https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

Systematic underestimation of type-specific ecosystem process variability in the Community Land Model v5 over Europe

Christian Poppe Terán[1,2,3], Bibi S. Naz[1,3], Harry Vereecken[1,3], Roland Baatz[4], Rosie Fisher[5], Harrie-Jan Hendricks Franssen[1,3]

Correspondence to: Christian Poppe Terán (c.poppe@fz-juelich.de)

[1]Institute of Bio and Geosciences – Agrosphere (IBG-3), Research Centre Jülich, Jülich, 52428, DE
[2]Faculty of Georesources and Materials Engineering, RWTH Aachen, Aachen, 52062, DE
[3]HPSC TerrSys, Geoverbund ABC/J, Jülich, 52428, DE
[4]Research Platform Data Analysis & Simulation, Leibniz Centre for Agricultural Landscape Research (ZALF), Müncheberg, 15374, DE
[5]CICERO Center for International Climate Research, Oslo, 0349, NO

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

Abstract

Evapotranspiration (ET) and gross primary production (GPP) are critical fluxes contributing to the energy, water, and carbon exchanges between the atmosphere and the land surface. Land surface models such as the Community Land Model v5 (CLM5) quantify these fluxes, contribute to a better understanding of climate change's impact on ecosystems, and estimate the state of carbon budgets and water resources. Past studies have shown the ability of CLM5 to model ET and GPP magnitudes well but emphasized systematic underestimations and lower variability than in the observations.

Here, we evaluate the simulated ET and GPP from CLM5 at the grid scale ($CLM5_{grid}$) and the plant functional type (PFT) scale ($CLM5_{PFT}$) with observations from eddy covariance stations from the Integrated Carbon Observation System (ICOS) over Europe. For most PFTs, $CLM5_{grid}$ and $CLM5_{PFT}$ compared better to ICOS than publicly available reanalysis data and estimates obtained from remote sensing. $CLM5_{PFT}$ exhibited a low systematic error in simulating the ET of the ICOS measurements (average bias of -5.05%), implying that the PFT-specific ET matches the magnitude of the observations closely. However, $CLM5_{PFT}$ severely underestimates GPP, especially in deciduous forests (bias of -43.76%). Furthermore, the simulated ET and GPP distribution moments across PFTs in $CLM5_{grid}$ and $CLM5_{PFT}$, reanalyses, and remote sensing data indicate an underestimated spatiotemporal variability compared to the observations across Europe. These results are essential insights for further evaluations in CLM5 by pointing to the limitations of CLM5 in simulating the spatiotemporal variability of ET and GPP across PFTs.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

1. Introduction

Ecosystem processes, such as evapotranspiration (ET) and gross primary production (GPP), play an important role in cycling water, carbon, and energy between ecosystems and the atmosphere. Changes in the magnitude and variability of these fluxes can indicate ecosystems' inhibited performance due to changing environment (Kühn et al., 2021; Migliavacca et al., 2021). These changes can lead to short-term alterations and long-term trends in water resources and carbon pools in the atmosphere and the land surface. Thus, the accurate quantification of the variability of ecosystem processes is pivotal for developing climate change projections and formulating effective mitigation policies (Friedlingstein et al., 2023; Graf et al., 2023).

Notably, an accurate, functional understanding of land surface processes is essential to identify threatened ecosystems in the present and the future and facilitate carbon budget calculations. Land surface models (LSMs) serve as deterministic and processed-based simulators of ecosystems, capturing energy, water, and carbon fluxes while considering their interactions and the heterogeneity of the land surface (Fisher and Koven, 2020). LSMs can complement point-scale observations from in-situ research infrastructures by providing spatiotemporally uniform and extensive high-resolution outputs. Their high-resolution process-based simulations contrast the often coarsely resolved remote sensing data. Hence, LSMs are frequently used tools for investigating and projecting the current understanding of ecosystem processes, such as GPP and ET, on various scales. However, there is uncertainty in the LSM structure, the parameters, the input data, and the initial conditions, which carry over to the simulated variables. Therefore, assessing how well the general simulated ET and GPP variability compares to the observations is crucial. Such evaluations deliver essential context on LSM biases and form a basis for analyses of more complex ecosystem responses. Recent studies already found discrepancies between LSM simulations of ET and GPP and observations collected in the field and from remote sensing. For instance, these discrepancies are evident in their magnitude and variability (De Pue et al., 2023; Boas et al., 2023; Cheng et al., 2021; Strebel et al., 2023) and their response to drought (Ukkola et al., 2016; Wu et al., 2020; Green et al., 2024). Therefore, assessing the accuracy of LSMs in representing observed GPP and ET fluxes is crucial to test and improve our current understanding of ecosystem process variability and identify the limitations of state-of-the-art LSMs.

Current land surface models, e.g., the Joint UK Land Environment Simulator (JULES), the Community Land Model 5 (CLM5), or the Community Atmosphere Biosphere Land Exchange Modeel (CABLE), employ a tiling system within the grid cell to account for functional differences of distinct patches on the land surface. The

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

natural and crop vegetation is grouped into plant functional types (PFT), the entities for which ecosystem process calculations are resolved (Fisher and Koven, 2020; Bonan et al., 2002; Solomon and Shugart, 1993). Typically, PFTs are defined based on morphological and phenological characteristics of the vegetation (e.g., leaf type and leaf longevity) and climate (Bonan et al., 2002). However, the usefulness of this PFT definition, or at least its current coarsely resolved implementation, is a subject of debate (Caldararu et al., 2015; Van Bodegom et al., 2012). The primary argument against it is that observed plant traits implemented as PFT-related parameters vary to some extent in space and time in response to a changing environment. This spatiotemporal dependence of PFT traits is only marginally represented in LSMs. On top of that, most research assessing LSMs only used a handful of observation sites and did not analyze aggregated values for groups of sites observing the same PFT. Such analyses would provide essential insights, as a recent study highlighted the differences between vegetation type concepts used in observation networks, e.g., the International Geosphere-Biosphere Programme (IGBP) classification, and PFTs used in LSMs and underlined the importance of improving these PFT concepts (Cranko Page et al., 2024).

The phenology of ecosystem processes, i.e., their seasonal cycles and evolution through the year and the growing season length, have shifted in timing due to climate change. A recent study investigated which factors drive the changes in the mean annual dynamics of ecosystem processes in Europe (Rahmati et al., 2023), and many of these discovered feedbacks, for instance, the effect of increased atmospheric dryness on growing season length, are only implemented simplistically in LSMs. Furthermore, robust simulations of LSMs for impact assessments become even more critical as ecosystems experience more disturbances along with the changing climate. For example, projections show that droughts have recently become more frequent in Europe (Vautard et al., 2023; Rousi et al., 2022) and that these extreme events will become even more frequent and severe in the future (Lehner et al., 2017). While the combined effect of a higher occurrence of compound drought events is currently not fully understood, it is clear from observations that individual drought years, or droughts in general, have already had a profound impact on ecosystem processes in Europe (Graf et al., 2020; Van Der Woude et al., 2023; Poppe Terán et al., 2023). Given that the frequency and severity of extreme events affect GPP and ET's statistical distributions, investigating how the characteristics of the simulated distributions compare with the observed can contextualize findings of modeled ecosystem drought responses in Europe.

One predominantly used LSM is the Community Land Model version 5 (CLM5) (Lawrence et al., 2019, 2018). In the most recent version, CLM5 solves the biogeochemistry (BGC), i.e., the carbon and nitrogen cycles

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

between the atmosphere, vegetation, and soil. CLM5 has been widely employed for quantifying and examining ecosystems at various scales, including global (Xie et al., 2020; Sitch et al., 2015; Lawrence et al., 2019), regional (Cheng et al., 2021; Boas et al., 2023) and site-scale (Strebel et al., 2023; Umair et al., 2020; Song et al., 2020; Fisher et al., 2019a) applications. Several studies have highlighted the ability of CLM5 to simulate ecosystem processes close to the observations (Wozniak et al., 2020; Lawrence et al., 2019; Cheng et al., 2021; Zhang et al., 2023; Boas et al., 2023). However, they have also emphasized an underestimated magnitude and variability in the simulations across different time scales and under various conditions.

The present study assesses CLM5's ability to capture ecosystem processes at a continental scale. To ensure comparability to point scale observations, we conducted high-resolution simulations at 0.0275° (approx. 3 km) resolution over the European CORDEX domain (Giorgi et al., 2009), resulting in 1544 x 1592 grid cells. Notably, the output contained variables from the subgrid-scale, i.e., from within a 3 km grid cell, for PFT present in the grid cell. We then compared the CLM5 grid level ($CLM5_{grid}$) and PFT level data ($CLM5_{PFT}$) to observations from a continental network of sites: The Integrated Carbon Observation System (ICOS) provides the WARM-WINTER-2020 data (Warm Winter 2020 Team and ICOS Ecosystem Thematic Centre, 2022), which includes Eddy Covariance measurements over a dense network of over 70 sites in Europe. These ICOS data are regarded as the gold standard for calibrating and evaluating process-based models due to their ample spatial coverage as a network encompassing diverse land cover types. Thus, it offers an excellent opportunity to comprehensively assess simulated GPP and ET for specific PFT from our CLM5 setup over Europe.

Additionally, we include remote sensing data from the Global Land Surface Satellite (GLASS, (Liang et al., 2021)) and reanalyses from the European Center for Medium-range Weather Forecasts Reanalysis 5 - Land (ERA5L, Copernicus Climate Change Service (2019)) as well as from the Global Land Evaporation Amsterdam Model (GLEAM, Martens et al. (2017)) in our analyses to identify common patterns of ecosystem process variability between CLM5, in-situ observations, reanalysis, and remote sensing data.

