mathcal{F}(x_{batch}) \tag{2}$$

$$\mathcal{L}_{EC} = MSE(NEE_{obs}, N\hat{E}E) \tag{3}$$

3.4 Objective function - Atmospheric

To generate the atmospheric loss term $\mathcal{L}_{ATM}$ for each training step, an hourly daytime (13:00 to 17:00 local time) observation of the mole fraction of CO_2 for each tower is selected. For each observation, a pre-computed STILT footprint, the LBC and ocean data are selected to create the left-hand side of eq 1. This is the change in CO_2 mole fractions attributable to fluxes within the footprint (ΔPPM_{foot}) that we aim to reproduce with the NEE model. The ecosystem-level model is then run for each non-zero location in the STILT footprint, producing an estimate of local NEE. The NEE inferences and NBF values are transported with the footprint, and the two terms are then added element-wise. The result is summed to create the simulated change in PPM at the tower which is the right-hand side of eq 1. The loss term $\mathcal{L}_{tower,ATM}$ is the squared difference between the simulated change in concentration $\Delta P\hat{P}M_{foot}$, and the observed change in concentration ΔPPM_{foot}. For a single training step, a loss is calculated for one observation for every tower. These three tower losses are averaged, producing the final atmospheric loss term (Eq. 5).

$$\mathcal{L}_{tower,ATM} = (\Delta PPM_{foot} - \Delta P\hat{P}M_{foot})^2 \tag{4}$$

$$\mathcal{L}_{ATM} = \frac{\sum_t^{Num.Towers} \mathcal{L}_{t,ATM}}{Num.Towers} \tag{5}$$

The two terms of the objective function, $\mathcal{L}_{EC}$ and $\mathcal{L}_{ATM}$ are combined using a learned weighting scheme (Kendall et al., 2018). This allows the model to learn the appropriate relative weight to assign each loss term during training, based on their internally estimated uncertainty. The model has a set of additional parameters which are used to learn an estimate of the homoscedastic uncertainty for each term of the objective function. This uncertainty is dependent on the inherent noise in the

training data, rather than the quality or quantity of training data. For EC-STILT these mixing parameters, σ_{EC}^2 and σ_{ATM}^2, are added to the processing chain of the model after its initialization, and then during training, the normal backpropagation process that neural networks use to update their parameters, can also update these mixing parameters. For the individual tasks $L \in [EC, ATM]$, the σ_L^2 parameter is used to create two terms; w_L (Eq. 6), and s_L (Eq. 6). These are then used to calculate

210 the effective loss, balanced by the learned uncertainty of the terms (Eq. 8). After this final weighting calculation, the loss is combined in small batches to stabilize model training during backpropagation.

$$w_{EC} = \frac{1}{2\sigma_{EC}^2}, \quad w_{ATM} = \frac{1}{2\sigma_{ATM}^2} \tag{6}$$

$$s_{EC} = \log\sqrt{\sigma_{EC}^2}, \quad s_{ATM} = \log\sqrt{\sigma_{ATM}^2} \tag{7}$$

$$\mathcal{L}_{total} = (w_{EC} \times \mathcal{L}_{EC}) + (w_{ATM} \times \mathcal{L}_{ATM}) + s_{EC} + s_{ATM} \tag{8}$$

215 ## 3.5 Model Training

EC-STILT is trained using 10-fold cross-validation. The full set of eddy-covariance observations are split by tower location into 10 equal subsets. Ten ensemble members are created, each trained on 8 of the subsets, with one held out for validation and one held out for testing. The ensemble is used to define the flux uncertainty as given in the Results sections, as a 1-σ of the member spread. The available footprints for each tower were randomized and 80% are used for training, 20% for validation on

220 each member. The cross-validation results for the eddy-covariance level results below are the ensemble member results against the 10% of sites held out for testing.

The STILT footprints were run for a 0.25 degree grid. All MODIS driver variables and NBF data were interpolated to this grid. MODIS variables were interpolated bilinearly. NBF fluxes were interpolated using the nearest neighbor method which reduced the overestimation of emissions in pixels surrounding urban hotspots. When run globally, EC-STILT is run at full

225 MODIS 0.05 degree resolution.

4 Results

4.1 Global model performance

EC-STILT produces an estimate of annual global NEE of -3.90 $PgC\ year^{-1} \pm 3.63$, compared to the ensemble mean of atmospheric inversions in GCB23, adjusted for fire, -4.29 $PgC\ year^{-1} \pm 0.65$ (Fig. 3 B). The RMSE of the EC-STILT annual

230 mean with the inversion ensemble annual mean is 0.69 $PgC\ year^{-1}$, compared with 1.81 $PgC\ year^{-1}$ for X-BASE. The uncertainty of the global annual flux is very high (Fig. 3 A), being driven by uncertainties in tropical NEE (Fig. 4C).

EC-STILT has learned a relationship between the input variables and NEE, which captures a large part of the natural inter-annual variability of global annual NEE. When the model is run globally for years 2001-2021, with only local driver variables, and no atmospheric information, the IAV of the EC-STILT member mean, 0.69 PgC year^{-1} (1-σ), is substantially closer to
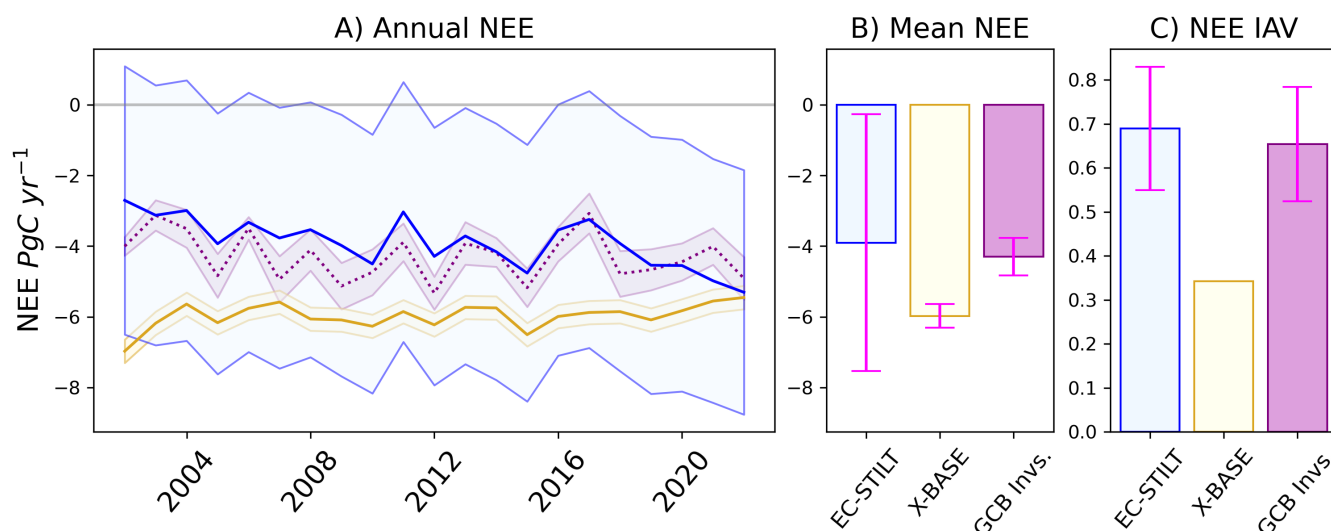
Figure 3. A) Annual NEE integrated globally in $PgC\,year^{-1}$. The blue line is the mean of the EC-STILT ensemble with the shaded area representing the 1-σ range across ensemble members. The yellow line is FLUXCOM X-BASE. The purple dotted line is the mean of the GCB23 ensemble of atmospheric inversions, and the shaded area is the standard deviation across the ensemble members. B) Mean NEE integrated globally over all years. The error bars represent the 1-σ range across ensemble members or published uncertainty. C)IAV of NEE, calculated as the standard deviation of the global integral. The error bars represent the 1-σ range of IAV across ensemble members

the GCB23 inversion ensemble IAV of 0.65 PgC year^{-1} than that of X-BASE (Fig. 3 C). The R^2 of EC-STILT annual mean NEE with the GCB23 inversion ensemble is 0.42 (N=21 years), compared with an R^2 of 0.02 for X-BASE. EC-STILT IAV is consistent across the model members despite the large uncertainty in the magnitude of the flux. The mean R^2 between the 10 members and GCB23 inversion ensemble is 0.34 (1-σ 0.08), and the mean IAV is 0.71 (1-σ 0.14) PgC y^{-1}.

Spatially, the distribution of the mean annual NEE is largely consistent with X-BASE (Figs. 4 B) with a spatial correlation of 0.71 for global mean NEE. EC-STILT has a stronger Amazonian sink and weaker boreal sink (Fig. 4, B). The boreal reduction is a reduction in the length and intensity of the growing season. The EC-STILT results have removed several hotspots of source, in the Sahel and in the Indian subcontinent. These potentially unrealistic hotspots in X-BASE can be attributed to an incorrect learned relationship for crop-cover PFTs in certain dry conditions (Nelson* and Walther* et al., 2024). The lack of specific Plant Functional Type (PFT) information in the EC-STILT training data may be responsible for the removal of these strong sources. The large model spread in the tropics (Fig. 3 C) across the Amazon Basin, the Congo Basin, and Oceania is the major source of global model spread in Fig. 3A.

4.2 Regional model performance

The atmospheric constraint has a strong impact on the EC-STILT estimate of regional fluxes. Similar to the global results, EC-STILT estimates regional magnitudes of NEE which are largely inconsistent (Fig. 5, Tab. 1) when compared with the ensemble
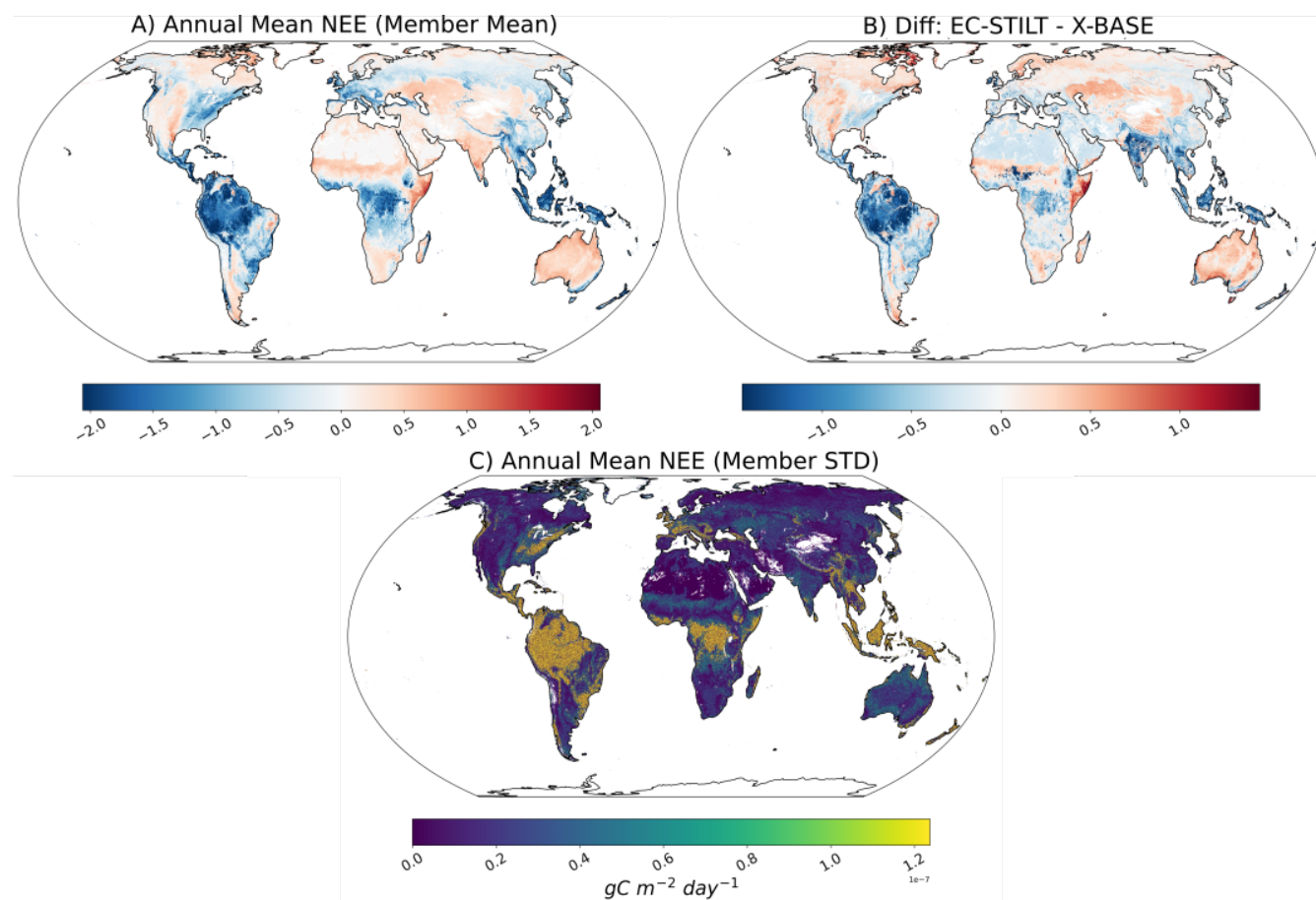
Figure 4. Spatial distribution of global annual NEE A) Mean annual NEE for EC-STILT in $gC\,m^{-2}\,day^{-1}$ B) The difference mean annual NEE between EC-STILT and X-BASE in $gC\,m^{-2}\,day^{-1}$. C) The standard deviation of mean annual NEE for the EC-STILT 10-member ensemble in $gC\,m^{-2}\,day^{-1}$

250 of atmospheric inversions used in Friedlingstein et al. (2023) corrected for fire using GFED4.1 fire emissions (Fig. 5, purple lines). The regional IAV is more in agreement with atmospheric inversions, producing similar or better R^2 and magnitude when compared with X-BASE regional NEE (Fig. 5, inset text, Tab. 1).

Regionally, there appears to be a relationship between the coverage of the eddy-covariance and atmospheric observational networks, and the impact of the atmospheric information (Tab. 1). For example, in Europe (Fig. 5), both the atmospheric and

255 eddy-covariance networks are extensive. It appears that in limited cases, the atmospheric information provides some additional information on the magnitude. EC-STILT is closer to the inversion mean with an RMSE of 0.2 $PgC\,y^{-1}$ compared with 0.37 $PgC\,y^{-1}$ for X-BASE. However, the IAV is very strongly conditioned on the information in eddy-covariance data; the IAV for Europe from EC-STILT is still very similar to X-BASE with a R^2 of 0.86 compared with X-BASE, and 0.31 compared with the inversion mean (Tab. 1). The spatial pattern of summer NEE during the peak growing season (mean JJA) is very consistent
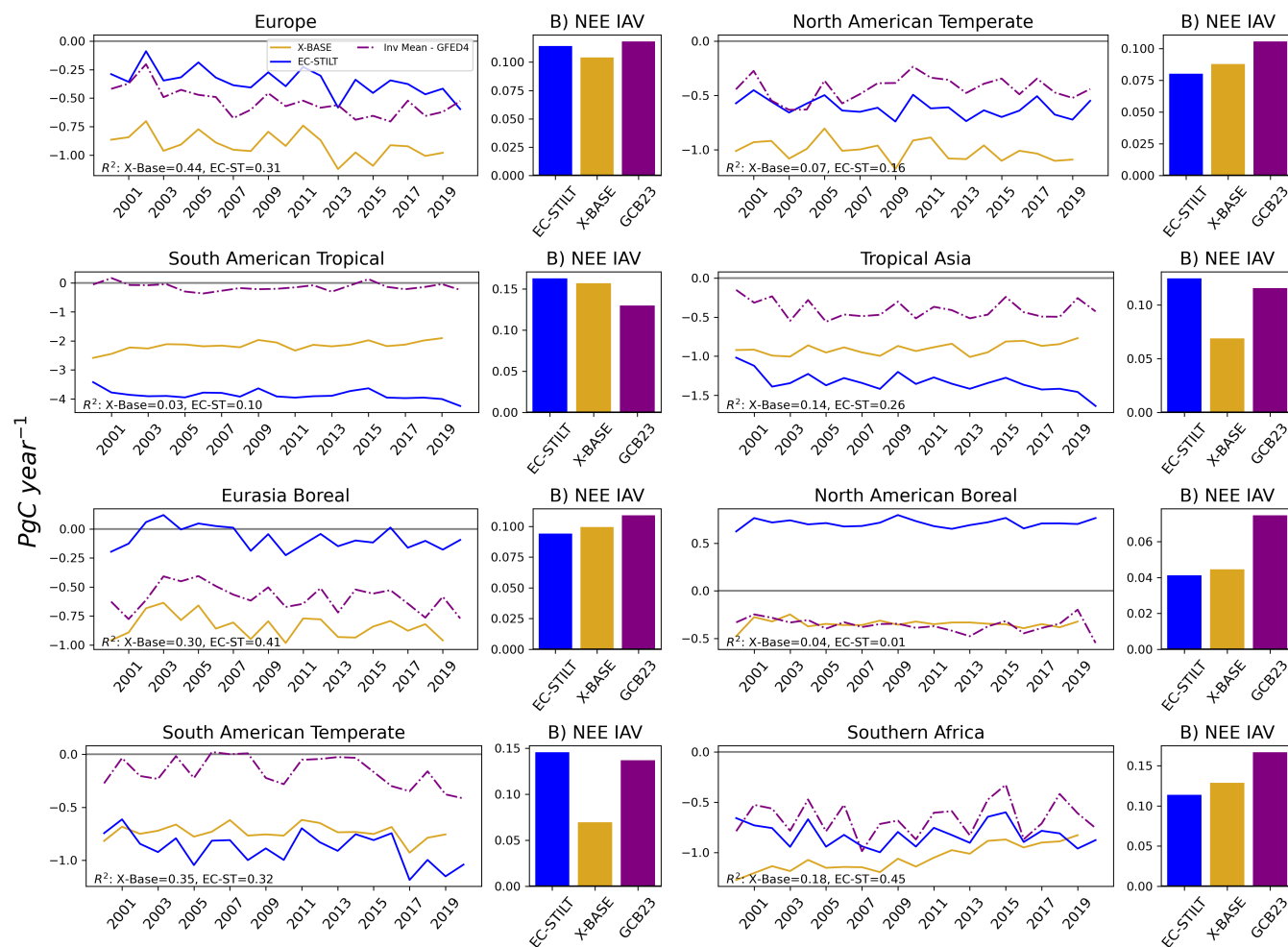
Figure 5. Annual regional NEE and IAV by TransCom region. Column A: Regional fluxes for TransCom regions directly impacted by atmospheric observations: Europe is impacted by observations in the HUN domain. Eurasia Boreal is impacted by observations in the ZOTTO domain. Both South American regions are directly impacted by observations in the ATTO domain. (left): the time series of the annual regional integral flux. (right) the magnitude of IAV for the region. Column B: Regions of similar biome and latitudinal band, by row. The blue is the ensemble mean of EC-STILT, the yellow is FLUXCOM X-BASE, the purple is the ensemble mean of the atmospheric inversions in GCB23 with GFED4.1 fire emissions removed. For each annual time series, the R^2 is reported for EC-STILT and X-BASE with regard to the fire-corrected ensemble mean

between EC-STILT and X-BASE (Fig. 6) with a spatial correlation of 0.86, demonstrating the strong influence eddy-covariance information retains when used in dual-constraint. The reduced sink in the region estimated by EC-STILT (Fig. 5), appears in Figure 6 to be largely an overall bias, rather than a learned difference in the spatial distribution of NEE.

Table 1. Comparison of different estimates with GCB: Column A shows the RMSE between the annual regional integrals in PgC year^{-1} of EC-STILT and X-BASE with the atmospheric inversion ensemble mean with fire removed (Invs). Column B shows the R^2 between the annual regional integrals from EC-STILT and X-BASE with regard to the atmospheric inversion ensemble mean, and between EC-STILT and X-BASE. Column C shows RMSE of the monthly IAV between EC-STILT and X-Base with regard to the atmospheric inversion ensemble mean. Bold indicates the higher relative performance between EC-STILT and X-BASE for each region and metric.

| | A) Annual RMSE | | B) Annual R^2 | | | C) IAV RMSE | |
| | EC-STILT, | X-BASE, | EC-STILT | X-BASE, | EC-STILT, | EC-STILT, | X-BASE, |
	Invs	Invs	Invs	Invs	X-BASE	Invs-f	Invs-f
North American Boreal	1.076	**0.094**	**0.009**	0.008	0.387	0.012	0.012
North American Temperate	**0.205**	0.579	**0.158**	0.067	0.554	**0.006**	0.008
South American Tropical	3.725	**2.030**	**0.098**	0.039	0.176	0.011	0.011
South American Temperate	0.735	**0.585**	**0.317**	0.310	0.424	**0.007**	0.010
Northern Africa	1.125	**0.659**	0.248	**0.354**	0.050	**0.014**	0.015
Southern Africa	**0.202**	0.419	**0.450**	0.094	0.010	**0.012**	0.014
Eurasia Boreal	0.520	**0.267**	**0.413**	0.272	0.750	**0.011**	0.012
Eurasia Temperate	1.893	**1.555**	0.266	**0.292**	0.876	**0.006**	0.007
Tropical Asia	0.941	**0.511**	**0.263**	0.124	0.007	0.009	0.009
Australia	**1.374**	0.418	0.083	**0.292**	0.009	0.005	0.005
Europe	**0.210**	0.397	0.306	**0.375**	0.873	0.008	**0.007**

In regions such as South American Tropical where there is limited eddy-covariance coverage to adequately constrain the model, additional atmospheric information does not improve the modeled NEE, and increases the model uncertainty. EC-STILT has moved away from the inversion mean, compared with X-BASE (RMSE of 3.73 $PgC\ y^{-1}$ compared with 2.05 $PgC\ y^{-1}$). The low R^2 for EC-STILT compared with both the inversion mean and X-BASE annual time series (0.10 compared with the inversions, 0.13 compared with X-BASE) indicate that neither the eddy-covariance data nor the atmospheric signal provide sufficient constraint for EC-STILT to find similar IAV to the atmospheric inversion mean, which itself might be wrong.

The increase in the IAV for the South American Temperate region demonstrates the potential for complex effects from the atmospheric constraint. The region is only directly impacted by the ATTO tower observations in unusual meteorological conditions, which is evident in the mean ATTO footprint in Figure 2. The eddy-covariance record for the region is sparse, but may have similarities to other, more densely sampled regions. We see a small move away from the inversion mean compared with FLUXCOM X-BASE in the magnitude of NEE (RMSE of 0.74 $PgC\ y^{-1}$ compared with 0.59 $PgC\ y^{-1}$ for X-BASE), but an increase in the magnitude of IAV compared with FLUXCOM X-BASE (0.14 $PgC\ y^{-1}$ compared with 0.07 $PgC\ y^{-1}$). The atmospheric constraint has added new information about this region despite very limited direct information available in training.

The atmospheric constraint also has strong effects in TransCom regions that do not directly contain or overlap the STILT regions used during training. This is because EC-STILT learns its land-surface response in environmental space of the features instead of in geographic space like an inversion, Across the globe, EC-STILT produces consistent changes over regions with
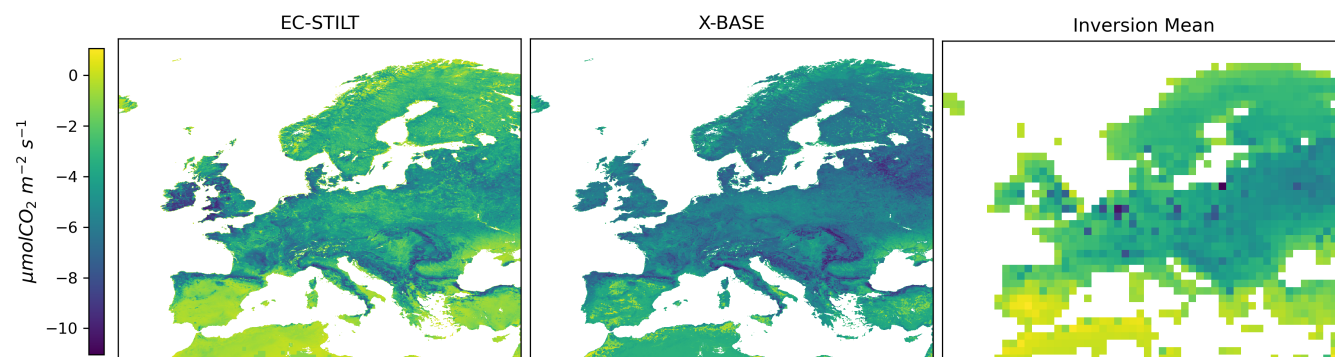
Figure 6. Mean summer NEE (JJA) in $\mu mol CO_2\ m^{-2}\ s^{-1}$ for EC-STILT, X-BASE, and the ensemble mean of the atmospheric inversions.

280 similar latitude ranges (Fig. 5, right column) with a dominance of the global north in observational networks (Chu et al., 2017). The magnitude of the northern hemispheric land carbon sink is part of ongoing discrepancies in the Global Carbon Budget (O'Sullivan et al., 2024).

4.3 IAV attribution

While the global IAV of EC-STILT is closer to IAV of the GCB land residual (Fig 3), the regional impact of atmospheric

285 constraint is more heterogeneous. The relationship between EC-STILT and the inversion mean appears to depend on the biome, and the strength of the eddy-covariance constraint. In certain regions EC-STILT has improved the R^2 (Southern Africa), or the magnitude of IAV (South American Temperate) relative to the fire-corrected ensemble mean of the atmospheric inversions, when compared to X-BASE (Fig. 5). In regions which are well constrained by eddy-covariance (Europe, North American Temperate), both the magnitude and annual R^2 relative the inversion mean are very similar to X-BASE. In these regions, the

290 annual R^2 relative to X-BASE is higher than with the inversion ensemble (Tab. 1).

EC-STILT mean monthly IAV, shows very little change compared with X-BASE (Tab. 1, IAV RMSE). EC-STILT has modified its response by biome; When IAV is broken down by month (Fig. 7) across boreal and temperate regions, the increases in monthly IAV occurs during the growing season while in tropical regions show a change during the dry season. Appendix Figure B1 shows the monthly spatial distribution of IAV. The monthly IAV also shows a clear dependence on the growing

295 season, with IAV distributed from the timing of the onset and end of season, and the magnitude of the carbon uptake. Tropical regions generally have lower IAV (Fig. 7), and EC-STILT locates the more of the IAV in the dry season.

4.4 Eddy-covariance site-level evaluation

EC-STILT maintains good overall performance for NEE at the eddy-covariance site level, particularly at daily time scale. We use R^2 and the Nash Sutcliffe model efficiency (NSE) (Nash and Sutcliffe, 1970) to evaluate model performance with regard
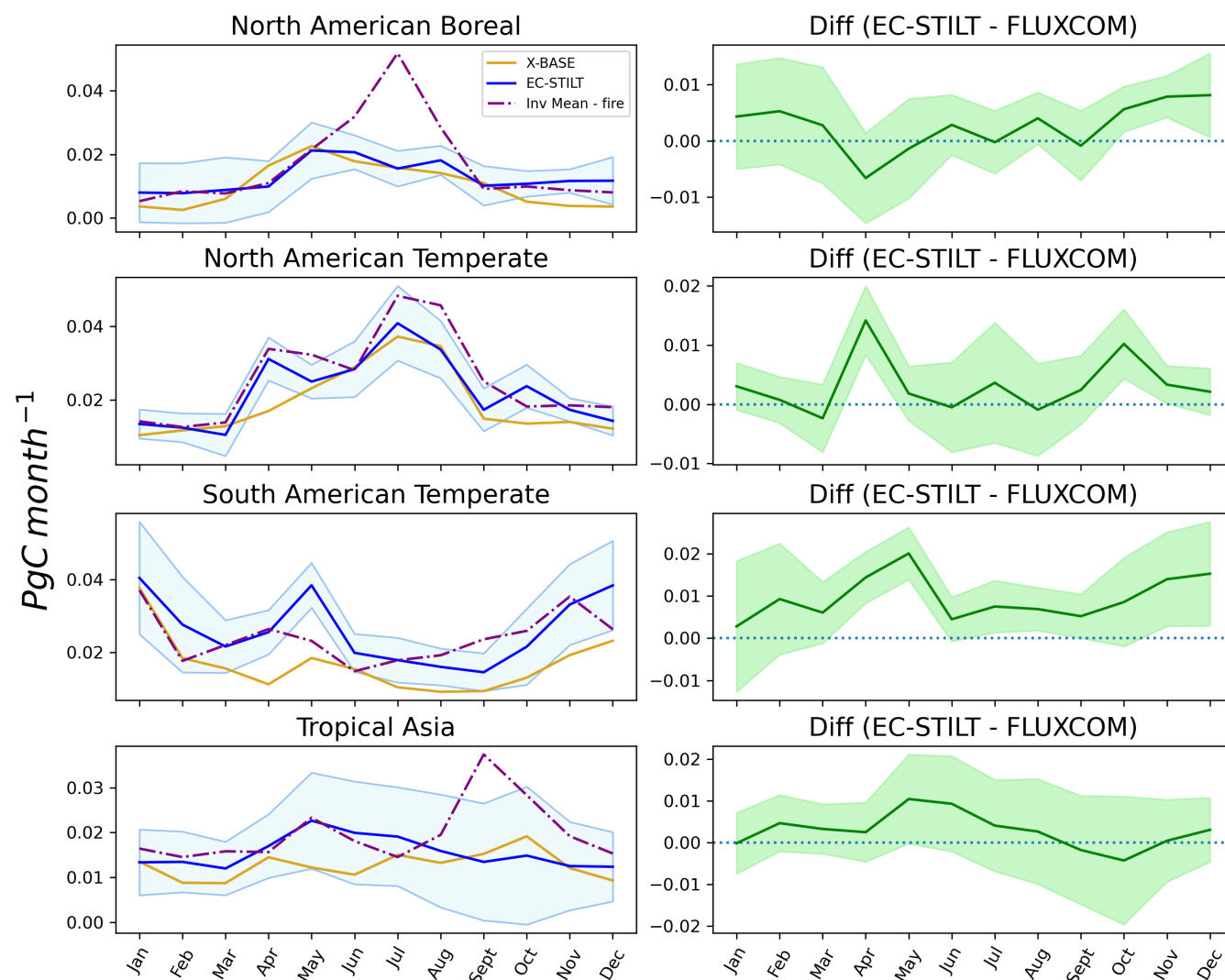
14

Figure 7. Attribution of IAV across the mean seasonal cycle: The left column is the monthly mean seasonality of IAV for the region, calculated as the standard deviation of the month across the period of analysis (2001-2021). The blue line is the member mean of EC-STILT, the yellow line is the X-BASE NEE, and the purple dotted line is the ensemble mean of the atmospheric inversions in GCB23. The right column is the difference (EC-STILT - X-BASE) across the year.

to the observed eddy-covariance observations, and for comparison with X-BASE. In the EC-STILT's 10-fold cross-validation for all sites, the model achieves a median NSE of 0.66, and median R^2 of 0.77 for hourly fluxes (X-BASE: NSE 0.71, R^2 0.71), and a median NSE of 0.76 and median R^2 of 0.87 for diurnal mean NEE (X-BASE: NSE 0.81, R^2 0.81), when comparing the observed and simulated time series. The data used for these performance metrics are the test data, held out from each cross-validation fold, and these tests are performed using the model member which did not see these data in training.

305   As with regional results, the site-level performance of EC-STILT appears to be strongly influenced by the level of local constraint from eddy-covariance. In figure 8 the hourly and daily averaged time series of NEE for three EC sites directly impacted by the atmospheric domains, one in boreal Russia (RU-Ha1), one in a German mixed forest (DE-Hai), and one Brazilian site in the Amazon basin (BR-Sa1) are shown. In the Russian site, EC-STILT (blue line) matches the temporal pattern of the EC observations (red line) at the hourly level (R^2 of 0.57). When averaged to the daily scale, both EC-STILT and

310   X-BASE produce a similar RMSE compared with the observations. EC-STILT appears to better represent the daily signal than X-BASE with a higher R^2 (0.43 compared with 0.18). EC-STILT produces a range in NEE which is slightly larger than the EC record and much greater than FLUXCOM X-BASE (yellow line). Boreal regions are only moderately constrained in EC-STILT by eddy-covariance given the large difference in annual NEE between EC-STILT and X-BASE (Fig. 5). This may allow for larger differences at the site level, as the boreal observations do not sufficiently weigh against the atmospheric information.

315   In the German site, DE-Hai, where the constraint from eddy-covariance is strongest, EC-STILT underestimates the hourly magnitude (Fig. 8, center left) with a mean RMSE of 0.92 for EC-STILT, 0.22 for X-BASE), but still largely captures the daily NEE, but with a larger bias than X-BASE (Fig 8, center right) with an R^2 of 0.61 and RMSE of 2.00 compared with an R^2 of 0.83 and RMSE of 1.04 for X-BASE. In the northern extra-tropics, as seen above, the marginal impact of the atmospheric information is minimized, and EC-STILT performs most similarly with X-BASE. EC-STILT has a stronger positive bias than

320   X-BASE (EC-STILT 1.75, X-BASE 0.83). This bias in the site level results is consistent with the slightly weaker regional sink seen in the northern temperate regions.