In summary, this study uses ICOS observations as ground truth data and compares them with grid and PFT level CLM5 data, reanalyses, and remote sensing derivatives to:

1. Compare performance indices (root mean square error and percent bias) between the models and ICOS measurements on a per-site and PFT-group basis to assess the systematic error and accuracy of ET and GPP simulations.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

2. Investigate how the models represent the observed ET and GPP for different PFTs regarding their sub-annual averaged and standard deviations.
3. Evaluate the simulated, PFT-level ET and GPP statistical distributions and their moments (mean, variance, skewness, and kurtosis) to contextualize assessments of factors, like droughts, which impact the shape of these distributions.
4. Compare the inter-site ET and GPP differences of PFT-grouped stations to estimate how different PFT-specific ET and GPP time series are in the models and the observations.

Thus, these findings offer critical information for comparisons of GPP and ET from the evaluated models. Furthermore, this study also paves the way for a better-informed analysis of the drought response of ET and GPP from the models being assessed over Europe. We expect that:

1. There is a lower systematic bias, and the simulation is closer to the observations by the PFT-scale than the grid-scale CLM5 outputs, remote sensing, and reanalysis data.
2. The remotely sensed and modeled data provide good approximations of the ICOS ET and GPP phenologies, but there are apparent differences in this ability between PFTs.
3. The remotely sensed and modeled data show a lower range of variability of ET and GPP values within and across the PFT groups than the ICOS measurements.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

2. Methods and Data

2.1. Community Land Model version 5

We use the CLM5 (Lawrence et al., 2018, 2019), forced offline with custom input data. The land surface of a region in CLM5 is first disaggregated into grid cells, which are uniformly distributed and simulated individually. These grid cells are tiled into land units (i.e., natural vegetation, crops, lakes, urban areas, and glaciers) with a relative area coverage within the grid cell. Importantly, plants in the naturally vegetated land units compete for water in a single soil column. The vegetation is grouped into PFTs (Lawrence and Chase, 2007), which are distinguished through leaf habit (evergreen or deciduous), morphology (needle and broad leaves, grass and shrubs), and the bioclimate of the grid cell location (boreal, temperate, and tropical). Here, we use CLM5-BGC, which calculates vertical carbon and nitrogen pools and fluxes between the vegetation, soil, and atmosphere. In the following subsections, we briefly describe the essential processes in CLM5 that are particularly relevant to this study and the input data and leading features of the European CLM5 setup.

2.1.1. Gross primary production and evapotranspiration

The stomatal conductance of plants (g_s) couples water exchange with carbon uptake between vegetation and the atmosphere. In CLM5, g_s is calculated by the Medlyn stomatal conductance model (Medlyn et al., 2011):

$$g_s = g_0 + 1.6\left(1 + \frac{g_1}{\sqrt{D}}\right)\frac{A}{c_s} \tag{1}$$

Where g_0 is the Medlyn intercept and defaults to 100 μmol m^{-2} s^{-1}, and g_1 is the Medlyn slope, a PFT-specific parameter. D is the vapor pressure deficit indicating atmospheric water demand, and c_s is the CO_2 partial pressure at the leaf surface relative to the total atmospheric pressure. A is the carbon assimilated through photosynthesis.

$$A = \frac{c_s - c_i}{1.6 r_s} \tag{2}$$

The calculation of A is adapted from Bonan et al. (2011) and is based on the Farquhar model (Farquhar et al., 1980) and limited by photosynthetic capacity given by the LUNA model (Ali et al., 2016). It requires knowledge of the gradient of CO_2 concentration from the outside to the inside of the leaf and neglects CO_2 storage at the leaf

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

surface. c_s and c_i are the leaf surface and internal partial CO_2 pressures, and r_s is the stomatal resistance, which is the inverse of g_s. Further, c_s and c_i are calculated.

$$c_s = c_a - 1.4 r_b A \tag{3}$$

$$c_i = c_a - (1.4 r_b + 1.6 r_s) A \tag{4}$$

The factor 1.4 refers to the diffusivity ratio between CO_2 and H_2O gases in the leaf boundary, and 1.6 is the same ratio in the stomata. The equations for A, g_s, c_i, and c_s are computed iteratively until c_i converges, using a hybrid algorithm with the secant method and Brent's method (Lawrence et al., 2018). The photosynthesis is scaled to the canopy GPP by considering the effect of sunlit to shaded area ratios of the total leaf area.

The water input from the atmosphere to the land surface can be snow accumulating on the ground, streamflow, lake water, intercepted by the vegetation canopy, or can infiltrate the ground. The water in the ground percolates through 20 soil layers and is stored, directly evaporated, or taken up by plant roots relative to their transpiration demand. Hydraulic stress in a plant is calculated in a hydraulic framework using Darcy's law for transient porous media flow (Bonan et al., 2014).

The transpiration flux T is calculated with the resulting r_s from above.

$$T = \frac{e_s - e_i}{r_s} \tag{5}$$

e_s is the H_2O vapor pressure at the leaf surface, and e_i is the saturation H_2O vapor pressure resulting from the leaf temperature.

The total evapotranspiration is then determined by summing the transpiration and the evaporation from vegetation interception, surface water, the ground, and potentially snow.

2.1.2. Setup of the European CLM5

The European CORDEX (Giorgi et al., 2009) domain delimited the extent of this study, matching with the extent of regional atmospheric models. With a resolution of 3 km (0.0275°), our grid contains 1544 × 1592 grid cells, including the ocean. We used stand-alone CLM5 with the activated BGC module and stub models for ice, sea, and waves.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

The simulations were forced by the COSMO Reanalysis 6 (Bollmeyer et al., 2015; Wahl et al., 2017), a 6 km resolution data set providing meteorological variables over the European CORDEX domain from 1995 to 2019. The main advantage of using this reanalysis is the high resolution and a better representation of seasonal precipitation intensities compared to a coarser resolved global reanalysis (Bollmeyer et al., 2015). Using this forcing in high-resolution LSM simulations should lead to a more accurate simulation of sub-surface and surface hydrological fluxes, especially in regions with a relatively heterogeneous land surface (Wahl et al., 2017; Prein et al., 2016).

The static surface information was initialized for the year 2000 and was determined using input data from a standard repository (Lawrence et al., 2018). These data include land use information from (Hurtt et al., 2020), PFT distribution maps from (Lawrence and Chase, 2007), soil texture from (IGBP, 2000), and slope and elevation taken from (Earth Resources Observation And Science (EROS) Center, 2017).

The CLM5-BGC needs initial conditions for the carbon pools. For that, a spin-up workflow is necessary to bring the carbon pools and fluxes of carbon to a steady state before starting with production simulations. The spin-up method consists of two steps. Firstly, an accelerated decomposition simulation step, where carbon pools are artificially minimized. Secondly, a conventional simulation step, growing the carbon pools to the desired equilibrium state. During both spin-up steps, the atmospheric forcing from 1995 to 2012 was cycled (i.e., a cycling period of 18 years). The progress towards a steady state is monitored by assessing the difference in total carbon fixed in the ecosystem between a selected year within the last 18-year cycling period and the same year in the previous cycling period. $C_{tot,y}$ is the total ecosystem carbon (including vegetation and soil) in the year y, and $C_{tot,y-t}$ is the complete ecosystem carbon in the year y-t. A grid cell's carbon pools are in carbon equilibrium if the following is fulfilled.

$$\frac{\Delta C_{tot}}{t} < 1\, gC\, m^2\, year^{-1} \tag{6}$$

The following conditions define the final steady state on the continental scale.

1. 97% of the grid cells (and the total area) are in equilibrium.
2. The change in continental ecosystem carbon across the continent is lower than 2 Tg C year^{-1} for the three preceding cycle periods.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

The soil organic matter carbon pools in high northern latitudes were the slowest to reach equilibrium, which was reached after just about 1500 simulation years.

After the spin-up, we conducted a 24-year (1995 until 2018) transient simulation starting with the initial conditions established by the spin-up. We output the simulated variables from two model levels for the analyses.

1. **CLM5$_{\mathbf{PFT}}$**: This is the model's native resolution of vegetation-related states and fluxes calculation. Using output at this level (not the default configuration) allows for multiple time series per grid cell, each corresponding to a single PFT. This enables a selection of modeled data as needed. For instance, when comparing model data to ecosystem-level measurements, CLM5$_{PFT}$ relates to the simulated time series of the corresponding PFT, resulting in an adequate assessment of model functions. When comparing to in-situ observations, we will refer to CLM5$_{PFT}$ when we subset the ICOS site location and the agreeing PFT from the CLM5 data.
2. **CLM5$_{\mathbf{grid}}$**: The grid cell level output aggregates the PFT and the other tiles (i.e., croplands, urban areas, and lakes) that compose the grid cell area. Consequently, this data does not relate to a single functional type. Instead, it informs about the average state and fluxes in the grid cell area. In this study, CLM5$_{grid}$ designates CLM5 data extracted from the grid cell closest to the station's location.

2.2. Evaluation data

2.2.1. Station data

As ground truth data in the comparisons, we used the ICOS research infrastructure, which has a station observation network spanning 14 European countries (ICOS RI, 2021). Each station has at least one eddy covariance measurement tower and incorporates a processing workflow following a standardized protocol. We use the curated data, the WARM-WINTER-2020 data set (Warm Winter 2020 Team and ICOS Ecosystem Thematic Centre, 2022), which consists of homogenized variable time series following the ONEFLUX data pipeline (Pastorello et al., 2020). The ICOS WARM-WINTER-2020 data has measurements of 73 stations totaling over 800 station-years corresponding to multiple land cover types (see Figure 1 for a map with the station locations and Table S1 for more information). Note that the land cover type indicated by the ICOS site metadata and represented in the measurements refers to the footprint of the eddy covariance station. We omitted the stations over wetland and mixed forest land cover types to ensure a coherent analysis because no PFT counterpart

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

is implemented in $CLM5_{PFT}$. Also, shrub PFTs were not included in our analyses because there were insufficient shrubland sites in the ICOS data to support a robust evaluation. The analyses also excluded stations whose land cover type was not included in metadata sites (e.g., DEIMS-SDR https://deims.org). Because the land cover types from the selected sites correspond well with PFTs in CLM5, we will also refer to them as PFTs.