In the Brazilian site, at the hourly scale EC-STILT strongly underestimates the range of NEE (hourly RMSE of 6.32), but infers a daily value when averaged, that is comparable with X-BASE (R^2 of 0.47 for EC-STILT and X-BASE). This may indicate that the eddy-covariance and atmospheric constraints are providing confounding information about the hourly flux,

325   which is discussed below.

5   Discussion

The core logic of our model is a synergistic sharing of information between the bottom-up and top-down constraints. The key innovation in our system is the combination of the atmospheric term with the ecosystem-level objective term derived from eddy-covariance data, and their joint projection onto a space that is not the "traditional" geographic distribution of NEE.

330   Instead, EC-STILT learns in "environmental response" space, or a set of relations between the driver datasets (features) and the output NEE. The STILT Atmospheric transport model provides the spatiotemporal locations where ecosystem NEE variations contribute to a downwind CO_2 observation. This large number of locations over the near-field can be modeled and their output constrained by the atmospheric observation. Considered independently this is underconstrained; one atmospheric observation is insufficient to constrain the multiple modeled NEE estimates under the STILT footprint as also experienced in many inverse

335   modeling problems using atmospheric data (see Peylin et al. (2013); Gaubert et al. (2023) for an overview). However, because the 'environmental response' space of the model is conditioned by the ecosystem-level objective term, the response across the STILT footprint is tightly bound to the local, better-constrained eddy-covariance record. Learning is thus jointly informed by
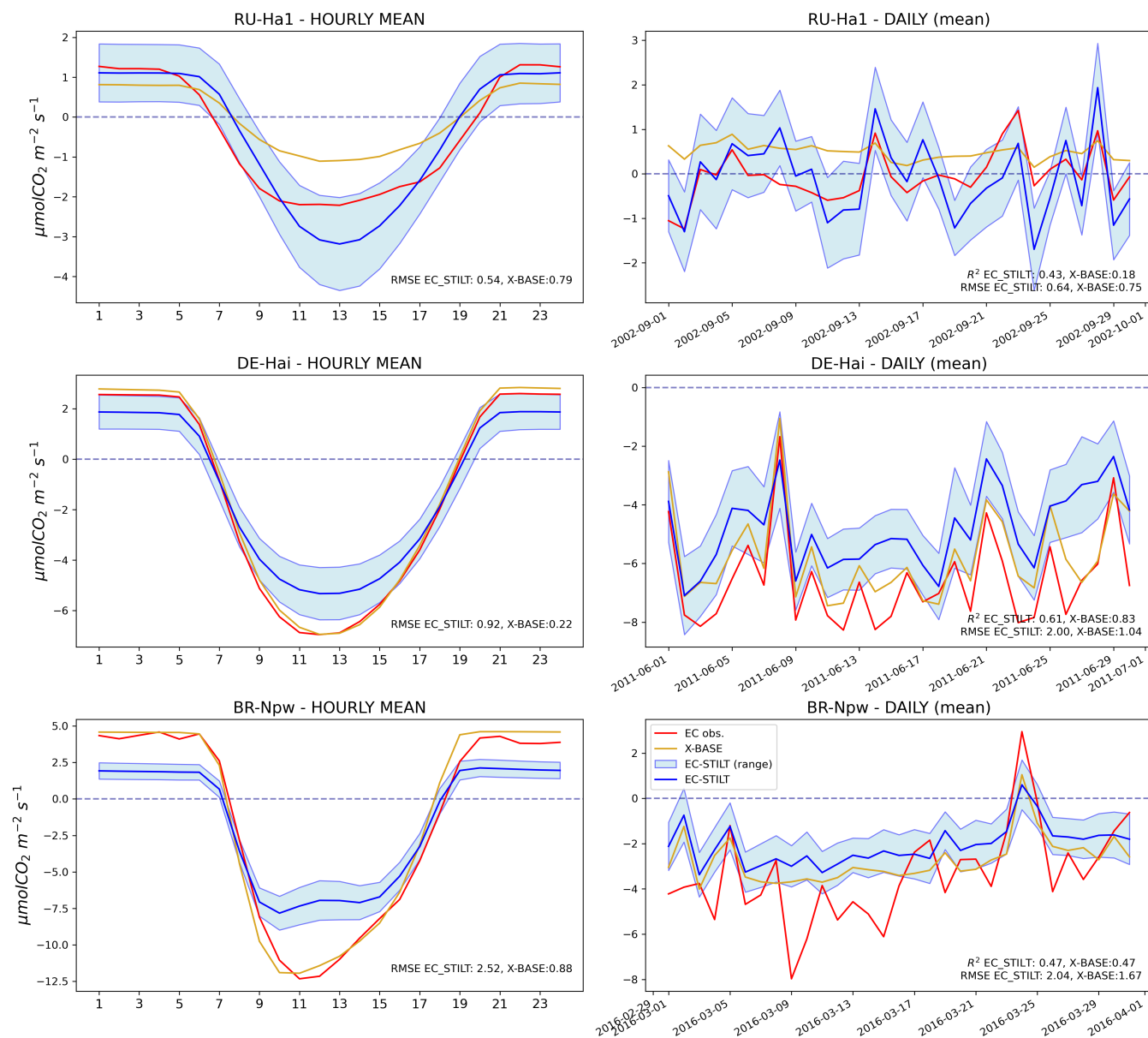
Figure 8. EC tower results across latitude bands: Each row is a tower. The left column is the mean hourly NEE, the right column is the daily mean across an example month. The blue line is the member mean of EC-STILT, the yellow line is the X-BASE NEE, and the red dotted line is the eddy-covariance observations for the tower. All values are in $\mu mol C\ m^{-2}\ s^{-1}$.

both data sources, while maintaining, outside of the tropics, the local daily performance of bottom-up models like FLUXCOM V1 (Jung et al., 2020) and X-BASE (Nelson* and Walther* et al., 2024).

340   EC-STILT efficiently learns driver/NEE relationships globally which adds new information on the regional and global dynamics of NEE. This is demonstrated by the improved representation of IAV of global NEE (Fig. 3) and regional NEE (Fig. 5). There is no explicit representation in the training data of the atmospheric IAV; The atmospheric constraint is only present for three regional domains (around the ZOTTO, ATTO, and HUN towers). These domains were chosen to represent a tropical (ATTO), extra-tropical (HUN), and boreal (ZOTTO) eco-region. Each tower has observations across a single year (2019 for

345   ZOTTO and Hegyhátsál, 2015 for ATTO) chosen for reasons of data availability. These relatively small samples of the available atmospheric information are sufficient to embed the atmospheric signal in the ecosystem response of EC-STILT, and modify the spatial and temporal distributions of estimated NEE. There is no explicit representation in the data of the atmospheric growth rate, nor mass-balance yet these are reproduced to an impressive degree. The atmospheric observations are regional and instant, with no direct information on the long-term state of the atmosphere. During training, the LBC signal does pro-

350   vide some indirect information on the atmosphere, as it is removed from the observation. Because the atmospheric transport system is in fact not even run during the calculation of global NEE. There is therefore no atmospheric information available to EC-STILT during the creation of our global NEE data. The potentially realistic global integral and spatial distribution (Fig. 3) demonstrates that the model is able to learn a more consistent ecosystem-level response globally from this sparse set of local observations. Of particular note is that these limited observations improve the estimate of IAV across the entire period of

355   analysis, and thus learned relations hold under a range of environmental conditions in each pixel.

       Although subject to large uncertainty, the magnitude of the annual flux has moved closer to the GCB23 atmospheric inversion estimate of global NEE. The global magnitude is strongly conditioned, however on the inclusion and handling of the non-biospheric flux data during training; As the total amount of NBF in a region increases, by necessity, the inferred NEE must decrease to keep the local instantaneous equation balanced (Eq. 1). The total contributions to the well-mixed atmosphere over

360   time are complex (Ciais et al., 2022), and often not available in spatially explicit data products. The current study used fire and fossil fuel emissions, but other terms such as riverine fluxes (see Botía et al. (2024) for a discussion at the ATTO tower) may play strong local or global roles in the regional CO_2 balance, and should be included in the tower loss calculation.

       In figure 3A, it appears that the atmospheric signal increases the uncertainty around the magnitude and distribution of tropical NEE, when compared with X-BASE. The ensemble members show a large spread across tropical regions, and at the

365   ecosystem-level, the ensemble mean has difficulty reproducing the magnitude of hourly observations from the eddy-covariance record (Fig. 8). EC-STILT's tropical shortcomings may be a combination of several structural factors; model type, potential mismatches in the atmospheric operator, and limited observational support, which are discussed below.

       When compared with X-BASE, the core model structure is important; X-BASE is a XGBoost regressor, which have been demonstrated to be robust in this domain compared with neural networks which have more capacity to overfit (Nelson* and

370   Walther* et al., 2024; Kraft et al., 2024; Jung et al., 2020). Because our objective function is sensitive to ecosystem and regional results across all three towers, it is possible that within the overall learning process, some local or regional overfitting is occurring. We attempt to reduce overfitting using early stopping of training according to the validation accuracy, and in the model using batch normalization. Potentially, this overfitting could lead to large intra-model uncertainty between neural network members compared with an XGBoost ensemble. Current implementations of XGBoost in the Python programming

additionally constrained by three atmospheric tall-towers over one year. As noted above, neither the eddy-covariance record nor the atmospheric observational record have strongly representative global coverage, and additional atmospheric information over new geographic regions could be a benefit. Therefore an important future development of a dual-constrained model is the inclusion of additional tower and non-tower data streams: Additional tall towers could provide stronger constraint in well-measured regions. Non-tower data could be leveraged to provide information on regions that are poorly covered by the current observational network. The current model could run, as-is, against aircraft observations. With a modification of the STILT transport system in the objective function, the structure of this study would allow a bottom-up model to be constrained by satellite retrievals of total column CO_2 (XCO_2). XCO_2 observations could be selected according to their potential to include biomes of specific interest in their upwind footprints. This would enrich our capacity to model NEE across un- and under-sampled biomes. This extension into XCO_2 introduces several possible issues: to model the full column of CO_2, multiple STILT footprints must be run for every observation at different elevations (Wu et al., 2018), and modeled fluxes transported multiple times. This may be a technical challenge to build effective training systems. The inclusion of the full column CO_2 may dilute the available information on the biogenic flux in the observation, reducing the information gain to the bottom-up model.

The performance of EC-STILT in the tropics show the limitations of adding atmospheric information without adequate support in the eddy-covariance record. In the harmonized model described above, the ecosystem-level information appears to provide the 'backbone' of the model's performance. Future modeling efforts will benefit from the ongoing efforts to improve and enlarge the corpus of available eddy-covariance sites and data, such as a potential extension of the FLUXNET 2015 dataset (Pastorello et al., 2020). As the eddy-covariance record better describes and constrains key biomes, additional atmospheric constraint may be able to provide improved information on the larger dynamics of regional and global NEE, while reducing the uncertainty.

An important aspect of transport modeling which is not addressed in the current study is the impact of model and meteorological uncertainty. In our analysis, we treat the output of the STILT model as definitive, relative to the meteorological fields in ECMWF short-term forecast. Computational schemes for addressing model uncertainty, multi-LPDM ensembling, or in-model ensembling for STILT by sampling within the stochastic particles could be attempted, without a major increase in computational and I/O workload. In this study, STILT runs are forced with ECMWF short-term forecast data (Botia et al., 2022), and the model is trained using ERA5 reanalysis data. Using identical meteorological data for STILT and the model drivers may also reduce uncertainty.

Conceptually, the current work accounts for two major terms in the non-biogenic flux budget; fires and fossil fuel emissions. In the Amazon and in boreal Russia, fire is the largest and the most relevant term, as the term accounts for both fire, and also for part of the instantaneous flux from land-cover/land-use change associated with biomass burning (Cochrane and Laurance, 2008). In the European domain, fossil fuels are the dominant non-biogenic flux signal. As noted in the section 2, NEE at the eddy-covariance tower and atmospheric CO_2 are not directly comparable. The dual-constraint in this study is created using the residual of an observation and non-biogenic fluxes. Any meaningful non-biogenic flux term which is not included, or adequately represented temporally and spatially will alias into the atmospheric target that the model is trying to match in

training. This reduces the effectiveness of the technique. Any additional spatially explicit NBF data may therefore, improve
445 this comparison, and increase the fidelity of the overall magnitude and may reduce the uncertainty in the global result.

Beyond an improved global NEE product, a key scientific goal is a detailed understanding of the different learned relationship
in EC-STILT between NEE and the driver variables, and how this mapping is transformed by the inclusion of atmospheric
information. There are many open questions remaining on how EC-STILT locates the IAV signal, and what new spatiotemporal
dependencies are learned outside of the eddy-covariance record. Methods for explainable machine-learning may allow for an
450 analysis of the model's learned 'environmental space'. Relationships between this internal representation and environmental
variables may provide insight into how and where EC-STILT locates the IAV.

6 Conclusions

This study demonstrates a novel method for combining bottom-up and top-down information in the objective function of a
data-driven flux model. We wanted to test the effectiveness of a physics-based operator that transports modeled flux density
455 into atmospheric concentration. And we wanted to determine whether a temporally and geographically-sparse constraint could
produce valid global results. The EC-STILT model in this study is still strongly conditioned by the bottom-up constraint from
eddy-covariance, and so maintains the strengths of the overall bottom-up approach, but now with access to information on land
surface far beyond the eddy-covariance network.

A classical top-down atmospheric inversion matches the observed atmospheric signal but lacks the framework to accurately
460 distribute that signal down into geographic/biophysical space. Bottom-up models can translate EC measurements into their
latent space, but do not have any additional information about how inferred fluxes influence the global signal. In contrast, our
hybrid system in this study can take information from the atmospheric observations and accurately distribute it into the learned
latent space of the model. Because this latent space still contains information from the eddy-covariance record, this allows for
a valid mapping down from the atmosphere to the model, and from the model to an estimate of NEE at multiple scales which
465 is comparable both with both atmospheric and ecosystem-level observations. Using our system, a dual-constrained model sits
directly between bottom-up and top-down approaches, directly addressing their limitations while maintaining their advantages.

Appendix A: Data

Table A1: Citation data for the 294 sites used in training.

AR-SLu (Garcia et al., 2016)	AR-TF1 (Kutzbach, 2021)	AR-Vir (Posse et al., 2016)	AT-Neu (Wohlfahrt et al., 2016)	AU-ASM (Cleverly and Eamus, 2016b)	AU-Ade (Beringer and Hutley, 2016c)

Table A1: Citation data for the 294 sites used in training.

AU-Cpr (Meyer et al., 2016)	AU-Cum (Pendall and Griebel, 2016)	AU-DaP (Beringer and Hutley, 2016b)	AU-DaS (Beringer and Hutley, 2016f)	AU-Dry (Beringer and Hutley, 2016e)	AU-Emr (Schroder et al., 2016)
AU-Fog (Beringer and Hutley, 2016a)	AU-Gin (Macfarlane et al., 2016)	AU-RDF (Beringer and Hutley, 2016d)	AU-Rob (Liddell, 2016)	AU-TTE (Cleverly and Eamus, 2016a)	AU-Tum (Woodgate et al., 2016)
AU-Wac (Beringer et al., 2016b)	AU-Whr (Beringer et al., 2016a)	AU-Wom (Arndt et al., 2016)	AU-Ync (Beringer and Walker, 2016)	BE-Bra (Team and Centre, 2022-02-01)	BE-Dor (Team and Centre, 2022-02-01)
BE-Lcr (RI, 2021-05-06)	BE-Lon (Team and Centre, 2022-02-01)	BE-Maa (Team and Centre, 2022-02-01)	BE-Vie (Team and Centre, 2022-02-01)	BR-Npw (Vourlitis et al., 2022)	BR-Sa1 (Saleska, 2016)
BR-Sa3 (Goulden, 2016d)	CA-Cbo (Staebler, 2022)	CA-DB2 (Knox, 2022)	CA-DBB (Christen and Knox, 2022)	CA-ER1 (Wagner-Riddle, 2021)	CA-Gro (McCaughey, 2016)
CA-LP1 (Black, 2021)	CA-Man (Amiro, 2016b)	CA-NS2 (Goulden, 2016a)	CA-NS3 (Goulden, 2016b)	CA-NS4 (Goulden, 2016c)	CA-NS5 (Goulden, 2016g)
CA-NS6 (Goulden, 2016e)	CA-NS7 (Goulden, 2016f)	CA-Oas (Black, 2016b)	CA-Obs (Black, 2016a)	CA-Qfo (Margolis, 2016)	CA-SF1 (Amiro, 2016c)
CA-SF2 (Amiro, 2016a)	CA-SF3 (Amiro, 2016d)	CA-TP1 (Arain, 2016b)	CA-TP2 (Arain, 2016a)	CA-TP3 (Arain, 2022b)	CA-TP4 (Arain, 2016c)
CA-TPD (Arain, 2022a)	CG-Tch (Nouvellon, 2016)	CH-Aws (Team and Centre, 2022-02-01)	CH-Cha (Team and Centre, 2022-02-01)	CH-Dav (Team and Centre, 2022-02-01)	CH-Fru (Team and Centre, 2022-02-01)
CH-Lae (Team and Centre, 2022-02-01)	CH-Oe1 (Ammann, 2016)	CH-Oe2 (Team and Centre, 2022-02-01)	CN-Cha (Zhang and Han, 2016)	CN-Cng (Dong, 2016)	CN-Dan (Shi et al., 2016)
CN-Din (Zhou and Yan, 2016)	CN-Du2 (Chen, 2016k)	CN-Du3 (Shao, 2016b)	CN-HaM (Tang et al., 2016)	CN-Qia (Wang and Fu, 2016)	CN-Sw2 (Shao, 2016a)

Table A1: Citation data for the 294 sites used in training.

CZ-BK1 (Team and Centre, 2022-02-01)	CZ-BK2 (Sigut et al., 2016)	CZ-KrP (Team and Centre, 2022-02-01)	CZ-Lnz (Team and Centre, 2022-02-01)	CZ-RAJ (Team and Centre, 2022-02-01)	CZ-Stn (Team and Centre, 2022-02-01)
CZ-wet (Team and Centre, 2022-02-01)	DE-Akm (Team and Centre, 2022-02-01)	DE-Geb (Team and Centre, 2022-02-01)	DE-Gri (Team and Centre, 2022-02-01)	DE-Hai (Team and Centre, 2022-02-01)	DE-HoH (Team and Centre, 2022-02-01)
DE-Hte (Team and Centre, 2020-03-10)	DE-Hzd (Team and Centre, 2022-02-01)	DE-Kli (Team and Centre, 2022-02-01)	DE-Lkb (Lindauer et al., 2016)	DE-Lnf (Knohl et al., 2016)	DE-Obe (Team and Centre, 2022-02-01)
DE-RuR (RI, 2022-06-14)	DE-RuS (Team and Centre, 2022-02-01)	DE-RuW (Team and Centre, 2022-02-01)	DE-Seh (Schneider and Schmidt, 2016)	DE-SfN (Klatt et al., 2016)	DE-Spw (Bernhofer et al., 2016)
DE-Tha (Team and Centre, 2022-02-01)	DE-Zrk (Sachs et al., 2016)	DK-Eng (Pilegaard and Ibrom, 2016)	DK-Fou (Olesen, 2016)	DK-Gds (RI, 2022-06-14)	DK-Sor (Team and Centre, 2022-02-01)
ES-Abr (Team and Centre, 2022-02-01)	ES-Agu (Team and Centre, 2022-02-01)	ES-Amo (Poveda et al., 2016)	ES-Cnd (Team and Centre, 2022-02-01)	ES-LJu (Team and Centre, 2022-02-01)	ES-LM1 (Team and Centre, 2022-02-01)
ES-LM2 (Team and Centre, 2022-02-01)	ES-LgS (Reverter et al., 2016b)	ES-Ln2 (Reverter et al., 2016a)	FI-Hyy (Team and Centre, 2022-02-01)	FI-Jok (Lohila et al., 2016)	FI-Ken (Team and Centre, 2022-02-01)
FI-Let (Team and Centre, 2022-02-01)	FI-Lom (Aurela et al., 2016a)	FI-Qvd (Team and Centre, 2022-02-01)	FI-Sii (Team and Centre, 2022-02-01)	FI-Sod (Aurela et al., 2016b)	FI-Var (RI, 2022-06-14)
FR-Aur (Team and Centre, 2022-02-01)	FR-Bil (Team and Centre, 2022-02-01)	FR-EM2 (RI, 2022-06-14)	FR-FBn (Team and Centre, 2022-02-01)	FR-Fon (Team and Centre, 2022-02-01)	FR-Gri (Team and Centre, 2022-02-01)
FR-Hes (Team and Centre, 2022-02-01)	FR-LBr (Berbigier and Loustau, 2016)	FR-LGt (RI, 2022-06-14)	FR-Lam (Team and Centre, 2022-02-01)	FR-Pue (Ourcival, 2016)	FR-Tou (RI, 2022-06-14)
GF-Guy (Team and Centre, 2022-02-01)	GH-Ank (Valentini et al., 2016b)	GL-Dsk (RI, 2022-06-14)	GL-NuF (Hansen, 2016)	GL-ZaF (Lund et al., 2016b)	GL-ZaH (Lund et al., 2016a)

Table A1: Citation data for the 294 sites used in training.

IE-Cra (Team and Centre, 2022-02-01)	IL-Yat (Team and Centre, 2022-02-01)	IT-BCi (Team and Centre, 2022-02-01)	IT-BFt (RI, 2022-06-14)	IT-CA1 (Sabbatini et al., 2016c)	IT-CA2 (Sabbatini et al., 2016a)
IT-CA3 (Sabbatini et al., 2016b)	IT-Col (Matteucci, 2016)	IT-Cp2 (Team and Centre, 2022-02-01)	IT-Cpz (Valentini et al., 2016a)	IT-Isp (Gruening et al., 2016b)	IT-La2 (Cescatti et al., 2016)
IT-Lav (Team and Centre, 2022-02-01)	IT-Lsn (RI, 2022-06-14)	IT-MBo (Team and Centre, 2022-02-01)	IT-Noe (Spano et al., 2016)	IT-PT1 (Manca and Goded, 2016)	IT-Ren (Team and Centre, 2022-02-01)
IT-Ro1 (Valentini et al., 2016c)	IT-Ro2 (Papale et al., 2016)	IT-SR2 (Team and Centre, 2022-02-01)	IT-SRo (Gruening et al., 2016a)	IT-Tor (Team and Centre, 2022-02-01)	JP-MBF (Kotani, 2016b)
JP-SMF (Kotani, 2016a)	MX-Tes (Yepez and Garatuza, 2021)	MY-PSO (Kosugi and Takanashi, 2016)	NL-Hor (Dolman et al., 2016a)	NL-Loo (Team and Centre, 2020-03-10)	PA-SPn (Wolf et al., 2016b)
PA-SPs (Wolf et al., 2016a)	PE-QFR (Griffis and Roman, 2021)	RU-Che (Merbold et al., 2016)	RU-Cok (Dolman et al., 2016b)	RU-Fy2 (Team and Centre, 2022-02-01)	RU-Fyo (Team and Centre, 2022-02-01)
RU-Ha1 (Belelli et al., 2016)	SD-Dem (Ardö et al., 2016)	SE-Deg (Team and Centre, 2022-02-01)	SE-Htm (Team and Centre, 2022-02-01)	SE-Lnn (Team and Centre, 2020-03-10)	SE-Nor (Team and Centre, 2022-02-01)
SE-Ros (Team and Centre, 2022-02-01)	SE-Svb (Team and Centre, 2022-02-01)	SJ-Adv (Christensen, 2016)	SJ-Blv (Boike et al., 2016)	SN-Dhr (Tagesson et al., 2016)	US-A32 (Billesbach et al., 2022)
US-AR1 (Billesbach et al., 2016b)	US-AR2 (Billesbach et al., 2016a)	US-ARM (Biraud et al., 2022)	US-ARb (Torn, 2016b)	US-ARc (Torn, 2016a)	US-Atq (Zona and Oechel, 2016a)
US-BZB (Euskirchen, 2022b)	US-BZF (Euskirchen, 2022c)	US-BZS (Euskirchen, 2022d)	US-BZo (Euskirchen, 2022a)	US-Bi1 (Rey-Sanchez et al., 2022b)	US-Bi2 (Rey-Sanchez et al., 2022a)

Table A1: Citation data for the 294 sites used in training.

US-Blo (Goldstein, 2016)	US-CF1 (Huggins, 2021)	US-CF2 (Huggins, 2022c)	US-CF3 (Huggins, 2022a)	US-CF4 (Huggins, 2022b)	US-CRT (Chen and Chu, 2016b)
US-CS1 (Desai, 2022a)	US-CS2 (Desai, 2022c)	US-CS3 (Desai, 2022d)	US-CS4 (Desai, 2022b)	US-Cop (Bowling, 2016)	US-EDN (Oikawa, 2021)
US-GBT (Massman, 2016)	US-GLE (Massman, 2022)	US-Goo (Meyers, 2016b)	US-HB1 (Forsythe et al., 2021)	US-HWB (Goslee, 2022)	US-Ha1 (Munger, 2016)
US-Hn3 (Liu et al., 2022)	US-Ho2 (Hollinger, 2022)	US-IB2 (Matamala, 2016)	US-ICs (Euskirchen et al., 2022a)	US-ICt (Euskirchen et al., 2022b)	US-Ivo (Zona and Oechel, 2016b)
US-Jo2 (Vivoni and Perez-Ruiz, 2022)	US-KFS (Brunsell, 2022a)	US-KLS (Brunsell, 2022b)	US-KS1 (Drake and Hinkle, 2016a)	US-KS2 (Drake and Hinkle, 2016b)	US-KS3 (Hinkle, 2022)
US-LWW (Meyers, 2016a)	US-Lin (Fares, 2016)	US-Los (Desai, 2016c)	US-MMS (Novick and Phillips, 2022)	US-MOz (Wood and Gu, 2022)	US-Me1 (Law, 2016c)
US-Me2 (Law, 2022)	US-Me3 (Law, 2016a)	US-Me4 (Law, 2016e)	US-Me5 (Law, 2016d)	US-Me6 (Law, 2016b)	US-Mpj (Litvak, 2021)
US-Myb (Sturtevant et al., 2016)	US-NGB (Torn and Dengel, 2021)	US-NR1 (Blanken et al., 2022)	US-Ne1 (Suyker, 2022)	US-Ne2 (Suyker, 2016b)	US-Ne3 (Suyker, 2016a)
US-ONA (Silveira, 2021)	US-ORv (Bohrer, 2021)	US-OWC (Bohrer and Kerns, 2022)	US-Oho (Chen et al., 2016)	US-PFa (Desai, 2016d)	US-Prr (Kobayashi and Suzuki, 2016)
US-Rms (Flerchinger, 2022c)	US-Ro1 (Baker et al., 2022)	US-Ro4 (Baker and Griffis, 2022a)	US-Ro5 (Baker and Griffis, 2021)	US-Ro6 (Baker and Griffis, 2022b)	US-Rwe (Flerchinger and Reba, 2022)
US-Rwf (Flerchinger, 2022a)	US-Rws (Flerchinger, 2022b)	US-SRC (Kurc, 2022)	US-SRG (Scott, 2016a)	US-SRM (Scott, 2016b)	US-Sne (Shortt et al., 2022)

Table A1: Citation data for the 294 sites used in training.

|					
US-Snf (Kusak et al., 2022)	US-Sta (Ewers and Pendall, 2016)	US-Syv (Desai, 2016b)	US-Ton (Baldocchi and Ma, 2016)	US-Tw1 (Valach et al., 2021)	US-Tw2 (Sturtevant et al., 2022)
US-Tw3 (Chamberlain et al., 2022)	US-Tw4 (Sanchez et al., 2016)	US-Tw5 (Valach et al., 2022)	US-Twt (Baldocchi, 2016)	US-UM3 (Bohrer, 2022)	US-UMB (Gough et al., 2016)
US-UMd (Gough et al., 2022)	US-Var (Baldocchi et al., 2016)	US-WCr (Desai, 2016a)	US-WPT (Chen and Chu, 2016a)	US-Whs (Scott, 2016d)	US-Wi0 (Chen, 2016g)
US-Wi1 (Chen, 2016e)	US-Wi2 (Chen, 2016j)	US-Wi3 (Chen, 2016b)	US-Wi4 (Chen, 2016d)	US-Wi5 (Chen, 2016a)	US-Wi6 (Chen, 2016h)
US-Wi7 (Chen, 2016i)	US-Wi8 (Chen, 2016c)	US-Wi9 (Chen, 2016f)	US-Wjs (Litvak, 2022)	US-Wkg (Scott, 2016c)	US-xBR (Network), 2022)
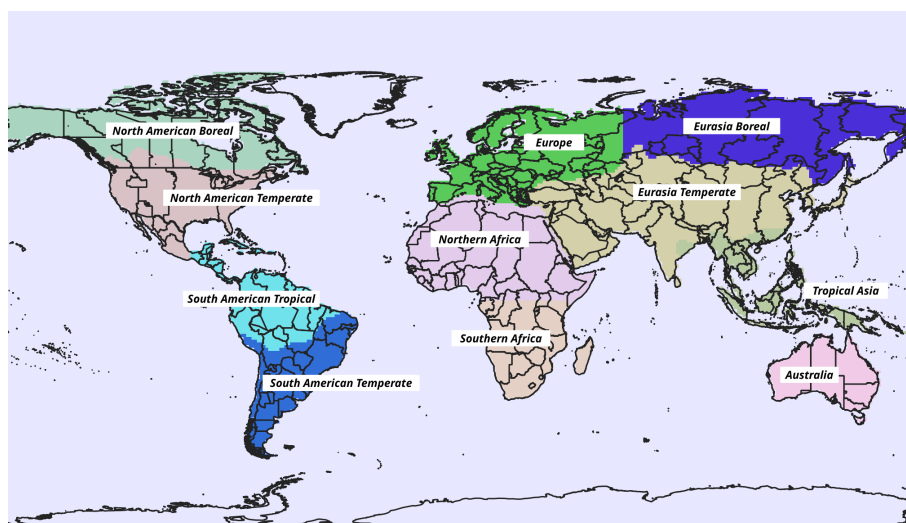
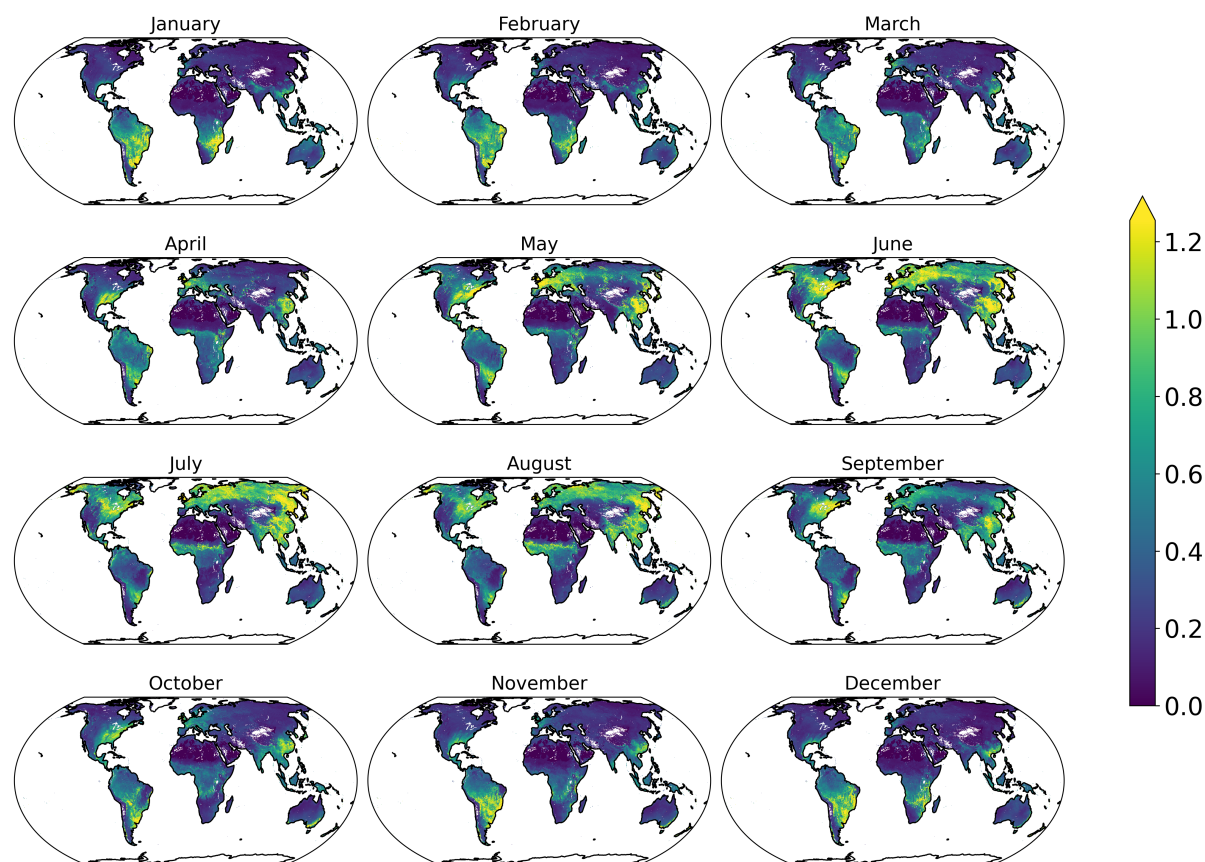**Figure A1.** TransCom land regions

Appendix B: Results

Figure B1. EC-STILT monthly mean IAV of NEE, calculated as the monthly mean of the day-of-year standard deviation

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

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.