The processing workflow of the WARM-WINTER-2020 data extracts daily time series for GPP, partitioned from the net ecosystem exchange (NEE) using the night-time method and a dependence on a variable friction velocity threshold (in g C day^{-1}, *GPP_NT_VUT_REF*). We retained negative GPP values in these data, which stem from the uncertainty of the NEE measurements and partitioning method, to avoid introducing bias into the GPP distributions (Reichstein et al., 2012; Pastorello et al., 2020). For the ET evaluation, we also extracted the gap-filled latent heat flux (W m^{-2}, *LE_F_MDS*). Importantly, we verified our results by checking for inconsistencies in the analysis of ICOS NEE (*NEE_VUT_REF*), ecosystem respiration (*RECO_NT_VUT_REF*), and energy balance corrected latent heat flux (*LE_CORR*).

The conversion of latent heat (W m^{-2}) into ET (mm day^{-1}) is achieved by multiplying with the factor 0.035, assuming a constant enthalpy of vaporization decoupled from temperature because variable enthalpy has a negligible effect on the overall outcome of the conversion.

2.2.2. Remote sensing and reanalysis data

To consider the CLM5 in the context of additional complementary data products, we include GPP data from the Global Land Surface Satellite (GLASS, Liang et al. (2021)). The GLASS GPP product uses the Moderate Resolution Imaging Spectroradiometer (MODIS) and Advanced Very High-Resolution Radiometer (AVHRR) sensors and the revised Light Use Efficiency (LUE) model (Zheng et al., 2020) in 8-daily resolution in time and 0.05° resolution in space.

We also compare the CLM5 outputs with GLASS ET data, which applies a multi-model ensemble (e.g., MODIS-ET, remote sensing Penman-Monteith ET) to remote sensing information to estimate 8-daily latent heat on a 0.05° grid. We convert latent heat to ET as described in Section 2.2.1.

Lastly, we use ET reanalysis data for evaluation, which fuse observations and models. Namely, they are the European Center for Medium-range Weather Forecasts Reanalysis 5 - Land product (ERA5L, Copernicus Climate Change Service (2019)), which has a spatial resolution of 0.1° and hourly temporal resolution, and the

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

Global Land Evaporation Amsterdam Model (GLEAM version 3.5a, Martens et al. (2017)), which has a spatial resolution of 0.25° and daily temporal resolution.

2.3. Data processing

First, the remote sensing and reanalysis data are bilinearly remapped to the 3 km European CORDEX grid and interpolated to 8-daily means for 1995 - 2018. Then, we extracted the $CLM5_{grid}$, GLASS, ERA5L, and GLEAM data from the grid cell closest to the location of each selected ICOS station. Further, we select the time series in $CLM5_{PFT}$ that coincides with that grid cell and the station's PFT. Importantly, we focus only on the four predominant PFTs represented in the ICOS network: Evergreen Needleleaf Forest (ENF), Deciduous Broadleaf Forest (DBF), Grasslands (GRA), and Croplands (CRO), as outlined in Table 1.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

Table 1: The predominant PFT in the ICOS WARM-WINTER-2020 observation dataset, the number of corresponding sites, and the accordant PFTs in CLM5.

ICOS IGBP PFT	Number of Stations	Corresponding CLM5 PFTs
Evergreen needleleaf forest (ENF)	18	Needleleaf evergreen tree – temperate Needleleaf evergreen tree - boreal
Deciduous Broadleaf forest (DBF)	8	Broadleaf deciduous tree – tropical Broadleaf deciduous tree – temperate Broadleaf deciduous tree – boreal
Grasslands (GRA)	8	C_3 arctic grass C_3 grass C_4 grass
Croplands (CRO)	8	C_3 Unmanaged Rainfed Crop C_3 Unmanaged Irrigated Crop

The ICOS observations were also interpolated to 8-daily means, encompassing a time scale with significant variability of ecosystem processes (De Pue et al., 2023), to match the coarsest time resolution of other data-sets (i.e., GLASS remote sensing) and thus to facilitate comparison of processes at the same scale. For a consistent comparison, the analyses only account for time steps where valid values are present for all data sources. We evaluate the data for each variable over each station and groups of stations with the same PFT.

2.4. Analyses

2.4.1. Yearly evolution and statistical distributions

We calculate ET and GPP PFT-specific phenology (mean sub-annual dynamics), resulting in day-of-year (DOY) plots. This is done by averaging the same 8-daily time step across years for each site and calculating the mean and standard deviation of site-specific DOY belonging to one PFT.

Further, we determined the statistical distributions as probability density functions resulting from the Gaussian kernel density estimate (Scott, 1992). Subsequently, the distribution moments (mean, variance, skewness, and kurtosis) are calculated. The distributions and their moments are based on all 8-daily values corresponding to one PFT for each data source. The uncertainties of the distribution moments are calculated based on Harding et al. (2014).

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

2.4.2. Performance metrics

The percent bias (PBIAS) measures systematic model error and is calculated as follows.

$$PBIAS=\frac{\sum_{i=1}^{n} X_{sim,i}-X_{obs,i}}{\sum_{i=1}^{n} X_{obs,i}}\times 100 \tag{7}$$

Where n is the number of time steps, $X_{sim,i}$ is the simulated value of the variable X at the time i, and $X_{obs,i}$ is the observed value of the variable X at the time i. If the PBIAS for variable X is positive, the model overestimates; if negative, it underestimates the observed variable X. In our analysis, X_i is the interpolated 8-daily mean.

Furthermore, we estimated the root mean square error (RMSE) to indicate model accuracy and the root mean square difference (RMSD) to indicate similarity. RMSE and RMSD are calculated the same. However, the term 'error' assumes the truthfulness of the reference data. Hence, we use the RMSD when comparing data only between models.

$$RMSE=RMSD=\sqrt{\frac{\sum_{i=1}^{n}\left(X_{sim,i}-X_{obs,i}\right)^2}{n}} \tag{8}$$

A RMSE close to zero indicates that the model approximates the observations nicely. Similarly, a low RMSD reveals a high similarity between the two analyzed series. We calculate these metrics on a per-station basis and a set of stations belonging to the same PFT.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

3. Results

3.1. Land surface representation

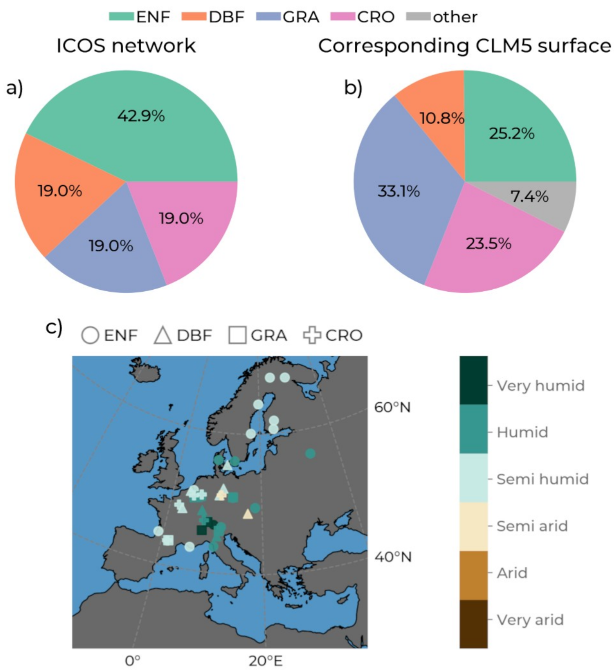


Figure 1: The share of represented plant functional types (color: Evergreen Needleleaf Forest (ENF), Deciduous Broadleaf Forest (DBF), Grasslands (GRA), and Croplands (CRO)) in a) the ICOS station network used in subsequent analyses and b) the corresponding grid cells in our European CLM5 setup. In c) is a map showing the locations of the ICOS stations, with the marker type indicating their PFT and the color of the marker indicating their hydro-climate (adapted from Jafari et al. (2018)) based on the mean annual precipitation from the COSMO-Reanalysis 6.

Before evaluating the GPP and ET variables from CLM5 and how they are compared with observations, we first assess if the PFT composition of the ICOS station network is comparable to the PFT composition in the respective cells selected in $CLM5_{grid}$. This is important, as GPP and ET magnitudes, variability, seasonality, drought responses, and trends strongly depend on the present vegetation type. In Figure 1 we observe that ENF, the PFT of almost half of the present ICOS stations, represents only around a quarter of the corresponding $CLM5_{grid}$ area. DBF also covers a smaller share of the area in those grid cells than in the ICOS station network. On the other hand, GRA and CRO are overrepresented in $CLM5_{grid}$ compared to the share of respective ICOS stations. Consequently, when comparing with the ICOS observations, the selected data from $CLM5_{grid}$ data are,

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

on average, overrepresenting the functionality of GRA and CRO and underrepresenting ENF and DBF, which hampers the evaluation of $CLM5_{grid}$ with in-situ ET and GPP. Hence, we also included the respective $CLM5_{PFT}$ GPP and ET in the subsequent analysis, enabling an accurate assessment of the functionality and relationships between PFT in the model. Additionally, we assess the similarities and differences between the two model scales, CLM_{grid} and CLM_{PFT}, and their approximation to the observations.

3.2. General model performance

This section presents model performance indices RMSE and PBIAS, comparing each model's ET and GPP estimates with measurements from the ICOS sites. We compared the RMSE and PBIAS on a per-site basis (Table S2 and Table S3), which yielded good results for most sites. The focus of this study, though, is the performance of PFT aggregations, combining data from sites that belong to the same PFT.

Table 2: The evapotranspiration (ET) root mean square error (RMSE) indicates the general model approximation and the percent bias (PBIAS), demonstrating systematic bias to the observations. Each value corresponds to a group of stations representing the same plant functional type (PFT; Evergreen Needleleaf Forest (ENF), Deciduous Broadleaf Forest (DBF), Grasslands (GRA), and Croplands (CRO)). The amount of data points (N) for each PFT is also indicated.