Amiro, B.: FLUXNET2015 CA-SF2 Saskatchewan - Western Boreal, Forest Burned in 1989, https://doi.org/10.18140/flx/1440047, accessed: 2022-11-01, 2016a.

Amiro, B.: FLUXNET2015 CA-Man Manitoba - Northern Old Black Spruce (Former BOREAS Northern Study Area), https://doi.org/10.18140/flx/1440035, accessed: 2022-11-01, 2016b.

Amiro, B.: FLUXNET2015 CA-SF1 Saskatchewan - Western Boreal, Forest Burned in 1977, https://doi.org/10.18140/flx/1440046, accessed: 2022-11-01, 2016c.

Amiro, B.: FLUXNET2015 CA-SF3 Saskatchewan - Western Boreal, Forest Burned in 1998, https://doi.org/10.18140/flx/1440048, accessed: 2022-11-01, 2016d.

Ammann, C.: FLUXNET2015 CH-Oe1 Oensingen Grassland, https://doi.org/10.18140/flx/1440135, accessed: 2022-11-01, 2016.

Angevine, W. M., Brioude, J., McKeen, S., and Holloway, J. S.: Uncertainty in Lagrangian pollutant transport simulations due to meteorological uncertainty from a mesoscale WRF ensemble, Geoscientific Model Development, 7, 2817–2829, https://doi.org/10.5194/gmd-7-2817-2014, publisher: Copernicus GmbH, 2014.

Arain, M.: AmeriFlux FLUXNET-1F CA-TPD Ontario - Turkey Point Mature Deciduous, https://doi.org/10.17190/amf/1881567, accessed: 2022-11-01, 2022a.

Arain, M.: AmeriFlux FLUXNET-1F CA-TP3 Ontario - Turkey Point 1974 Plantation White Pine, https://doi.org/10.17190/amf/1881566, accessed: 2022-11-01, 2022b.

Arain, M. A.: FLUXNET2015 CA-TP2 Ontario - Turkey Point 1989 Plantation White Pine, https://doi.org/10.18140/flx/1440051, accessed: 2022-11-01, 2016a.

Arain, M. A.: FLUXNET2015 CA-TP1 Ontario - Turkey Point 2002 Plantation White Pine, https://doi.org/10.18140/flx/1440050, accessed: 2022-11-01, 2016b.

Arain, M. A.: FLUXNET2015 CA-TP4 Ontario - Turkey Point 1939 Plantation White Pine, https://doi.org/10.18140/flx/1440053, accessed: 2022-11-01, 2016c.

Ardö, J., El Tahir, B. A., and ElKhidir, H. A. M.: FLUXNET2015 SD-Dem Demokeya, https://doi.org/10.18140/flx/1440186, accessed: 2022-11-01, 2016.

Arndt, S., Hinko-Najera, N., and Griebel, A.: FLUXNET2015 AU-Wom Wombat, https://doi.org/10.18140/flx/1440207, accessed: 2022-11-01, 2016.

Aubinet, M., Berbigier, P., Bernhofer, C., Cescatti, A., Feigenwinter, C., Granier, A., Grünwald, T., Havrankova, K., Heinesch, B., Longdoz, B., Marcolla, B., Montagnani, L., and Sedlak, P.: Comparing CO2 Storage and Advection Conditions at Night at Different Carboeuroflux Sites, Boundary-Layer Meteorology, 116, 63–93, https://doi.org/10.1007/s10546-004-7091-8, 2005.

Aurela, M., Lohila, A., Tuovinen, J.-P., Hatakka, J., Rainne, J., Mäkelä, T., and Lauria, T.: FLUXNET2015 FI-Lom Lompolojankka, https://doi.org/10.18140/flx/1440228, accessed: 2022-11-01, 2016a.

Aurela, M., Tuovinen, J.-P., Hatakka, J., Lohila, A., Mäkelä, T., Rainne, J., and Lauria, T.: FLUXNET2015 FI-Sod Sodankyla, https://doi.org/10.18140/flx/1440160, accessed: 2022-11-01, 2016b.

Badgley, G., Anderegg, L. D. L., Berry, J. A., and Field, C. B.: Terrestrial gross primary production: Using NIRV
520   to scale from site to globe, Global Change Biology, 25, 3731–3740, https://doi.org/10.1111/gcb.14729, _eprint:
https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.14729, 2019.

Baker, D. F., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P., Denning, A. S., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fung, I. Y.,
Heimann, M., John, J., Maki, T., Maksyutov, S., Masarie, K., Prather, M., Pak, B., Taguchi, S., and Zhu, Z.: TransCom 3 inversion inter-
comparison: Impact of transport model errors on the interannual variability of regional CO2 fluxes, 1988–2003, Global Biogeochemical
525   Cycles, 20, https://doi.org/10.1029/2004GB002439, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2004GB002439,
2006.

Baker, J. and Griffis, T.: AmeriFlux FLUXNET-1F US-Ro5 Rosemount I18South, https://doi.org/10.17190/amf/1818371, accessed: 2022-
11-01, 2021.

Baker, J. and Griffis, T.: AmeriFlux FLUXNET-1F US-Ro4 Rosemount Prairie, https://doi.org/10.17190/amf/1881589, accessed: 2022-11-
530   01, 2022a.

Baker, J. and Griffis, T.: AmeriFlux FLUXNET-1F US-Ro6 Rosemount I18North, https://doi.org/10.17190/amf/1881590, accessed: 2022-
11-01, 2022b.

Baker, J., Griffis, T., and Griffis, T.: AmeriFlux FLUXNET-1F US-Ro1 Rosemount- G21, https://doi.org/10.17190/amf/1881588, accessed:
2022-11-01, 2022.

535   Baldocchi, D.: FLUXNET2015 US-Twt Twitchell Island, https://doi.org/10.18140/flx/1440106, accessed: 2022-11-01, 2016.

Baldocchi, D. and Ma, S.: FLUXNET2015 US-Ton Tonzi Ranch, https://doi.org/10.18140/flx/1440092, accessed: 2022-11-01, 2016.

Baldocchi, D., Ma, S., and Xu, L.: FLUXNET2015 US-Var Vaira Ranch- Ione, https://doi.org/10.18140/flx/1440094, accessed: 2022-11-01,
2016.

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,
540   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.

Belelli, L., Papale, D., and Valentini, R.: FLUXNET2015 RU-Ha1 Hakasia Steppe, https://doi.org/10.18140/flx/1440184, accessed: 2022-
11-01, 2016.

545   Berbigier, P. and Loustau, D.: FLUXNET2015 FR-LBr Le Bray, https://doi.org/10.18140/flx/1440163, accessed: 2022-11-01, 2016.

Beringer, J. and Hutley, L.: FLUXNET2015 AU-Fog Fogg Dam, https://doi.org/10.18140/flx/1440124, accessed: 2022-11-01, 2016a.

Beringer, J. and Hutley, L.: FLUXNET2015 AU-DaP Daly River Savanna, https://doi.org/10.18140/flx/1440123, accessed: 2022-11-01,
2016b.

Beringer, J. and Hutley, L.: FLUXNET2015 AU-Ade Adelaide River, https://doi.org/10.18140/flx/1440193, accessed: 2022-11-01, 2016c.

550   Beringer, J. and Hutley, L.: FLUXNET2015 AU-RDF Red Dirt Melon Farm, Northern Territory, https://doi.org/10.18140/flx/1440201, ac-
cessed: 2022-11-01, 2016d.

Beringer, J. and Hutley, L.: FLUXNET2015 AU-Dry Dry River, https://doi.org/10.18140/flx/1440197, accessed: 2022-11-01, 2016e.

Beringer, J. and Hutley, P. L.: FLUXNET2015 AU-DaS Daly River Cleared, https://doi.org/10.18140/flx/1440122, accessed: 2022-11-01,
2016f.

555   Beringer, J. and Walker, J.: FLUXNET2015 AU-Ync Jaxa, https://doi.org/10.18140/flx/1440208, accessed: 2022-11-01, 2016.

Beringer, J., Cunningham, S., Baker, P., Cavagnaro, T., MacNally, R., Thompson, R., and McHugh, I.: FLUXNET2015 AU-Whr Whroo, https://doi.org/10.18140/flx/1440206, accessed: 2022-11-01, 2016a.

Beringer, J., Hutley, L., McGuire, D., and U, P.: FLUXNET2015 AU-Wac Wallaby Creek, https://doi.org/10.18140/flx/1440127, accessed: 2022-11-01, 2016b.

560 Bernhofer, C., Grünwald, T., Moderow, U., Hehn, M., Eichelmann, U., and Prasse, H.: FLUXNET2015 DE-Spw Spreewald, https://doi.org/10.18140/flx/1440220, accessed: 2022-11-01, 2016.

Billesbach, D., Bradford, J., and Torn, M.: FLUXNET2015 US-AR2 ARM USDA UNL OSU Woodward Switchgrass 2, https://doi.org/10.18140/flx/1440104, accessed: 2022-11-01, 2016a.

Billesbach, D., Bradford, J., and Torn, M.: FLUXNET2015 US-AR1 ARM USDA UNL OSU Woodward Switchgrass 1, 565 https://doi.org/10.18140/flx/1440103, accessed: 2022-11-01, 2016b.

Billesbach, D., Kueppers, L., Torn, M., and Biraud, S.: AmeriFlux FLUXNET-1F US-A32 ARM-SGP Medford Hay Pasture, https://doi.org/10.17190/amf/1881568, accessed: 2022-11-01, 2022.

Biraud, S., Fischer, M., Chan, S., and Torn, M.: AmeriFlux FLUXNET-1F US-ARM ARM Southern Great Plains Site- Lamont, https://doi.org/10.17190/amf/1854366, accessed: 2022-11-01, 2022.

570 Black, T.: AmeriFlux FLUXNET-1F CA-LP1 British Columbia - Mountain Pine Beetle-Attacked Lodgepole Pine Stand, https://doi.org/10.17190/amf/1832155, accessed: 2022-11-01, 2021.

Black, T. A.: FLUXNET2015 CA-Obs Saskatchewan - Western Boreal, Mature Black Spruce, https://doi.org/10.18140/flx/1440044, accessed: 2022-11-01, 2016a.

Black, T. A.: FLUXNET2015 CA-Oas Saskatchewan - Western Boreal, Mature Aspen, https://doi.org/10.18140/flx/1440043, accessed: 2022-575 11-01, 2016b.

Blanken, P., Monson, R., Burns, S., Bowling, D., and Turnipseed, A.: AmeriFlux FLUXNET-1F US-NR1 Niwot Ridge Forest (LTER NWT1), https://doi.org/10.17190/amf/1871141, accessed: 2022-11-01, 2022.

Bodesheim, P., Jung, M., Gans, F., Mahecha, M. D., and Reichstein, M.: Upscaled diurnal cycles of land–atmosphere fluxes: a new global half-hourly data product, Earth System Science Data, 10, 1327–1365, https://doi.org/10.5194/essd-10-1327-2018, publisher: Copernicus 580 GmbH, 2018.

Bohrer, G.: AmeriFlux FLUXNET-1F US-ORv Olentangy River Wetland Research Park, https://doi.org/10.17190/amf/1832164, accessed: 2022-11-01, 2021.

Bohrer, G.: AmeriFlux FLUXNET-1F US-UM3 Douglas Lake, https://doi.org/10.17190/amf/1881596, accessed: 2022-11-01, 2022.

Bohrer, G. and Kerns, J.: AmeriFlux FLUXNET-1F US-OWC Old Woman Creek, https://doi.org/10.17190/amf/1871142, accessed: 2022-585 11-01, 2022.

Boike, J., Westermann, S., Lüers, J., Langer, M., and Piel, K.: FLUXNET2015 SJ-Blv Bayelva, Spitsbergen, https://doi.org/10.18140/flx/1440242, accessed: 2022-11-01, 2016.

Botia, S., Komiya, S., Marshall, J., Koch, T., Galkowski, M., Lavric, J., Gomes-Alves, E., Walter, D., Fisch, G., Pinho, D. M., Nelson, B. W., Martins, G., Luijkx, I. T., Koren, G., Florentie, L., de Araujo, A. C., Sa, M., Andreae, M. O., Heimann, M., Peters, W., and Gerbig, C.: 590 The CO2 record at the Amazon Tall Tower Observatory: A new opportunity to study processes on seasonal and inter-annual scales, Global Change Biology, 28, 588–611, https://doi.org/10.1111/gcb.15905, place: Hoboken Publisher: Wiley WOS:000710981900001, 2022.

Botía, S., Munassar, S., Koch, T., Custodio, D., Basso, L. S., Komiya, S., Lavric, J. V., Walter, D., Gloor, M., Martins, G., Naus, S., Koren, G., Luijkx, I., Hantson, S., Miller, J. B., Peters, W., Rödenbeck, C., and Gerbig, C.: Combined CO_2 measurement record indicates decreased

Amazon forest carbon uptake, offset by Savannah carbon release, EGUsphere, pp. 1–55, https://doi.org/10.5194/egusphere-2024-1735,

595   publisher: Copernicus GmbH, 2024.

Bowling, D.: FLUXNET2015 US-Cop Corral Pocket, https://doi.org/10.18140/flx/1440100, accessed: 2022-11-01, 2016.

Brunsell, N.: AmeriFlux FLUXNET-1F US-KFS Kansas Field Station, https://doi.org/10.17190/amf/1881585, accessed: 2022-11-01, 2022a.

Brunsell, N.: AmeriFlux FLUXNET-1F US-KLS Kansas Land Institute, https://doi.org/10.17190/amf/1854367, accessed: 2022-11-01, 2022b.

600   Cescatti, A., Marcolla, B., Zorer, R., and Gianelle, D.: FLUXNET2015 IT-La2 Lavarone2, https://doi.org/10.18140/flx/1440235, accessed: 2022-11-01, 2016.

Chamberlain, S., Oikawa, P., Sturtevant, C., Szutu, D., Verfaillie, J., and Baldocchi, D.: AmeriFlux FLUXNET-1F US-Tw3 Twitchell Alfalfa, https://doi.org/10.17190/amf/1881594, accessed: 2022-11-01, 2022.

Chen, J.: FLUXNET2015 US-Wi5 Mixed Young Jack Pine (MYJP), https://doi.org/10.18140/flx/1440059, accessed: 2022-11-01, 2016a.

605   Chen, J.: FLUXNET2015 US-Wi3 Mature Hardwood (MHW), https://doi.org/10.18140/flx/1440057, accessed: 2022-11-01, 2016b.

Chen, J.: FLUXNET2015 US-Wi8 Young Hardwood Clearcut (YHW), https://doi.org/10.18140/flx/1440062, accessed: 2022-11-01, 2016c.

Chen, J.: FLUXNET2015 US-Wi4 Mature Red Pine (MRP), https://doi.org/10.18140/flx/1440058, accessed: 2022-11-01, 2016d.

Chen, J.: FLUXNET2015 US-Wi1 Intermediate Hardwood (IHW), https://doi.org/10.18140/flx/1440054, accessed: 2022-11-01, 2016e.

Chen, J.: FLUXNET2015 US-Wi9 Young Jack Pine (YJP), https://doi.org/10.18140/flx/1440063, accessed: 2022-11-01, 2016f.

610   Chen, J.: FLUXNET2015 US-Wi0 Young Red Pine (YRP), https://doi.org/10.18140/flx/1440055, accessed: 2022-11-01, 2016g.

Chen, J.: FLUXNET2015 US-Wi6 Pine Barrens 1 (PB1), https://doi.org/10.18140/flx/1440060, accessed: 2022-11-01, 2016h.

Chen, J.: FLUXNET2015 US-Wi7 Red Pine Clearcut (RPCC), https://doi.org/10.18140/flx/1440061, accessed: 2022-11-01, 2016i.

Chen, J.: FLUXNET2015 US-Wi2 Intermediate Red Pine (IRP), https://doi.org/10.18140/flx/1440056, accessed: 2022-11-01, 2016j.

Chen, J. and Chu, H.: FLUXNET2015 US-WPT Winous Point North Marsh, https://doi.org/10.18140/flx/1440116, accessed: 2022-11-01,

615   2016a.

Chen, J. and Chu, H.: FLUXNET2015 US-CRT Curtice Walter-Berger Cropland, https://doi.org/10.18140/flx/1440117, accessed: 2022-11-01, 2016b.

Chen, J., Chu, H., and Noormets, A.: FLUXNET2015 US-Oho Oak Openings, https://doi.org/10.18140/flx/1440088, accessed: 2022-11-01, 2016.

620   Chen, S.: FLUXNET2015 CN-Du2 Duolungrassland (D01), https://doi.org/10.18140/flx/1440140, accessed: 2022-11-01, 2016k.

Chen, T. and Guestrin, C.: XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16, pp. 785–794, Association for Computing Machinery, New York, NY, USA, ISBN 978-1-4503-4232-2, https://doi.org/10.1145/2939672.2939785, 2016.

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F.-M., Chédin, A., and Ciais, P.: Inferring CO2 sources and

625   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.

Chevallier, F., Remaud, M., O'Dell, C. W., Baker, D., Peylin, P., and Cozic, A.: Objective evaluation of surface- and satellite-driven carbon dioxide atmospheric inversions, Atmospheric Chemistry and Physics, 19, 14 233–14 251, https://doi.org/10.5194/acp-19-14233-2019, publisher: Copernicus GmbH, 2019.

630   Christen, A. and Knox, S.: AmeriFlux FLUXNET-1F CA-DBB Delta Burns Bog, https://doi.org/10.17190/amf/1881565, accessed: 2022-11-01, 2022.

Christensen, T.: FLUXNET2015 SJ-Adv Adventdalen, https://doi.org/10.18140/flx/1440241, accessed: 2022-11-01, 2016.

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:
635   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,
640   15, 1289–1316, https://doi.org/10.5194/gmd-15-1289-2022, publisher: Copernicus GmbH, 2022.

Cleverly, J. and Eamus, D.: FLUXNET2015 AU-TTE Ti Tree East, https://doi.org/10.18140/flx/1440205, accessed: 2022-11-01, 2016a.

Cleverly, J. and Eamus, D.: FLUXNET2015 AU-ASM Alice Springs, https://doi.org/10.18140/flx/1440194, accessed: 2022-11-01, 2016b.

Cochrane, M. A. and Laurance, W. F.: Synergisms among Fire, Land Use, and Climate Change in the Amazon, Ambio, 37, 522–527, https://www.jstor.org/stable/25547943, publisher: [Springer, Royal Swedish Academy of Sciences], 2008.

645   Crisp, D., Dolman, H., Tanhua, T., McKinley, G. A., Hauck, J., Bastos, A., Sitch, S., Eggleston, S., and Aich, V.: How Well Do We Understand the Land-Ocean-Atmosphere Carbon Cycle?, Reviews of Geophysics, 60, e2021RG000 736, https://doi.org/10.1029/2021RG000736, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2021RG000736, 2022.

Desai, A.: FLUXNET2015 US-WCr Willow Creek, https://doi.org/10.18140/flx/1440095, accessed: 2022-11-01, 2016a.

Desai, A.: FLUXNET2015 US-Syv Sylvania Wilderness Area, https://doi.org/10.18140/flx/1440091, accessed: 2022-11-01, 2016b.

650   Desai, A.: FLUXNET2015 US-Los Lost Creek, https://doi.org/10.18140/flx/1440076, accessed: 2022-11-01, 2016c.

Desai, A.: FLUXNET2015 US-PFa Park Falls/WLEF, https://doi.org/10.18140/flx/1440089, accessed: 2022-11-01, 2016d.

Desai, A.: AmeriFlux FLUXNET-1F US-CS1 Central Sands Irrigated Agricultural Field, https://doi.org/10.17190/amf/1881576, accessed: 2022-11-01, 2022a.

Desai, A.: AmeriFlux FLUXNET-1F US-CS4 Central Sands Irrigated Agricultural Field, https://doi.org/10.17190/amf/1881579, accessed:
655   2022-11-01, 2022b.

Desai, A.: AmeriFlux FLUXNET-1F US-CS2 Tri County School Pine Forest, https://doi.org/10.17190/amf/1881577, accessed: 2022-11-01, 2022c.

Desai, A.: AmeriFlux FLUXNET-1F US-CS3 Central Sands Irrigated Agricultural Field, https://doi.org/10.17190/amf/1881578, accessed: 2022-11-01, 2022d.

660   Dolman, H., Hendriks, D., Parmentier, F.-J., Marchesini, L. B., Dean, J., and Van Huissteden, K.: FLUXNET2015 NL-Hor Horstermeer, https://doi.org/10.18140/flx/1440177, accessed: 2022-11-01, 2016a.

Dolman, H., Van Der Molen, M., Parmentier, F.-J., Marchesini, L. B., Dean, J., Van Huissteden, K., and Maximov, T.: FLUXNET2015 RU-Cok Chokurdakh, https://doi.org/10.18140/flx/1440182, accessed: 2022-11-01, 2016b.

Dong, G.: FLUXNET2015 CN-Cng Changling, https://doi.org/10.18140/flx/1440209, accessed: 2022-11-01, 2016.

665   Drake, B. and Hinkle, R.: FLUXNET2015 US-KS1 Kennedy Space Center (Slash Pine), https://doi.org/10.18140/flx/1440074, accessed: 2022-11-01, 2016a.

Drake, B. and Hinkle, R.: FLUXNET2015 US-KS2 Kennedy Space Center (Scrub Oak), https://doi.org/10.18140/flx/1440075, accessed: 2022-11-01, 2016b.

Euskirchen, E.: AmeriFlux FLUXNET-1F US-BZo Bonanza Creek Old Thermokarst Bog, https://doi.org/10.17190/amf/1881571, accessed: 670   2022-11-01, 2022a.

Euskirchen, E.: AmeriFlux FLUXNET-1F US-BZB Bonanza Creek Thermokarst Bog, https://doi.org/10.17190/amf/1881569, accessed: 2022-11-01, 2022b.

Euskirchen, E.: AmeriFlux FLUXNET-1F US-BZF Bonanza Creek Rich Fen, https://doi.org/10.17190/amf/1881570, accessed: 2022-11-01, 2022c.

675 Euskirchen, E.: AmeriFlux FLUXNET-1F US-BZS Bonanza Creek Black Spruce, https://doi.org/10.17190/amf/1881572, accessed: 2022-11-01, 2022d.

Euskirchen, E., Shaver, G., and Bret-Harte, S.: AmeriFlux FLUXNET-1F US-ICs Imnavait Creek Watershed Wet Sedge Tundra, https://doi.org/10.17190/amf/1871138, accessed: 2022-11-01, 2022a.

Euskirchen, E., Shaver, G., and Bret-Harte, S.: AmeriFlux FLUXNET-1F US-ICt Imnavait Creek Watershed Tussock Tundra, 680   https://doi.org/10.17190/amf/1881583, accessed: 2022-11-01, 2022b.

Ewers, B. and Pendall, E.: FLUXNET2015 US-Sta Saratoga, https://doi.org/10.18140/flx/1440115, accessed: 2022-11-01, 2016.

Faassen, K. A. P., Vilà-Guerau de Arellano, J., González-Armas, R., Heusinkveld, B. G., Mammarella, I., Peters, W., and Luijkx, I. T.: Separating above-canopy CO_2 and O_2 measurements into their atmospheric and biospheric signatures, Biogeosciences, 21, 3015–3039, https://doi.org/10.5194/bg-21-3015-2024, publisher: Copernicus GmbH, 2024.

685 Fares, S.: FLUXNET2015 US-Lin Lindcove Orange Orchard, https://doi.org/10.18140/flx/1440107, accessed: 2022-11-01, 2016.

Flerchinger, G.: AmeriFlux FLUXNET-1F US-Rwf RCEW Upper Sheep Prescibed Fire, https://doi.org/10.17190/amf/1881591, accessed: 2022-11-01, 2022a.

Flerchinger, G.: AmeriFlux FLUXNET-1F US-Rws Reynolds Creek Wyoming Big Sagebrush, https://doi.org/10.17190/amf/1881592, accessed: 2022-11-01, 2022b.

690 Flerchinger, G.: AmeriFlux FLUXNET-1F US-Rms RCEW Mountain Big Sagebrush, https://doi.org/10.17190/amf/1881587, accessed: 2022-11-01, 2022c.

Flerchinger, G. and Reba, M.: AmeriFlux FLUXNET-1F US-Rwe RCEW Reynolds Mountain East, https://doi.org/10.17190/amf/1871143, accessed: 2022-11-01, 2022.

Forsythe, J., Kline, M., and O'Halloran, T.: AmeriFlux FLUXNET-1F US-HB1 North Inlet Crab Haul Creek, 695   https://doi.org/10.17190/amf/1832160, accessed: 2022-11-01, 2021.

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

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., 700   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, O., 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, 705   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.,
710 Watanabe, M., Wimart-Rousseau, C., Yang, D., Yang, X., Yuan, W., Yue, X., Zaehle, S., Zeng, J., and Zheng, B.: Global Carbon Budget 2023, Earth System Science Data, 15, 5301–5369, https://doi.org/10.5194/essd-15-5301-2023, publisher: Copernicus GmbH, 2023.

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,
715 Agricultural and Forest Meteorology, 263, 292–307, https://doi.org/10.1016/j.agrformet.2018.09.001, 2018.

Gao, B.-c.: NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space, Remote Sensing of Environment, 58, 257–266, https://doi.org/10.1016/S0034-4257(96)00067-3, 1996.

Garcia, A., Di Bella, C., Houspanossian, J., Magliano, P., Jobbágy, E., Posse, G., Fernández, R., and Nosetto, M.: FLUXNET2015 AR-SLu San Luis, https://doi.org/10.18140/flx/1440191, accessed: 2022-11-01, 2016.

720 Gaubert, B., Edwards, D. P., Anderson, J. L., Arellano, A. F., Barré, J., Buchholz, R. R., Darras, S., Emmons, L. K., Fillmore, D., Granier, C., Hannigan, J. W., Ortega, I., Raeder, K., Soulié, A., Tang, W., Worden, H. M., and Ziskin, D.: Global Scale Inversions from MOPITT CO and MODIS AOD, Remote Sensing, 15, 4813, https://doi.org/10.3390/rs15194813, number: 19 Publisher: Multidisciplinary Digital Publishing Institute, 2023.

Goldstein, A.: FLUXNET2015 US-Blo Blodgett Forest, https://doi.org/10.18140/flx/1440068, accessed: 2022-11-01, 2016.

725 Goslee, S.: AmeriFlux FLUXNET-1F US-HWB USDA ARS Pasture Sytems and Watershed Management Research Unit- Hawbecker Site, https://doi.org/10.17190/amf/1881582, accessed: 2022-11-01, 2022.

Gough, C., Bohrer, G., and Curtis, P.: FLUXNET2015 US-UMB Univ. Of Mich. Biological Station, https://doi.org/10.18140/flx/1440093, accessed: 2022-11-01, 2016.

Gough, C., Bohrer, G., and Curtis, P.: AmeriFlux FLUXNET-1F US-UMd UMBS Disturbance, https://doi.org/10.17190/amf/1881597, ac-
730 cessed: 2022-11-01, 2022.

Goulden, M.: FLUXNET2015 CA-NS2 UCI-1930 Burn Site, https://doi.org/10.18140/flx/1440037, accessed: 2022-11-01, 2016a.

Goulden, M.: FLUXNET2015 CA-NS3 UCI-1964 Burn Site, https://doi.org/10.18140/flx/1440038, accessed: 2022-11-01, 2016b.

Goulden, M.: FLUXNET2015 CA-NS4 UCI-1964 Burn Site Wet, https://doi.org/10.18140/flx/1440039, accessed: 2022-11-01, 2016c.

Goulden, M.: FLUXNET2015 BR-Sa3 Santarem-Km83-Logged Forest, https://doi.org/10.18140/flx/1440033, accessed: 2022-11-01, 2016d.

735 Goulden, M.: FLUXNET2015 CA-NS6 UCI-1989 Burn Site, https://doi.org/10.18140/flx/1440041, accessed: 2022-11-01, 2016e.

Goulden, M.: FLUXNET2015 CA-NS7 UCI-1998 Burn Site, https://doi.org/10.18140/flx/1440042, accessed: 2022-11-01, 2016f.

Goulden, M.: FLUXNET2015 CA-NS5 UCI-1981 Burn Site, https://doi.org/10.18140/flx/1440040, accessed: 2022-11-01, 2016g.

Griffis, T. and Roman, T.: AmeriFlux FLUXNET-1F PE-QFR Quistococha Forest Reserve, https://doi.org/10.17190/amf/1832157, accessed: 2022-11-01, 2021.

740 Gruening, C., Goded, I., Cescatti, A., Manca, G., and Seufert, G.: FLUXNET2015 IT-SRo San Rossore, https://doi.org/10.18140/flx/1440176, accessed: 2022-11-01, 2016a.

Gruening, C., Goded, I., Cescatti, A., and Pokorska, O.: FLUXNET2015 IT-Isp Ispra ABC-IS, https://doi.org/10.18140/flx/1440234, ac-cessed: 2022-11-01, 2016b.

Hansen, B. U.: FLUXNET2015 GL-NuF Nuuk Fen, https://doi.org/10.18140/flx/1440222, accessed: 2022-11-01, 2016.

745 Haszpra, L.: Multi-decadal atmospheric carbon dioxide measurements in Hungary, central Europe, Atmospheric Measurement Techniques, 17, 4629–4647, https://doi.org/10.5194/amt-17-4629-2024, publisher: Copernicus GmbH, 2024.

Hayek, M. N., Wehr, R., Longo, M., Hutyra, L. R., Wiedemann, K., Munger, J. W., Bonal, D., Saleska, S. R., Fitzjarrald, D. R., and Wofsy, S. C.: A novel correction for biases in forest eddy covariance carbon balance, Agricultural and Forest Meteorology, 250-251, 90–101, https://doi.org/10.1016/j.agrformet.2017.12.186, 2018.

750 Hegarty, J., Draxler, R. R., Stein, A. F., Brioude, J., Mountain, M., Eluszkiewicz, J., Nehrkorn, T., Ngan, F., and Andrews, A.: Evaluation of Lagrangian Particle Dispersion Models with Measurements from Controlled Tracer Releases, Journal of Applied Meteorology and Climatology, 52, 2623–2637, https://doi.org/10.1175/JAMC-D-13-0125.1, publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology, 2013.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Sim-
755 mons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Vil- laume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3803, 2020.

760 Hinkle, R.: AmeriFlux FLUXNET-1F US-KS3 Kennedy Space Center (Salt Marsh), https://doi.org/10.17190/amf/1881586, accessed: 2022-11-01, 2022.

Hollinger, D.: AmeriFlux FLUXNET-1F US-Ho2 Howland Forest (West Tower), https://doi.org/10.17190/amf/1881581, accessed: 2022-11-01, 2022.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G.: Overview of the radiometric and biophysical performance of
765 the MODIS vegetation indices, Remote Sensing of Environment, 83, 195–213, https://doi.org/10.1016/S0034-4257(02)00096-2, 2002.

Huggins, D.: AmeriFlux FLUXNET-1F US-CF1 CAF-LTAR Cook East, https://doi.org/10.17190/amf/1832158, accessed: 2022-11-01, 2021.

Huggins, D.: AmeriFlux FLUXNET-1F US-CF3 CAF-LTAR Boyd North, https://doi.org/10.17190/amf/1881574, accessed: 2022-11-01, 2022a.

Huggins, D.: AmeriFlux FLUXNET-1F US-CF4 CAF-LTAR Boyd South, https://doi.org/10.17190/amf/1881575, accessed: 2022-11-01,
770 2022b.

Huggins, D.: AmeriFlux FLUXNET-1F US-CF2 CAF-LTAR Cook West, https://doi.org/10.17190/amf/1881573, accessed: 2022-11-01, 2022c.

Jones, M. W., Andrew, R. M., Peters, G. P., Janssens-Maenhout, G., De-Gol, A. J., Ciais, P., Patra, P. K., Chevallier, F., and Le Quéré, C.: Gridded fossil CO2 emissions and related O2 combustion consistent with national inventories 1959–2018, Scientific Data, 8, 2,
775 https://doi.org/10.1038/s41597-020-00779-6, number: 1 Publisher: Nature Publishing Group, 2021.

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,
780 Biogeosciences, 17, 1343–1365, https://doi.org/10.5194/bg-17-1343-2020, publisher: Copernicus GmbH, 2020.

Jung, M., Nelson, J., Migliavacca, M., El-Madany, T., Papale, D., Reichstein, M., Walther, S., and Wutzler, T.: Technical note: Flagging inconsistencies in flux tower data, Biogeosciences, 21, 1827–1846, https://doi.org/10.5194/bg-21-1827-2024, publisher: Copernicus GmbH, 2023.

Jung, M., Nelson, J., Migliavacca, M., El-Madany, T., Papale, D., Reichstein, M., Walther, S., and Wutzler, T.: Technical note: Flagging inconsistencies in flux tower data, Biogeosciences, 21, 1827–1846, https://doi.org/10.5194/bg-21-1827-2024, publisher: Copernicus GmbH, 2024.

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.

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: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7482–7491, IEEE, Salt Lake City, UT, USA, ISBN 978-1-5386-6420-9, https://doi.org/10.1109/CVPR.2018.00781, 2018.