	PFT	N	$CLM5_{grid}$	$CLM5_{PFT}$	ERA5L	GLASS	GLEAM
RMSE	**ENF**	6784	0.71	0.72	0.83	0.83	0.67
	DBF	2302	0.55	0.61	0.72	0.69	0.56
	GRA	3745	0.65	0.86	0.59	0.57	0.59
	CRO	4647	0.7	0.99	0.86	0.84	0.61
	ø	4369.5	0.65	0.80	0.75	0.73	0.61
PBIAS	**ENF**	6784	-21.24	-16.15	20.31	12.57	14.14
	DBF	2302	-9.96	-0.41	43.57	29.02	15.67
	GRA	3745	-18.62	-13.55	3.51	2.38	2.08
	CRO	4647	-4.67	9.91	44.18	26.26	6.74
	ø	4369.5	-13.62	-5.05	27.89	17.56	9.66

In Table 2, we list the performance indices for ET and the number of 8-daily time-steps across the corresponding stations that went into their calculation. $CLM5_{PFT}$ has a higher RMSE and a lower PBIAS than $CLM5_{grid}$ for ET

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

across PFTs, except in CRO. Notably, the systematic bias in CLM5 is generally negative, with the same exception. On the other hand, ERA5L, GLASS, and GLEAM exhibit a general positive systematic bias for ET. ERA5L and GLASS show more significant deviations from the ICOS ET observations at ENF and DBF than $CLM5_{PFT}$ and CLM_{grid} but perform similarly at GRA and CRO. GLEAM has generally low RMSEs and performs best among the models simulating ET at ENF and CRO. The most considerable systematic ET biases are found for ERA5L at CRO and DBF sites, followed by GLASS for the same PFTs.

Table 3: The gross primary production (GPP) root mean square error (RMSE) indicates the general model approximation and the percent bias (PBIAS), demonstrating systematic bias to the observations. Each value corresponds to a group of stations representing the same plant functional type (PFT: Evergreen Needleleaf Forest (ENF), Deciduous Broadleaf Forest (DBF), Grasslands (GRA), and Croplands (CRO)). The amount of data points (N) for each PFT is also indicated.

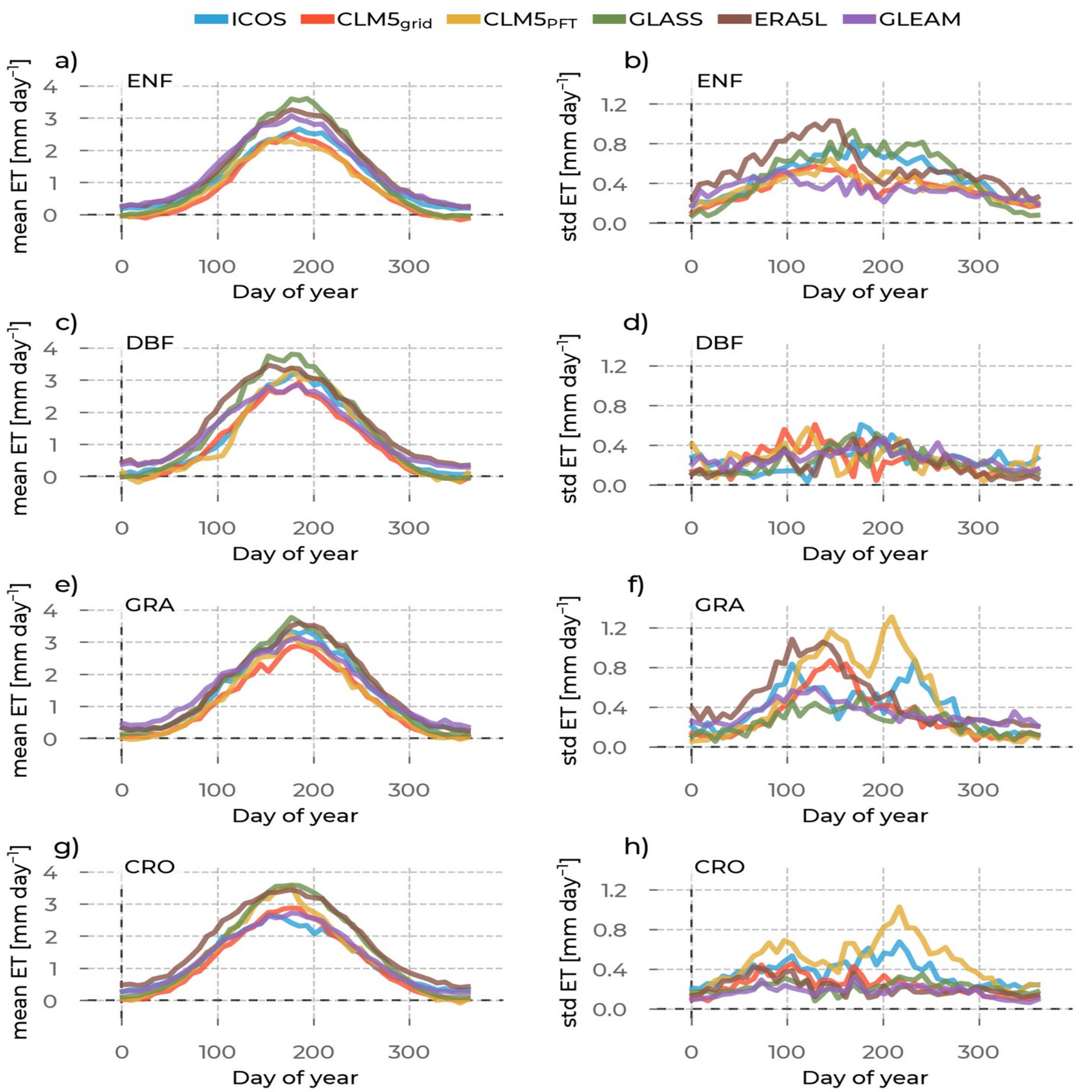
	PFT	**N**	**$CLM5_{grid}$**	**$CLM5_{PFT}$**	**GLASS**
RMSE	**ENF**	5976	2.25	2.44	1.75
	DBF	2473	3.71	3.35	2.81
	GRA	2838	3.14	3.01	2.63
	CRO	3607	3.85	4.21	3.55
	⌀	3723.5	3.24	3.25	2.69
PBIAS	**ENF**	5976	-26	-7.7	-14.53
	DBF	2473	-38.88	-43.76	-24.51
	GRA	2838	-30.73	-25.5	-21.34
	CRO	3607	-14.99	-1.48	-6.29
	⌀	3723.5	-27.65	-19.61	-16.67

Table 3 shows the performance indices for GPP. $CLM5_{PFT}$ performed better than $CLM5_{grid}$ in approximating the ICOS GPP observations at DBF and GRA sites. Conversely, $CLM5_{grid}$ is closer to the observations for ENF and CRO PFTs. The GLASS data show the lowest GPP RMSEs concerning ICOS measurements across all PFTs. All models approximated the ICOS GPP best at ENF, and the worst performance was at CRO sites. Furthermore, all models exhibit a negative, systematic bias in simulating the observed GPP across all PFTs. Especially at DBF and GRA PFTs, $CLM5_{grid}$, $CLM5_{PFT,}$ and GLASS show large systematic underestimations of the measurements. $CLM5_{PFT}$ has a low PBIAS related to the ICOS data for ENF and CRO sites.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

3.3. PFT phenology and its variability

3.3.1. ET



355

Figure 2: In the left column are the yearly evapotranspiration (ET) evolutions averaged across stations belonging to one plant functional type (rows). We differentiate the data source by color (ICOS observations: blue, $CLM5_{grid}$: red, $CLM5_{PFT}$: yellow, GLASS: green, ERA5L: brown, GLEAM: purple). The standard deviations across the sites are plotted in the right column to
360 **measure spread around this mean. Each row shows these plots for one plant functional type: Evergreen Needleleaf Forest (ENF), Deciduous Broadleaf Forest (DBF), Grasslands (GRA), and Croplands (CRO).**

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

This section describes the results of the investigation on the mean and the standard deviation of the yearly evolution of ET across PFTs and data sources (Figure 2). We will analyze the ET mean and standard deviation for each PFT sequentially. On average, the annual evolution of ET for $CLM5_{grid}$ and $CLM5_{PFT}$ compares well to the ICOS measurements. They also capture the observed seasonal transitions between low winter ET and high summer ET well.

However, the summer peak and the autumn decreasing period of ET at ENF sites from $CLM5_{grid}$ and $CLM5_{PFT}$ are simulated earlier than observed by ICOS (Figure 2 a). Furthermore, ET from $CLM5_{grid}$ and $CLM5_{PFT}$ underestimate the observations slightly throughout the year, including a lower summer peak and a lower minimum in winter. This underestimation results in a PBIAS of -21.2% for $CLM5_{grid}$ and -16.2% for $CLM5_{PFT}$ related to the ICOS observations. Oppositely, GLASS, ERA5L, and GLEAM ET values overestimated the measurements by ICOS in summer by a substantial margin. ERA5L and GLEAM overestimate the ET observations throughout the year. Meanwhile, GLASS ET was lower than ICOS during winter, totaling an overall PBIAS of +12.6%. $CLM5_{PFT}$ does not compare better to ICOS observations than $CLM5_{grid}$ (RMSEs of 0.71 and 0.72 mm day^{-1}). Still, $CLM5_{PFT}$ exhibits a lower summer peak and higher winter ET, equalling a more considerable underestimation of observations in summer and a minor underestimation of observations in winter. We also noted a lower standard deviation of $CLM5_{grid}$ and $CLM5_{PFT}$ across stations than in the ICOS measurements throughout the year (Figure 2 b). Interestingly, GLASS represents the ET variation across ENF sites better than CLM5, especially in summer. ERA5L, however, overestimated inter-site variability in the first half of the year, but this drops significantly around the 150th day of the year.