Klatt, J., Schmid, H. P., Mauder, M., and Steinbrecher, R.: FLUXNET2015 DE-SfN Schechenfilz Nord, https://doi.org/10.18140/flx/1440219, accessed: 2022-11-01, 2016.

Knohl, A., Tiedemann, F., Kolle, O., Schulze, E.-D., Anthoni, P., Kutsch, W., Herbst, M., and Siebicke, L.: FLUXNET2015 DE-Lnf Leinefelde, https://doi.org/10.18140/flx/1440150, accessed: 2022-11-01, 2016.

Knox, S.: AmeriFlux FLUXNET-1F CA-DB2 Delta Burns Bog 2, https://doi.org/10.17190/amf/1881564, accessed: 2022-11-01, 2022.

Kobayashi, H. and Suzuki, R.: FLUXNET2015 US-Prr Poker Flat Research Range Black Spruce Forest, https://doi.org/10.18140/flx/1440113, accessed: 2022-11-01, 2016.

Kondo, M., Ichii, K., Takagi, H., and Sasakawa, M.: Comparison of the data-driven top-down and bottom-up global terrestrial CO 2 exchanges: GOSAT CO 2 inversion and empirical eddy flux upscaling: Data-driven Terrestrial CO 2 Exchanges, Journal of Geophysical Research: Biogeosciences, 120, https://doi.org/10.1002/2014JG002866, 2015.

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.

Kosugi, Y. and Takanashi, S.: FLUXNET2015 MY-PSO Pasoh Forest Reserve (PSO), https://doi.org/10.18140/flx/1440240, accessed: 2022-11-01, 2016.

Kotani, A.: FLUXNET2015 JP-SMF Seto Mixed Forest Site, https://doi.org/10.18140/flx/1440239, accessed: 2022-11-01, 2016a.

Kotani, A.: FLUXNET2015 JP-MBF Moshiri Birch Forest Site, https://doi.org/10.18140/flx/1440238, accessed: 2022-11-01, 2016b.

Kraft, B., Nelson, J. A., Walther, S., Gans, F., Weber, U., Duveiller, G., Reichstein, M., Zhang, W., Rußwurm, M., Tuia, D., Körner, M., Hamdi, Z. M., and Jung, M.: On the added value of sequential deep learning for upscaling evapotranspiration, EGUsphere, pp. 1–30, https://doi.org/10.5194/egusphere-2024-2896, publisher: Copernicus GmbH, 2024.

Kurc, S.: AmeriFlux FLUXNET-1F US-SRC Santa Rita Creosote, https://doi.org/10.17190/amf/1871145, accessed: 2022-11-01, 2022.

Kusak, K., Sanchez, C., Szutu, D., and Baldocchi, D.: AmeriFlux FLUXNET-1F US-Snf Sherman Barn, https://doi.org/10.17190/amf/1854371, accessed: 2022-11-01, 2022.

Kutzbach, L.: AmeriFlux FLUXNET-1F AR-TF1 Rio Moat Bog, https://doi.org/10.17190/amf/1818370, accessed: 2022-11-01, 2021.

Law, B.: FLUXNET2015 US-Me3 Metolius-second Young Aged Pine, https://doi.org/10.18140/flx/1440080, accessed: 2022-11-01, 2016a.

820 Law, B.: FLUXNET2015 US-Me6 Metolius Young Pine Burn, https://doi.org/10.18140/flx/1440099, accessed: 2022-11-01, 2016b.

Law, B.: FLUXNET2015 US-Me1 Metolius - Eyerly Burn, https://doi.org/10.18140/flx/1440078, accessed: 2022-11-01, 2016c.

Law, B.: FLUXNET2015 US-Me5 Metolius-first Young Aged Pine, https://doi.org/10.18140/flx/1440082, accessed: 2022-11-01, 2016d.

Law, B.: FLUXNET2015 US-Me4 Metolius-old Aged Ponderosa Pine, https://doi.org/10.18140/flx/1440081, accessed: 2022-11-01, 2016e.

Law, B.: AmeriFlux FLUXNET-1F US-Me2 Metolius Mature Ponderosa Pine, https://doi.org/10.17190/amf/1854368, accessed: 2022-11-01,

825 2022.

Liddell, M. J.: FLUXNET2015 AU-Rob Robson Creek, Queensland, Australia, https://doi.org/10.18140/flx/1440203, accessed: 2022-11-01, 2016.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, C. A.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Trans-

830 port (STILT) model, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/10.1029/2002JD003161, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2002JD003161, 2003.

Lindauer, M., Steinbrecher, R., Wolpert, B., Mauder, M., and Schmid, H. P.: FLUXNET2015 DE-Lkb Lackenberg, https://doi.org/10.18140/flx/1440214, accessed: 2022-11-01, 2016.

Litvak, M.: AmeriFlux FLUXNET-1F US-Mpj Mountainair Pinyon-Juniper Woodland, https://doi.org/10.17190/amf/1832161, accessed:

835 2022-11-01, 2021.

Litvak, M.: AmeriFlux FLUXNET-1F US-Wjs Willard Juniper Savannah, https://doi.org/10.17190/amf/1871146, accessed: 2022-11-01, 2022.

Liu, H., Huang, M., and Chen, X.: AmeriFlux FLUXNET-1F US-Hn3 Hanford 100H Sagebrush, https://doi.org/10.17190/amf/1881580, accessed: 2022-11-01, 2022.

840 Lohila, A., Aurela, M., Tuovinen, J.-P., Hatakka, J., and Laurila, T.: FLUXNET2015 FI-Jok Jokioinen, https://doi.org/10.18140/flx/1440159, accessed: 2022-11-01, 2016.

Luijkx, I. T., Chevallier, F., Roedenbeck, C., Niwa, Y., Liu, J., Feng, L., Palmer, P. I., Bowman, K., Peters, W., Tian, X., Piao, S., Zheng, B., Lloret, Z., Cozic, A., Jacobson, A. R., Yun, J., Byrne, B., Bloom, A., Jin, Z., Wang, Y., Zhang, H., Zhao, M., Wang, T., Ding, J., Liu, Z., Jiang, F., Ju, W., Yang, D., Chandra, N., and Patra, P.: Global CO2 gridded flux fields from 14 atmospheric inversions in GCB2023,

845 https://doi.org/10.18160/4M52-VCRU, 2024.

Lund, M., Jackowicz-Korczyński, M., and Abermann, J.: FLUXNET2015 GL-ZaH Zackenberg Heath, https://doi.org/10.18140/flx/1440224, accessed: 2022-11-01, 2016a.

Lund, M., Jackowicz-Korczyński, M., and Abermann, J.: FLUXNET2015 GL-ZaF Zackenberg Fen, https://doi.org/10.18140/flx/1440223, accessed: 2022-11-01, 2016b.

850 Macfarlane, C., Lambert, P., Byrne, J., Johnstone, C., and Smart, N.: FLUXNET2015 AU-Gin Gingin, https://doi.org/10.18140/flx/1440199, accessed: 2022-11-01, 2016.

Manca, G. and Goded, I.: FLUXNET2015 IT-PT1 Parco Ticino Forest, https://doi.org/10.18140/flx/1440172, accessed: 2022-11-01, 2016.

Margolis, H. A.: FLUXNET2015 CA-Qfo Quebec - Eastern Boreal, Mature Black Spruce, https://doi.org/10.18140/flx/1440045, accessed: 2022-11-01, 2016.

855 Massman, B.: FLUXNET2015 US-GBT GLEES Brooklyn Tower, https://doi.org/10.18140/flx/1440118, accessed: 2022-11-01, 2016.

Massman, B.: AmeriFlux FLUXNET-1F US-GLE GLEES, https://doi.org/10.17190/amf/1871136, accessed: 2022-11-01, 2022.

Matamala, R.: FLUXNET2015 US-IB2 Fermi National Accelerator Laboratory- Batavia (Prairie Site), https://doi.org/10.18140/flx/1440072, accessed: 2022-11-01, 2016.

Matteucci, G.: FLUXNET2015 IT-Col Collelongo, https://doi.org/10.18140/flx/1440167, accessed: 2022-11-01, 2016.

860 McCaughey, H.: FLUXNET2015 CA-Gro Ontario - Groundhog River, Boreal Mixedwood Forest, https://doi.org/10.18140/flx/1440034, accessed: 2022-11-01, 2016.

Merbold, L., Rebmann, C., and Corradi, C.: FLUXNET2015 RU-Che Cherski, https://doi.org/10.18140/flx/1440181, accessed: 2022-11-01, 2016.

Meyer, W., Cale, P., Koerber, G., Ewenz, C., and Sun, Q.: FLUXNET2015 AU-Cpr Calperum, https://doi.org/10.18140/flx/1440195, ac-
865 cessed: 2022-11-01, 2016.

Meyers, T.: FLUXNET2015 US-LWW Little Washita Watershed, https://doi.org/10.18140/flx/1440077, accessed: 2022-11-01, 2016a.

Meyers, T.: FLUXNET2015 US-Goo Goodwin Creek, https://doi.org/10.18140/flx/1440070, accessed: 2022-11-01, 2016b.

Moncrieff, J., Malhi, Y., and Leuning, R.: The propagation of errors in long-term measurements of land-atmosphere fluxes of carbon and water, Global Change Biology, 2, 231–240, https://doi.org/10.1111/j.1365-2486.1996.tb00075.x, _eprint:
870 https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2486.1996.tb00075.x, 1996.

Munassar, S., Rödenbeck, C., Koch, F.-T., Totsche, K. U., Gałkowski, M., Walther, S., and Gerbig, C.: Net ecosystem exchange (NEE) estimates 2006–2019 over Europe from a pre-operational ensemble-inversion system, Atmospheric Chemistry and Physics, 22, 7875–7892, https://doi.org/10.5194/acp-22-7875-2022, publisher: Copernicus GmbH, 2022.

Munger, J. W.: FLUXNET2015 US-Ha1 Harvard Forest EMS Tower (HFR1), https://doi.org/10.18140/flx/1440071, accessed: 2022-11-01,
875 2016.

Nash, J. and Sutcliffe, J.: River Flow Forecasting Through Conceptual Models Part I — A Discussion of Principles, Journal of Hydrology, 10, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6, publisher: Elsevier BV, 1970.

Nelson* and Walther*, Gans, F., Kraft, B., Weber, U., Novick, K., Buchmann, N., Migliavacca, M., Wohlfahrt, G., Šigut, L., Ibrom, A., Papale, D., Göckede, M., Duveiller, G., Knohl, A., Hörtnagl, L., Scott, R. L., Zhang, W., Hamdi, Z. M., Reichstein, M., Aranda-Barranco,
880 S., Ardö, J., Op de Beeck, M., Billdesbach, D., Bowling, D., Bracho, R., Brümmer, C., Camps-Valls, G., Chen, S., Cleverly, J. R., Desai, A., Dong, G., El-Madany, T. S., Euskirchen, E. S., Feigenwinter, I., Galvagno, M., Gerosa, G., Gielen, B., Goded, I., Goslee, S., Gough, C. M., Heinesch, B., Ichii, K., Jackowicz-Korczynski, M. A., Klosterhalfen, A., Knox, S., Kobayashi, H., Kohonen, K.-M., Korkiakoski, M., Mammarella, I., Mana, G., Marzuoli, R., Matamala, R., Metzger, S., Montagnani, L., Nicolini, G., O'Halloran, T., Ourcival, J.-M., Peichl, M., Pendall, E., Ruiz Reverter, B., Roland, M., Sabbatini, S., Sachs, T., Schmidt, M., Schwalm, C. R., Shekhar, A., Silberstein,
885 R., Silveira, M. L., Spano, D., Tagesson, T., Tramontana, G., Trotta, C., Turco, F., Vesala, T., Vincke, C., Vitale, D., Vivoni, E. R., Wang, Y., Woodgate, W., Yepez, E. A., Zhang, J., Zona, D., and Jung, M.: X-BASE: the first terrestrial carbon and water flux products from an extended data-driven scaling framework, FLUXCOM-X, EGUsphere, pp. 1–51, https://doi.org/10.5194/egusphere-2024-165, publisher: Copernicus GmbH, 2024.

Network), N.: AmeriFlux FLUXNET-1F US-xBR NEON Bartlett Experimental Forest (BART), https://doi.org/10.17190/amf/1881598, ac-
890 cessed: 2022-11-01, 2022.

Nouvellon, Y.: FLUXNET2015 CG-Tch Tchizalamou, https://doi.org/10.18140/flx/1440142, accessed: 2022-11-01, 2016.

Novick, K. and Phillips, R.: AmeriFlux FLUXNET-1F US-MMS Morgan Monroe State Forest, https://doi.org/10.17190/amf/1854369, ac-
cessed: 2022-11-01, 2022.

Oikawa, P.: AmeriFlux FLUXNET-1F US-EDN Eden Landing Ecological Reserve, https://doi.org/10.17190/amf/1832159, accessed: 2022-
895   11-01, 2021.

Olesen, J.: FLUXNET2015 DK-Fou Foulum, https://doi.org/10.18140/flx/1440154, accessed: 2022-11-01, 2016.

Ourcival, J.-M.: FLUXNET2015 FR-Pue Puechabon, https://doi.org/10.18140/flx/1440164, accessed: 2022-11-01, 2016.

O'Sullivan, M., Sitch, S., Friedlingstein, P., Luijkx, I. T., Peters, W., Rosan, T. M., Arneth, A., Arora, V. K., Chandra, N., Chevallier, F., Ciais,
        P., Falk, S., Feng, L., Gasser, T., Houghton, R. A., Jain, A. K., Kato, E., Kennedy, D., Knauer, J., McGrath, M. J., Niwa, Y., Palmer, P. I.,
900   Patra, P. K., Pongratz, J., Poulter, B., Rödenbeck, C., Schwingshackl, C., Sun, Q., Tian, H., Walker, A. P., Yang, D., Yuan, W., Yue, X., and
        Zaehle, S.: The key role of forest disturbance in reconciling estimates of the northern carbon sink, Communications Earth & Environment,
        5, 1–10, https://doi.org/10.1038/s43247-024-01827-4, publisher: Nature Publishing Group, 2024.

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:
905   Nature Publishing Group, 2019.

Papale, D., Tirone, G., Valentini, R., Arriga, N., Belelli, L., Consalvo, C., Dore, S., Manca, G., Mazzenga, F., Sabbatini, S., and Stefani, P.:
        FLUXNET2015 IT-Ro2 Roccarespampani 2, https://doi.org/10.18140/flx/1440175, accessed: 2022-11-01, 2016.

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

Siebicke, L., Šigut, L., Silberstein, R., Sirca, C., Spano, D., Steinbrecher, R., Stevens, R. M., Sturtevant, C., Suyker, A., Tagesson, T., Takanashi, S., Tang, Y., Tapper, N., Thom, J., Tomassucci, M., Tuovinen, J.-P., Urbanski, S., Valentini, R., van der Molen, M., van Gorsel, E., van Huissteden, K., Varlagin, A., Verfaillie, J., Vesala, T., Vincke, C., Vitale, D., Vygodskaya, N., Walker, J. P., Walter-Shea, E., Wang,
935 H., Weber, R., Westermann, S., Wille, C., Wofsy, S., Wohlfahrt, G., Wolf, S., Woodgate, W., Li, Y., Zampedri, R., Zhang, J., Zhou, G., Zona, D., Agarwal, D., Biraud, S., Torn, M., and Papale, D.: The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data, Scientific Data, 7, 225, https://doi.org/10.1038/s41597-020-0534-3, number: 1 Publisher: Nature Publishing Group, 2020.

Pendall, E. and Griebel, A.: FLUXNET2015 AU-Cum Cumberland Plains, https://doi.org/10.18140/flx/1440196, accessed: 2022-11-01, 2016.

940 Peters, W., Krol, M. C., Van Der WERF, G. R., Houweling, S., Jones, C. D., Hughes, J., Schaefer, K., Masarie, K. A., Jacobson, A. R., Miller, J. B., Cho, C. H., Ramonet, M., Schmidt, M., Ciattaglia, L., Apadula, F., Heltai, D., Meinhardt, F., Di Sarra, A. G., Piacentino, S., Sferlazzo, D., Aalto, T., Hatakka, J., Ström, J., Haszpra, L., Meijer, H. a. J., Van Der Laan, S., Neubert, R. E. M., Jordan, A., Rodó, X., Morguí, J.-A., Vermeulen, A. T., Popa, E., Rozanski, K., Zimnoch, M., Manning, A. C., Leuenberger, M., Uglietti, C., Dolman, A. J., Ciais, P., Heimann, M., and Tans, P. P.: Seven years of recent European net terrestrial carbon dioxide exchange con-
945 strained by atmospheric observations, Global Change Biology, 16, 1317–1337, https://doi.org/10.1111/j.1365-2486.2009.02078.x, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2486.2009.02078.x, 2010.

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.

950 Pilegaard, K. and Ibrom, A.: FLUXNET2015 DK-Eng Enghave, https://doi.org/10.18140/flx/1440153, accessed: 2022-11-01, 2016.

Posse, G., Lewczuk, N., Richter, K., and Cristiano, P.: FLUXNET2015 AR-Vir Virasoro, https://doi.org/10.18140/flx/1440192, accessed: 2022-11-01, 2016.

Poveda, F. D., Ballesteros, A. L., Cañete, E. P. S., Ortiz, P. S., Jiménez, M. R. M., Priego, O. P., and Kowalski, A. S.: FLUXNET2015 ES-Amo Amoladeras, https://doi.org/10.18140/flx/1440156, accessed: 2022-11-01, 2016.

955 RANDERSON, J., VAN DER WERF, G., GIGLIO, L., COLLATZ, G., and KASIBHATLA, P.: Global Fire Emissions Database, Version 4.1 (GFEDv4), https://doi.org/10.3334/ornldaac/1293, 2017.

Reverter, B. R., Perez-Cañete, E. S., and Kowalski, A. S.: FLUXNET2015 ES-Ln2 Lanjaron-Salvage Logging, https://doi.org/10.18140/flx/1440226, accessed: 2022-11-01, 2016a.

Reverter, B. R., Perez-Cañete, E. S., and Kowalski, A. S.: FLUXNET2015 ES-LgS Laguna Seca, https://doi.org/10.18140/flx/1440225,
960 accessed: 2022-11-01, 2016b.

Rey-Sanchez, C., Wang, C., Szutu, D., Hemes, K., Verfaillie, J., and Baldocchi, D.: AmeriFlux FLUXNET-1F US-Bi2 Bouldin Island Corn, https://doi.org/10.17190/amf/1871135, accessed: 2022-11-01, 2022a.

Rey-Sanchez, C., Wang, C., Szutu, D., Shortt, R., Chamberlain, S., Verfaillie, J., and Baldocchi, D.: AmeriFlux FLUXNET-1F US-Bi1 Bouldin Island Alfalfa, https://doi.org/10.17190/amf/1871134, accessed: 2022-11-01, 2022b.

965 RI, I.: Ecosystem Final Quality (L2) Product in ETC-Archive Format - Release 2021-1, https://doi.org/10.18160/fzmy-pg92, accessed: 2022-11-01, 2021-05-06.

RI, I.: Ecosystem Final Quality (L2) Product in ETC-Archive Format - Release 2022-1, https://doi.org/10.18160/pad9-hqhu, accessed: 2022-11-01, 2022-06-14.

Rödenbeck, C., Houweling, S., Gloor, M., and Heimann, M.: CO2 flux history 1982–2001 inferred from atmospheric data using a global in-
version 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.: History of El Niño impacts on the global carbon cycle 1957–2017: a
quantification from atmospheric CO2 data, Philosophical Transactions of the Royal Society B: Biological Sciences, 373, 20170 303,
https://doi.org/10.1098/rstb.2017.0303, publisher: Royal Society, 2018.

Sabbatini, S., Arriga, N., Gioli, B., and Papale, D.: FLUXNET2015 IT-CA2 Castel d'Asso2, https://doi.org/10.18140/flx/1440231, accessed:
2022-11-01, 2016a.

Sabbatini, S., Arriga, N., Matteucci, G., and Papale, D.: FLUXNET2015 IT-CA3 Castel d'Asso 3, https://doi.org/10.18140/flx/1440232,
accessed: 2022-11-01, 2016b.

Sabbatini, S., Arriga, N., and Papale, D.: FLUXNET2015 IT-CA1 Castel d'Asso1, https://doi.org/10.18140/flx/1440230, accessed: 2022-11-
01, 2016c.

Sachs, T., Wille, C., Larmanou, E., and Franz, D.: FLUXNET2015 DE-Zrk Zarnekow, https://doi.org/10.18140/flx/1440221, accessed: 2022-
11-01, 2016.

Saleska, S.: FLUXNET2015 BR-Sa1 Santarem-Km67-Primary Forest, https://doi.org/10.18140/flx/1440032, accessed: 2022-11-01, 2016.

Sanchez, C. R., Sturtevant, C., Szutu, D., Baldocchi, D., Eichelmann, E., and Knox, S.: FLUXNET2015 US-Tw4 Twitchell East End Wetland,
https://doi.org/10.18140/flx/1440111, accessed: 2022-11-01, 2016.

Schaaf, C. and Wang, Z.: MCD43A2 MODIS/Terra+Aqua BRDF/Albedo Quality Daily L3 Global - 500m V006,
https://doi.org/10.5067/modis/mcd43a2.006, accessed: 2022-11-01, 2015a.

Schaaf, C. and Wang, Z.: MCD43A4 MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF Adjusted Ref Daily L3 Global - 500m V006,
https://doi.org/10.5067/modis/mcd43a4.006, accessed: 2022-11-01, 2015b.

Schneider, K. and Schmidt, M.: FLUXNET2015 DE-Seh Selhausen, https://doi.org/10.18140/flx/1440217, accessed: 2022-11-01, 2016.

Schroder, I., Zegelin, S., Palu, T., and Feitz, A.: FLUXNET2015 AU-Emr Emerald, https://doi.org/10.18140/flx/1440198, accessed: 2022-
11-01, 2016.

Scott, R.: FLUXNET2015 US-SRG Santa Rita Grassland, https://doi.org/10.18140/flx/1440114, accessed: 2022-11-01, 2016a.

Scott, R.: FLUXNET2015 US-SRM Santa Rita Mesquite, https://doi.org/10.18140/flx/1440090, accessed: 2022-11-01, 2016b.

Scott, R.: FLUXNET2015 US-Wkg Walnut Gulch Kendall Grasslands, https://doi.org/10.18140/flx/1440096, accessed: 2022-11-01, 2016c.

Scott, R.: FLUXNET2015 US-Whs Walnut Gulch Lucky Hills Shrub, https://doi.org/10.18140/flx/1440097, accessed: 2022-11-01, 2016d.

Shao, C.: FLUXNET2015 CN-Sw2 Siziwang Grazed (SZWG), https://doi.org/10.18140/flx/1440212, accessed: 2022-11-01, 2016a.

Shao, C.: FLUXNET2015 CN-Du3 Duolun Degraded Meadow, https://doi.org/10.18140/flx/1440210, accessed: 2022-11-01, 2016b.

Shi, P., Zhang, X., and He, Y.: FLUXNET2015 CN-Dan Dangxiong, https://doi.org/10.18140/flx/1440138, accessed: 2022-11-01, 2016.

Shortt, R., Hemes, K., Szutu, D., Verfaillie, J., and Baldocchi, D.: AmeriFlux FLUXNET-1F US-Sne Sherman Island Restored Wetland,
https://doi.org/10.17190/amf/1871144, accessed: 2022-11-01, 2022.

Sigut, L., Havrankova, K., Jocher, G., Pavelka, M., and Janouš, D.: FLUXNET2015 CZ-BK2 Bily Kriz Grassland,
https://doi.org/10.18140/flx/1440144, accessed: 2022-11-01, 2016.

Silveira, M.: AmeriFlux FLUXNET-1F US-ONA Florida Pine Flatwoods, https://doi.org/10.17190/amf/1832163, accessed: 2022-11-01,
2021.

Spano, D., Duce, P., Marras, S., Sirca, C., Arca, A., Zara, P., and Ventura, A.: FLUXNET2015 IT-Noe Arca Di Noe - Le Prigionette, https://doi.org/10.18140/flx/1440171, accessed: 2022-11-01, 2016.

Staebler, R.: AmeriFlux FLUXNET-1F CA-Cbo Ontario - Mixed Deciduous, Borden Forest Site, https://doi.org/10.17190/amf/1854365, accessed: 2022-11-01, 2022.

1010 Sturtevant, C., Szutu, D., Baldocchi, D., Matthes, J. H., Oikawa, P., and Chamberlain, S. D.: FLUXNET2015 US-Myb Mayberry Wetland, https://doi.org/10.18140/flx/1440105, accessed: 2022-11-01, 2016.

Sturtevant, C., Verfaillie, J., and Baldocchi, D.: AmeriFlux FLUXNET-1F US-Tw2 Twitchell Corn, https://doi.org/10.17190/amf/1881593, accessed: 2022-11-01, 2022.

Suyker, A.: FLUXNET2015 US-Ne3 Mead - Rainfed Maize-Soybean Rotation Site, https://doi.org/10.18140/flx/1440086, accessed: 2022-1015   11-01, 2016a.

Suyker, A.: FLUXNET2015 US-Ne2 Mead - Irrigated Maize-Soybean Rotation Site, https://doi.org/10.18140/flx/1440085, accessed: 2022-11-01, 2016b.

Suyker, A.: AmeriFlux FLUXNET-1F US-Ne1 Mead - Irrigated Continuous Maize Site, https://doi.org/10.17190/amf/1871140, accessed: 2022-11-01, 2022.

1020 Tagesson, T., Ardö, J., and Fensholt, R.: FLUXNET2015 SN-Dhr Dahra, https://doi.org/10.18140/flx/1440246, accessed: 2022-11-01, 2016.

Tang, Y., Kato, T., and Du, M.: FLUXNET2015 CN-HaM Haibei Alpine Tibet Site, https://doi.org/10.18140/flx/1440190, accessed: 2022-11-01, 2016.

Team, D. . and Centre, I. E. T.: Drought-2018 Ecosystem Eddy Covariance Flux Product for 52 Stations in FLUXNET-Archive Format, https://doi.org/10.18160/yvr0-4898, accessed: 2022-11-01, 2020-03-10.

1025 Team, W. W. . and Centre, I. E. T.: Warm Winter 2020 Ecosystem Eddy Covariance Flux Product for 73 Stations in FLUXNET-Archive Format—Release 2022-1, https://doi.org/10.18160/2g60-zhak, accessed: 2022-11-01, 2022-02-01.

Torn, M.: FLUXNET2015 US-ARc ARM Southern Great Plains Control Site- Lamont, https://doi.org/10.18140/flx/1440065, accessed: 2022-11-01, 2016a.

Torn, M.: FLUXNET2015 US-ARb ARM Southern Great Plains Burn Site- Lamont, https://doi.org/10.18140/flx/1440064, accessed: 2022-1030   11-01, 2016b.

Torn, M. and Dengel, S.: AmeriFlux FLUXNET-1F US-NGB NGEE Arctic Barrow, https://doi.org/10.17190/amf/1832162, accessed: 2022-11-01, 2021.

Tran, D. A., Gerbig, C., Rödenbeck, C., Panov, A., Prokushkin, A., Winderlich, J., Lavrič, J., Heimann, M., Seifert, T., Schultz, U., Schmidt, S., and Zaehle, S.: Zotino Tall Tower Observatory (ZOTTO) hourly CO2 measurement dataset, https://doi.org/10.17617/3.YBPFG2, 1035   2024a.

Tran, D. A., Gerbig, C., Rödenbeck, C., and Zaehle, S.: Interannual Variations in Siberian Carbon Uptake and Carbon Release Period, preprint, Climate and Earth System/Atmospheric Modelling and Data Analysis/Troposphere/Physics (physical properties and processes), https://doi.org/10.5194/egusphere-2023-2573, 2024b.

Upton, S., Reichstein, M., Gans, F., Peters, W., Kraft, B., and Bastos, A.: Constraining biospheric carbon dioxide fluxes by combined 1040   top-down and bottom-up approaches, Atmospheric Chemistry and Physics, 24, 2555–2582, https://doi.org/10.5194/acp-24-2555-2024, publisher: Copernicus GmbH, 2024.

Valach, A., Shortt, R., Szutu, D., Eichelmann, E., Knox, S., Hemes, K., Verfaillie, J., and Baldocchi, D.: AmeriFlux FLUXNET-1F US-Tw1 Twitchell Wetland West Pond, https://doi.org/10.17190/amf/1832165, accessed: 2022-11-01, 2021.

Valach, A., Kasak, K., Szutu, D., Verfaillie, J., and Baldocchi, D.: AmeriFlux FLUXNET-1F US-Tw5 East Pond Wetland, https://doi.org/10.17190/amf/1881595, accessed: 2022-11-01, 2022.

Valentini, R., Dore, S., Mazzenga, F., Sabbatini, S., Stefani, P., Tirone, G., and Papale, D.: FLUXNET2015 IT-Cpz Castelporziano, https://doi.org/10.18140/flx/1440168, accessed: 2022-11-01, 2016a.

Valentini, R., Nicolini, G., Stefani, P., De Grandcourt, A., and Stivanello, S.: FLUXNET2015 GH-Ank Ankasa, https://doi.org/10.18140/flx/1440229, accessed: 2022-11-01, 2016b.

Valentini, R., Tirone, G., Vitale, D., Papale, D., Arriga, N., Belelli, L., Dore, S., Manca, G., Mazzenga, F., Pegoraro, E., Sabbatini, S., and Stefani, P.: FLUXNET2015 IT-Ro1 Roccarespampani 1, https://doi.org/10.18140/flx/1440174, accessed: 2022-11-01, 2016c.

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.

Vivoni, E. and Perez-Ruiz, E.: AmeriFlux FLUXNET-1F US-Jo2 Jornada Experimental Range Mixed Shrubland, https://doi.org/10.17190/amf/1881584, accessed: 2022-11-01, 2022.

Vourlitis, G., Dalmagro, H., De S. Nogueira, J., Johnson, M., and Arruda, P.: AmeriFlux FLUXNET-1F BR-Npw Northern Pantanal Wetland, https://doi.org/10.17190/amf/1881563, accessed: 2022-11-01, 2022.

Wagner-Riddle, C.: AmeriFlux FLUXNET-1F CA-ER1 Elora Research Station, https://doi.org/10.17190/amf/1832154, accessed: 2022-11-01, 2021.

Walther* and Besnard*, Nelson, J. A., El-Madany, T. S., Migliavacca, M., Weber, U., Carvalhais, N., Ermida, S. L., Brümmer, C., Schrader, F., Prokushkin, A. S., Panov, A. V., and Jung, M.: Technical note: A view from space on global flux towers by MODIS and Landsat: the FluxnetEO data set, Biogeosciences, 19, 2805–2840, https://doi.org/10.5194/bg-19-2805-2022, publisher: Copernicus GmbH, 2022.

Wang, H. and Fu, X.: FLUXNET2015 CN-Qia Qianyanzhou, https://doi.org/10.18140/flx/1440141, accessed: 2022-11-01, 2016.

Wohlfahrt, G., Hammerle, A., and Hörtnagl, L.: FLUXNET2015 AT-Neu Neustift, https://doi.org/10.18140/flx/1440121, accessed: 2022-11-01, 2016.

Wolf, S., Eugster, W., and Buchmann, N.: FLUXNET2015 PA-SPs Sardinilla-Pasture, https://doi.org/10.18140/flx/1440179, accessed: 2022-11-01, 2016a.

Wolf, S., Eugster, W., and Buchmann, N.: FLUXNET2015 PA-SPn Sardinilla Plantation, https://doi.org/10.18140/flx/1440180, accessed: 2022-11-01, 2016b.

Wood, J. and Gu, L.: AmeriFlux FLUXNET-1F US-MOz Missouri Ozark Site, https://doi.org/10.17190/amf/1854370, accessed: 2022-11-01, 2022.

Woodgate, W., Van Gorsel, E., and Leuning, R.: FLUXNET2015 AU-Tum Tumbarumba, https://doi.org/10.18140/flx/1440126, accessed: 2022-11-01, 2016.

Wu, D., Lin, J. C., Fasoli, B., Oda, T., Ye, X., Lauvaux, T., Yang, E. G., and Kort, E. A.: A Lagrangian approach towards extracting signals of urban CO_2 emissions from satellite observations of atmospheric column CO_2 (XCO_2): X-Stochastic Time-Inverted Lagrangian Transport model ("X-STILT v1"), Geoscientific Model Development, 11, 4843–4871, https://doi.org/10.5194/gmd-11-4843-2018, publisher: Copernicus GmbH, 2018.

Yepez, E. and Garatuza, J.: AmeriFlux FLUXNET-1F MX-Tes Tesopaco, Secondary Tropical Dry Forest, https://doi.org/10.17190/amf/1832156, accessed: 2022-11-01, 2021.

Zhang, J. and Han, S.: FLUXNET2015 CN-Cha Changbaishan, https://doi.org/10.18140/flx/1440137, accessed: 2022-11-01, 2016.

Zhou, G. and Yan, J.: FLUXNET2015 CN-Din Dinghushan, https://doi.org/10.18140/flx/1440139, accessed: 2022-11-01, 2016.

Zona, D. and Oechel, W.: FLUXNET2015 US-Atq Atqasuk, https://doi.org/10.18140/flx/1440067, accessed: 2022-11-01, 2016a.

1085   Zona, D. and Oechel, W.: FLUXNET2015 US-Ivo Ivotuk, https://doi.org/10.18140/flx/1440073, accessed: 2022-11-01, 2016b.