The observed timings of the ET summer peak and transition periods at DBF sites from ICOS are captured better by $CLM5_{PFT}$ than $CLM5_{grid}$ (Figure 2 c), and the magnitude of the total simulated ET is close to the observations there (PBIAS of -0.4% on the PFT level and -10.0% on the grid level). However, on average, $CLM5_{PFT}$ ET approximates the station observations worse than $CLM5_{grid}$ (RMSE of 0.61 versus 0.55 mm day^{-1}). Nevertheless, the average ET summer peak from $CLM5_{PFT}$ is very close to the ICOS data, while $CLM5_{grid}$ shows a smoother and lower peak. ET values from GLASS and ERA5L are more significant than the observations for most of the year. Meanwhile, GLEAM overestimates ICOS ET in spring, similar to ET from $CLM5_{grid}$ in summer. ET from the ICOS observations here at DBF sites show a reduced standard deviation during summer and a less pronounced seasonal cycle than at ENF sites (Figure 2 d). The annual evolution of ET standard deviations across DBF sites from $CLM5_{grid}$ and $CLM5_{PFT}$ is larger than the observations for the year's first half. GLASS better

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

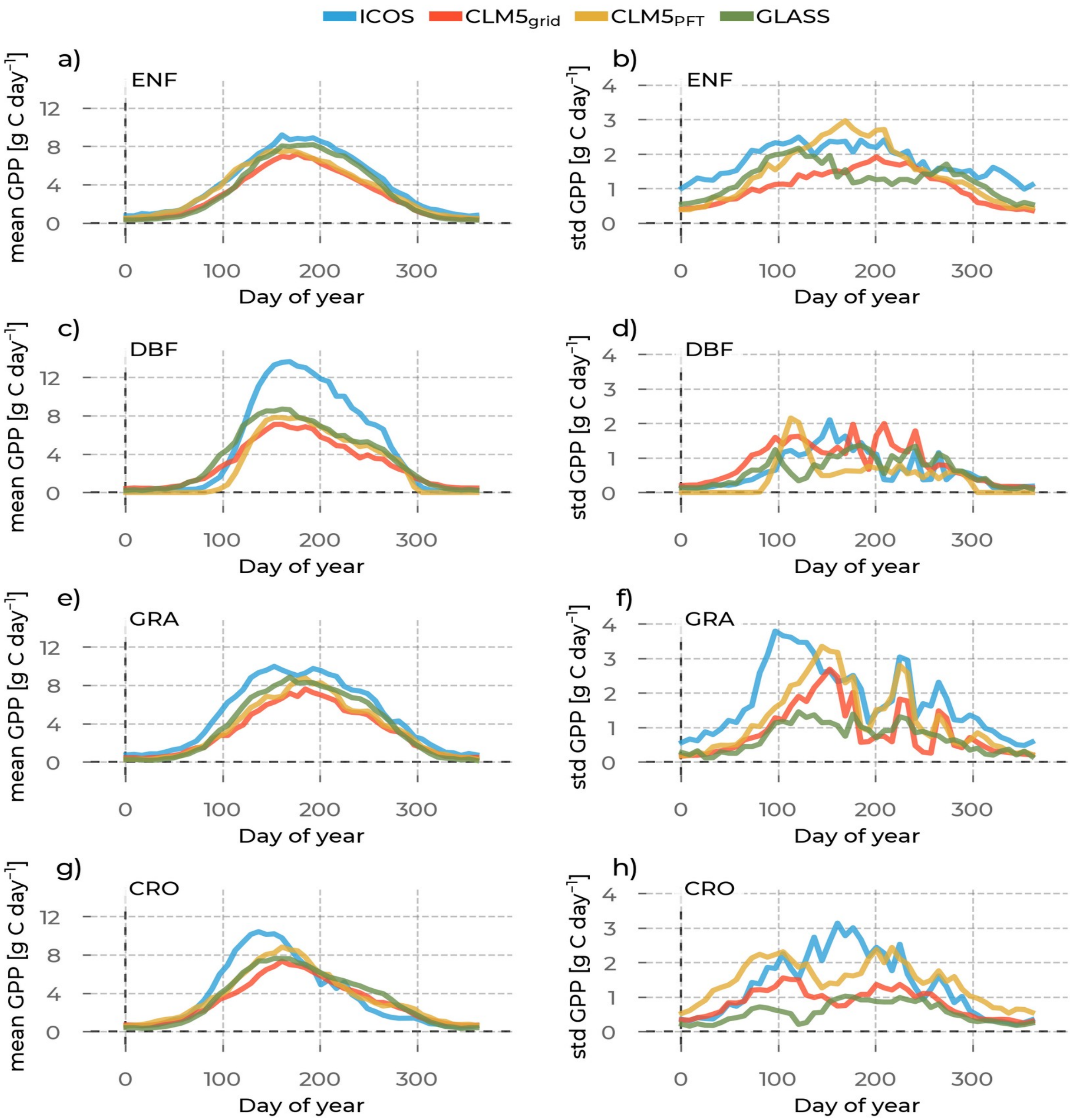
captures this site variability. Generally, the models reproduce the magnitude of the standard deviation across DBF sites better than across ENF sites.

While the timing of the ET summer peak and transition periods at GRA sites from $CLM5_{grid}$ and $CLM5_{PFT}$ represent ICOS observations well (Figure 2 e) CLM5 generally underestimates the ET observations (PBIAS of -18.6% for grid level and -13.6% for PFT level values). Oppositely, GLASS, ERA5L, and GLEAM overestimate the corresponding observations. Once again, $CLM5_{PFT}$ does not simulate ET closer to the observations than $CLM5_{grid}$, exhibited by the RMSEs of 0.86 and 0.65 mm day^{-1}. While the ET mean yearly evolution of $CLM5_{grid}$ and $CLM5_{PFT}$ is very similar, the PFT level annual standard deviation across sites indicates a much higher variance during the summer peak than the cell-level data and in-situ observations (Figure 2 f). The GLASS ET underestimates the in-situ observed standard deviation across sites throughout the year. Interestingly, $CLM5_{PFT}$ is the only model that captures the reduction of inter-site variability during the summer and the second inter-site variability peak in the year's second half in the ICOS data.

Finally, the ET summer peak at CRO sites from $CLM5_{grid}$ and $CLM5_{PFT}$ (Figure 2 g) is simulated slightly later, and the winter values are underestimated compared to ICOS measurements. Conversely, with the other analyzed PFT, CLM5grid is higher than the observations in summer at CRO sites, and $CLM5_{PFT}$ is even higher. Again, ET from GLEAM has a similar summer peak to $CLM5_{grid}$. Furthermore, GLASS and ERA5L are again higher than the observations during most of the year. $CLM5_{grid}$ underestimates observations with a PBIAS of -4.7%, and $CLM5_{PFT}$ overestimates them with +9.9%. $CLM5_{PFT}$ performs worse than $CLM5_{grid}$ in approximating the observations, with a higher RMSE of 0.99 compared to 0.70 mm day^{-1}. Moreover, the standard deviation across sites is more significant in $CLM5_{PFT}$ than in the observations, while for $CLM5_{grid,}$ it is lower than the observed (Figure 2 h). Interestingly, the variability across CRO sites in GLASS evolves through the year similarly to $CLM5_{grid}$.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

3.3.2. GPP



415 **Figure 3: In the left column are the yearly Gross Primary Production (GPP) evolutions averaged across stations belonging to one plant functional type (rows). We differentiate the data source by color (ICOS observations: blue, $CLM5_{grid}$: red, $CLM5_{PFT}$: yellow, GLASS: green). The standard deviations across the sites are plotted in the right column to measure the spread around this**

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

mean. Each row shows these plots for one plant functional type: Evergreen Needleleaf Forest (ENF), Deciduous Broadleaf Forest (DBF), Grasslands (GRA), and Croplands (CRO).

The GPP values of all PFTs show a summer peak and a low period in winter (Figure 3). The negative values present in the ICOS measurements are caused by the processing of the measurements by ICOS and are, therefore, not represented by CLM5 or GLASS.

$CLM5_{grid}$ underestimates the ICOS GPP at ENF sites throughout the year, and $CLM5_{PFT}$ underestimates them mostly in summer and autumn (Figure 3 a). The ET summer peak timing from $CLM5_{PFT}$ is earlier than that of $CLM5_{grid}$ and the observations. Consequently, The autumn transition period starts earlier in $CLM5_{PFT}$ than in the other data sources. Notably, GPP from GLASS approximates the high summer values and the autumn transition period better than CLM5. During summer, $CLM5_{PFT}$ shows a more significant standard deviation of GPP across ENF sites than $CLM5_{grid}$ (Figure 3 b). While this site variability in $CLM5_{PFT}$ compares better with the ICOS observations, its seasonality is more amplified in the model. Hence, the range between low winter and very high summer site variability of GPP in $CLM5_{PFT}$ is more extensive than in the observations. GLASS and $CLM5_{grid}$ GPP values show a generally lower diversity across sites throughout the year than ICOS, especially in summer. These characteristics result in a PBIAS of -26.0% by $CLM5_{grid}$ and -7.7% by $CLM5_{PFT}$. Further, the RMSEs are 2.25 and 2.44 g C day^{-1}, indicating a better approximation to the observations by $CLM5_{grid}$ than $CLM5_{PFT.}$

The most significant mismatch between ICOS GPP and the models at DBF sites is during summer: the average observed peak across sites is almost twice as prominent as the model peak (Figure 3 c). This results in a PBIAS of -38.9% for $CLM5_{grid}$ and -43.8% for $CLM5_{PFT}$ relating to ICOS data. While the GPP phenology of $CLM5_{PFT}$ captures the timing and the steepness of the ICOS reference in the transition period during spring, it peaks earlier and much lower than ICOS measurements, thereby substantially underestimating observations from spring to autumn. The timing of the GPP spring increase at DBF sites of GLASS is earlier than the observed one, and the transition is less steep. Further, the GLASS GPP summer peak is only slightly higher than in CLM5, therefore also underestimating ICOS observations strongly. $CLM5_{PFT}$ does somewhat better than $CLM5_{grid}$ in approximating ICOS GPP time series at DBF sites (RMSEs of 3.35 versus 3.71 g C day^{-1}). Interestingly, the standard deviation across DBF sites of the observations and the models is relatively small compared to ENF, GRA, and CRO throughout the year (Figure 3 d). However, the emergence of the peak standard deviation in the measurements during summer and gradually lower values during autumn and spring is not well represented in the models.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

On average, ICOS GPP at GRA sites is underestimated by $CLM5_{grid}$, $CLM5_{PFT}$, and GLASS throughout the year (Figure 3 e). The yearly evolution of the ICOS GPP measurements has steeper slopes in spring and autumn and peaks higher and earlier in summer than $CLM5_{grid}$ and $CLM5_{PFT}$. Furthermore, in the observations, the high values during summer are maintained high on a plateau with a slight negative slope until that slope becomes steeper in the transition period in autumn. At the same time, the models show a relatively pointed peak. The GRA GPP from $CLM5_{grid}$ and $CLM5_{PFT}$ perform similarly well in approximating the observations, with RMSEs of 3.14 and 3.01 g C day^{-1} for cell and PFT scale, respectively. They underestimate the observations from ICOS, evident in the PBIAS's -30.7% and -25.5% for cell and PFT scale, respectively. There is a high standard deviation across GRA sites in ICOS, especially in spring and autumn (Figure 3 f), which is well represented only by $CLM5_{PFT}$. The GRA inter-site GPP variability is lower in the GLASS data than in the ICOS observations throughout the year.

The most striking difference between the average yearly GPP evolution at CRO sites from $CLM5_{grid}$, $CLM5_{PFT}$, and ICOS is the shifted peak (Figure 3 g). Specifically, the ICOS observations show the peak around 50 days earlier and around 2.5 g C day^{-1} higher than $CLM5_{grid}$ and 1.25 g C day^{-1} higher than $CLM5_{PFT}$. The GPP slopes from ICOS in the spring and autumn transition periods are steeper than from $CLM5_{grid}$ and $CLM5_{PFT}$, but $CLM5_{grid}$ and GLASS accurately estimate the observed mean winter GPP. $CLM5_{grid}$ underestimates the in-situ GPP observations with a PBIAS of -15.0%, while $CLM5_{PFT}$ underestimates them with a PBIAS of -1.5%. Regarding modeling the observations accurately, $CLM5_{PFT}$ performs slightly worse than $CLM5_{grid}$ (RMSEs of 3.85 versus 4.21 g C day^{-1}). The phenology of the standard deviation across CRO sites from ICOS increases towards the summer peak and is low during winter (Figure 3 h). $CLM5_{grid}$ and $CLM5_{PFT}$, however, show a lower standard deviation across sites in summer.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

470 ### 3.4. Statistical distributions

3.4.1. ET

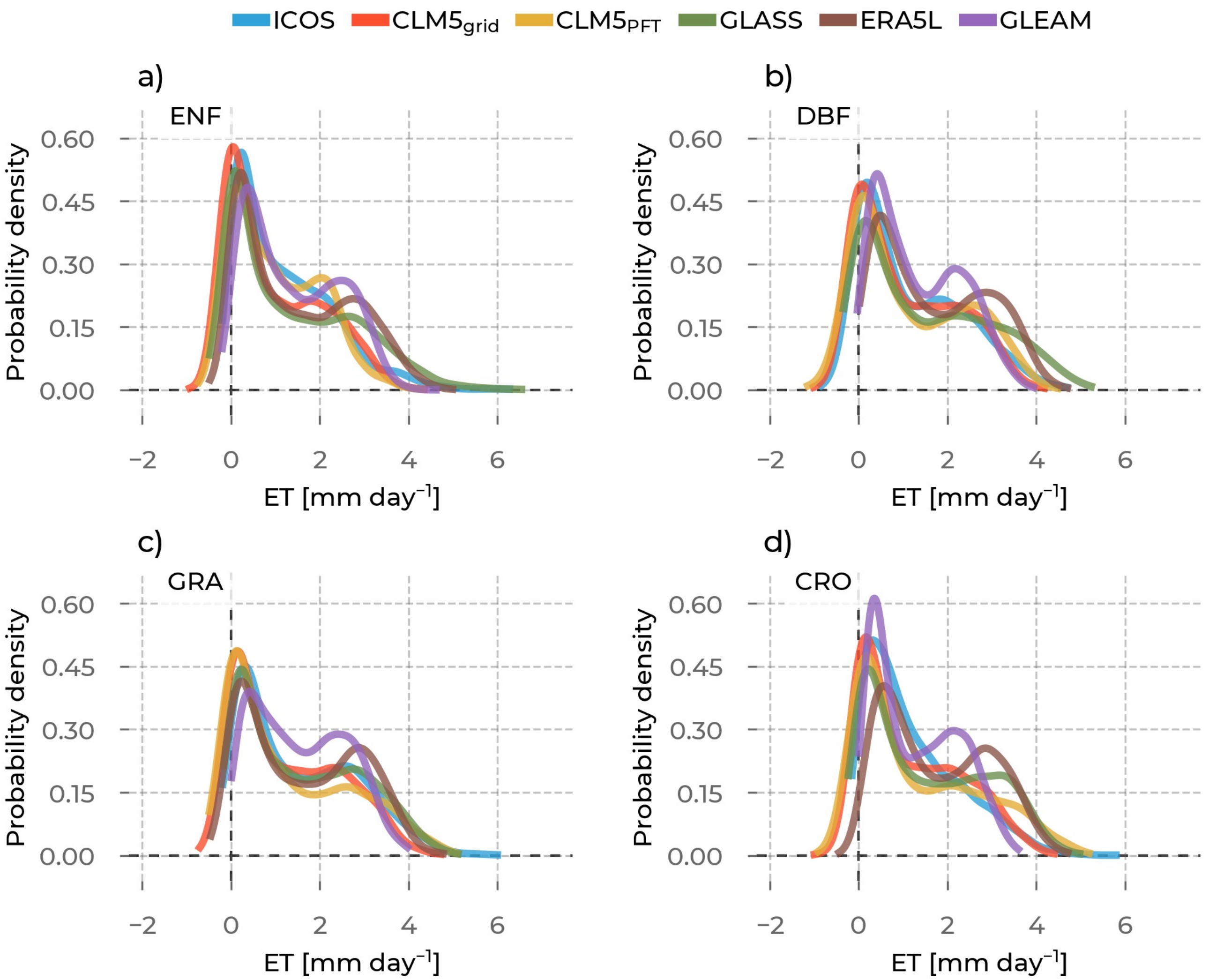


Figure 4: The probability density curves for all evapotranspiration (ET) values from stations belonging to the selected plant functional types are shown: Evergreen Needleleaf Forest (ENF),
475 **Deciduous Broadleaf Forest (DBF), Grasslands (GRA), and Croplands (CRO). The data source differs by color (ICOS observations: blue, CLM5$_{grid}$: red, CLM5$_{PFT}$: yellow, GLASS: green, ERA5L: brown, GLEAM: purple).**

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

In this section, we describe the results of the statistical distributions of ET in the model and the observations for each PFT. Then, we give more details on the moments of these distributions and how the models compare to the observations.

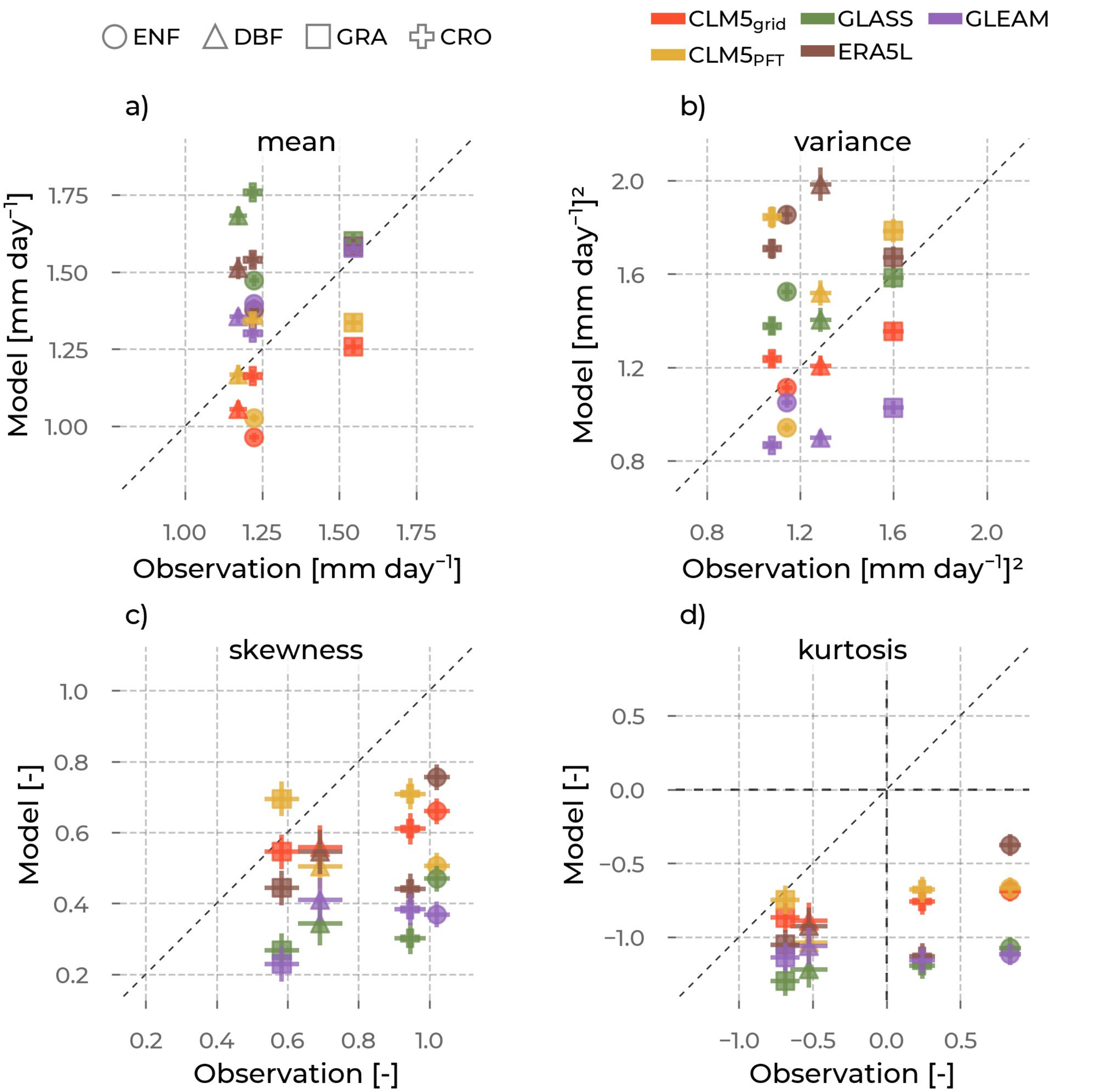
The ET probability density functions from ENF sites from $CLM5_{grid}$ and $CLM5_{PFT}$ (Figure 4 a) are generally shifted towards lower values compared to ICOS. Notably, the lack of high summer values >5 mm day^{-1} and overestimated probability of negative values is striking. Meanwhile, GLASS has a similar frequency of negative ET values as $CLM5_{grid}$ and $CLM5_{PFT}$. Still, it performs better in representing high summer values to the cost of under-representing the observed mid-ET value range of 1 - 3 mm day^{-1}. ET at ENF sites from GLEAM approximates the lower range nicely. However, it exhibits a second mode in the mid-high values, similar to ERA5L, GLASS, and $CLM5_{PFT}$, which is not present in the ICOS observations.

Again, there is a higher tendency to bimodality in the probability density of ET at DBF sites from $CLM5_{PFT}$, GLEAM, and ERA5L (Figure 4 b) than from $CLM5_{grid}$and the ICOS observations, exhibiting a second peak in the mid-high range values. In general, the distribution of DBF ET is very well represented by CLM5 on the grid and PFT scale.

The probability density curves of GRA ET (Figure 4 c) show that shallow values <0.5 mm day^{-1} that correspond to the low winter ET are more likely in $CLM5_{grid}$ and $CLM5_{PFT}$ than in the ICOS measurements. Additionally, the probability of higher values >2 mm day^{-1} from ICOS is underestimated by both CLM5 scales, and values >5 mm day^{-1} are not represented at all in CLM5 and GLASS. The tendency to bimodality of the ICOS ET distribution at GRA sites, showing a second peak at around 3 mm day^{-1}, is represented less pronouncedly in $CLM5_{PFT}$ and more pronouncedly in GLEAM and ERA5L.

Low ET values at CRO sites <0.5 mm day^{-1} and mid-range values from 2 to 4 mm day^{-1} have a higher frequency in $CLM5_{grid}$, $CLM5_{PFT}$, and GLASS than in-situ observations. On the other hand, modeled values from 0.5 to 2 mm day^{-1} occur at a lower frequency than in ICOS (Figure 4 d). Again, GLEAM and ERA5L exhibit a second peak of the distribution between 2 to 3 g C day^{-1}.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024



505 **Figure 5: The mean (a), variance (b), skewness (c), and kurtosis (d) of the evapotranspiration (ET) distributions from the models (color, y-axis), as opposed to the corresponding values from observations (x-axis) aggregated for each plant functional type (marker type): Evergreen Needleleaf Forest (ENF), Deciduous Broadleaf Forest (DBF), Grasslands (GRA), Croplands**

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

(CRO). The error bars are the standard errors of the respective moment, depending on the sample size.

In Figure 5 we show the moments of the ET distributions for each data source per PFT compared to ICOS observations. Ideally, the simulated versus observed distribution moments for a given PFT would lie on the 1:1 line. In that case, the ranking of the moments between PFTs was simulated well (e.g., ascending order of mean ET of PFTs). The uncertainties of the moments, measured by their standard error, are low and indicate that our results described below are robust.

Generally, $CLM5_{PFT}$ has higher means across PFTs than $CLM5_{grid}$ (Figure 5 a, yellow versus red markers). Consequently, except for CRO, the $CLM5_{PFT}$ ET means across PFTs, the values are closer to the observed ones. The ET mean across CRO sites from $CLM5_{grid}$ approximated the in-situ data well. On the other hand, the $CLM5_{PFT}$ mean ET overestimated the ICOS observations. The ET means across DBF and CRO sites from $CLM5_{grid}$, and $CLM5_{PFT}$ are close to the observed values, while the ones across ENF and GRA sites are off. This underestimation of the ICOS ET mean by $CLM5_{PFT}$ and $CLM5_{grid}$ by 0.2 mm day^{-1} also alters the observed mean ET ranking order between the PFT. GLASS, GLEAM, and ERA5L averages generally overestimate observations and show a similarly changed ranking as $CLM5_{PFT}$ and $CLM5_{grid}$. However, compared to $CLM5_{PFT}$ and $CLM5_{grid}$, the mean ET values at ENF and GRA sites from GLASS are closer to the observations, while DBF's and CRO's mean ET are overestimated.

The ranking of $CLM5_{grid}$ variance between PFT is represented nicely, but its simulated range is lower than in observations, i.e., 1.1 - 1.6 for ICOS and 1.1 - 1.4 mm day^{-1} for $CLM5_{grid}$ (Figure 5 b). This range is more extensive for $CLM5_{PFT}$because of the overvalued variances for CRO, DBF, and GRA, whose simulated variances are very close. The ET variances for each PFT from GLASS and ERA5L are higher than in the ICOS data, and their ranking order also differs. GLEAM underestimates the ET variance across all PFTs.

There is a notable improvement in the approximation of the observed magnitude of CRO skewness when going from $CLM5_{grid}$ to $CLM5_{PFT}$ (Figure 5c). Conversely, $CLM5_{PFT}$ overestimates ET averages at GRA sites and underestimates at DBF and ENF sites at larger magnitudes compared to $CLM5_{grid}$ relative to the ICOS data. $CLM5_{grid,}$GLEAM, and ERA5L distributions have a lower skewness than ICOS for all PFTs. Similar to the variance, here, the range of skewness in the models is also substantially lower among PFTs than in the ICOS measurements.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

535 The ET kurtoses across PFT from $CLM5_{PFT}$ are closer to the ICOS measurements than the ones from $CLM5_{grid}$ (Figure 5 d). However, the range of simulated ET kurtoses, like the variances and skewnesses across PFTs, is lower in all the models than in ICOS. Furthermore, all models show a generally lower kurtosis for all the PFT-specific ET distributions compared to ICOS. The ICOS observations show a leptokurtic ET distribution at CRO and ENF sites, while all models show platykurtic distributions for all PFTs.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

540 ### 3.4.2. GPP

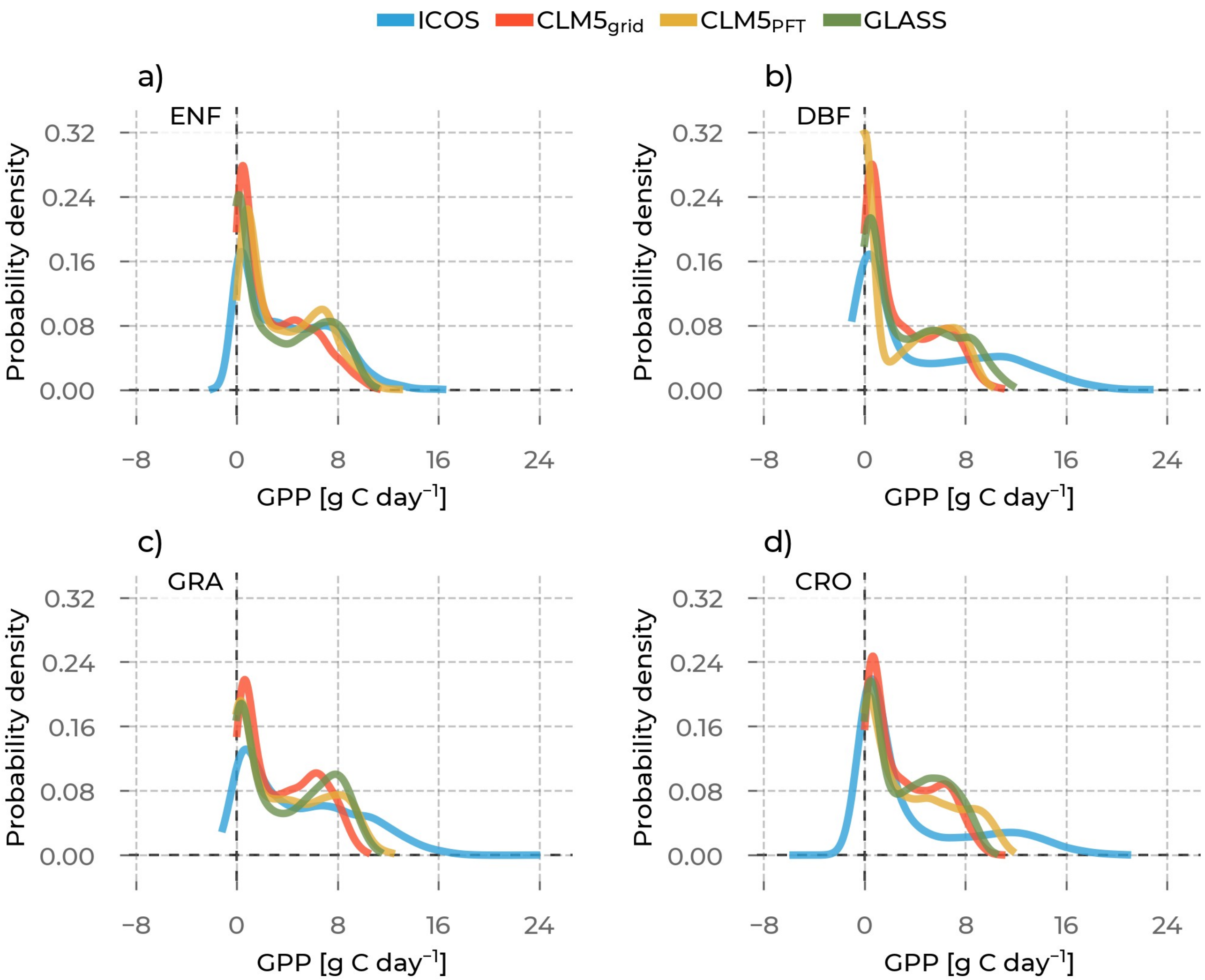


Figure 6: The probability density curves for all Gross Primary Production (GPP) values from stations belonging to the selected plant functional types are shown: Evergreen Needleleaf Forest (ENF), Deciduous Broadleaf Forest (DBF), Grasslands (GRA), Croplands (CRO). The data
545 **source differs by color (ICOS observations: blue, $CLM5_{grid}$: red, $CLM5_{PFT}$: yellow, GLASS: green).**

We continue to delineate the results of the same analyses for the GPP distributions and their moments.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

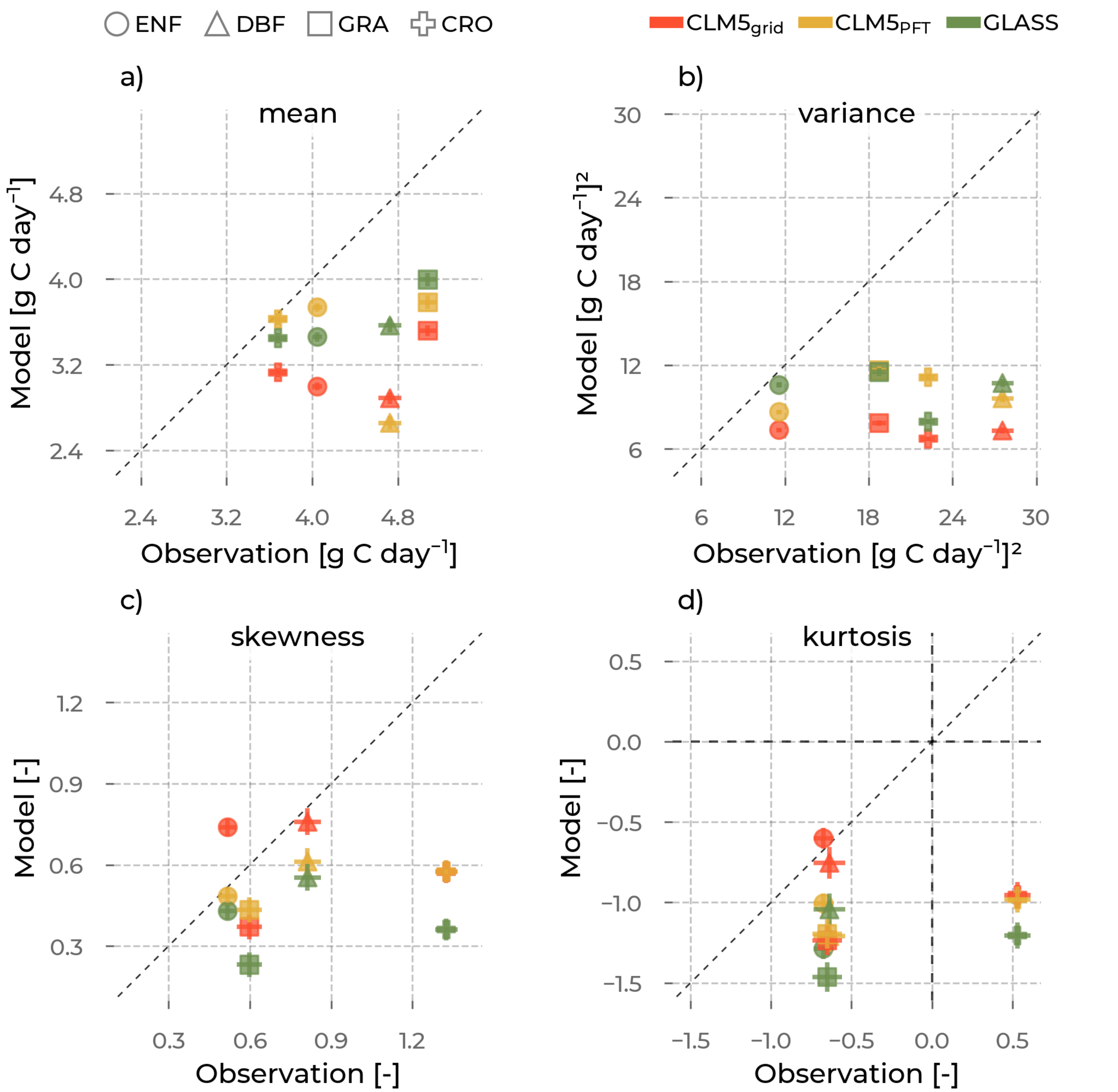
In ENF sites, the GPP of both $CLM5_{grid}$ and $CLM5_{PFT}$ overestimate the measured likelihood of low positive values of 0 - 2 g C day^{-1} (Figure 6 a). Both CLM5 scales underestimate the occurrence of GPP values >8 g C day^{-1} at ENF sites. Although $CLM5_{PFT}$ better represents the frequency of these observed high values, the exceptionally high observed values >12 g C day-1 are absent in the models. Notably, $CLM5_{PFT}$ shows a bimodal character that is also apparent in GLASS but not in $CLM5_{grid}$ and ICOS.

Low positive GPP values at DBF sites <1 g C day^{-1} occur more often in $CLM5_{grid}$, $CLM5_{PFT}$, and GLASS than in the ICOS network (Figure 6 b). $CLM5_{PFT}$ underestimates the observed likelihood of values from 2 to 3.5 g C day^{-1}, which are, in turn, overestimated by $CLM5_{grid}$. Further, GPP distributions of $CLM5_{grid}$ and $CLM5_{PFT}$ show a second mode between 4 and 8 g C day^{-1}, corresponding with the summer peak values. Another consequence of the much lower $CLM5_{grid}$ and $CLM5_{PFT}$ summer peak compared to ICOS is that the observed very high-value range from 12 to 22 g C day^{-1} is absent in the model data.

As found before for ENF and DBF, also the GPP at GRA sites from $CLM5_{grid}$ and $CLM5_{PFT}$ distributions are narrower than the one from ICOS, indicating a lower diversity of GPP values (Figure 6 c). This narrower distribution results in approximately half of the observed GPP value range not represented in the models. Therefore, the GPP peaks from $CLM5_{grid}$ and $CLM5_{PFT}$ distributions, located at 2 g C day^{-1}, are higher than observed by ICOS. However, the frequency of observed values between 4 and 7 g C day^{-1} is approximated nicely by both CLM5 scales and GLASS. The tendency to bimodality is evident in $CLM5_{grid}$ and GLASS and absent in $CLM5_{PFT}$ and the ICOS observations.

In contrast to the other PFT, the GPP distribution peak at CRO sites from ICOS aligns with the modeled ones in the low positive values (Figure 6 d). The frequency of mid-high range values >3 and <10 g C day^{-1} is overestimated in $CLM5_{grid}$, $CLM5_{PFT}$, and GLASS compared to the local ICOS observations. Finally, very high values >12 g C day^{-1} still occur in ICOS measurements but not in $CLM5_{grid}$, $CLM5_{PFT}$, or GLASS.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024



570

Figure 7: The mean (a), variance (b), skewness (c), and kurtosis (d) of the gross primary production (GPP) distributions from the models (color, y-axis), as opposed to the corresponding values from observations (x-axis) aggregated for each plant functional type (marker type): Evergreen Needleleaf Forest (ENF), Deciduous Broadleaf Forest (DBF), Grasslands (GRA),

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

Croplands (CRO). The error bars are the standard errors of the respective moment, depending on the sample size.

Analog to the ET analyses, we show the moments of the ET probability density functions per PFT for $CLM5_{PFT}$, $CLM5_{grid}$, and GLASS compared to ICOS observations in Figure 7. Generally, the GPP distribution moments show an underestimation of diversity in the model compared to the in-situ observations, exhibited by a smaller range of the PFT-related moments (Figure 7). In particular, the mean, variance, skewness, and kurtosis variation between the different PFTs is larger for ICOS data than for $CLM5_{PFT}$, $CLM5_{grid}$, and GLASS. The standard errors in calculations for the moments of GPP distributions are relatively small, so the confidence in these results is high.

GPP means from $CLM5_{grid}$ are around 3.0 - 3.5 g C day^{-1}, while the means of the observations range between 3.5 and 5.0 g C day^{-1} (Figure 7 a). Although this observed range of mean GPP per PFT is more accurately modeled by $CLM5_{PFT}$ (2.8 to 3.8 g C day^{-1}), it still underestimates the respective values from ICOS. Furthermore, the order between the GPP means from $CLM5_{PFT}$ differs from ICOS. For example, the mean GPP at DBF sites from $CLM5_{PFT}$ is the lowest among the PFTs but the second largest in the in-situ data. Notably, the GPP averages from GLASS exhibit a similar low range to $CLM5_{grid}$ but shifted to higher values.

The GPP variances from $CLM5_{grid}$ and $CLM5_{PFT}$ show an underestimation of the observed variability (Figure 7 b). Across DBF sites, the $CLM5_{grid}$ and $CLM5_{PFT}$ variances are much lower (7 and 9 g C day^{-1}) than in the ICOS observations (27 g C day^{-1}). However, ICOS GPP PFT-specific variances vary between 12 and 27 g C day^{-1}. Although the observed range and ranking order of GPP variances across PFT are better represented by $CLM5_{PFT}$than $CLM5_{grid}$those ranges, magnitudes of variance are always lower than observed. The range of variance in GLASS is similar to that in $CLM5_{PFT}$ and $CLM5_{grid}$, albeit at a higher variance magnitude.

The skewnesses of the GPP distributions in ICOS are well approximated by $CLM5_{grid}$ and $CLM5_{PFT}$ at ENF, DBF, and GRA sites (Figure 7 c). However, at CRO sites, ICOS skewnesses are again underestimated by $CLM5_{PFT}$ and $CLM5_{grid}$. Once more, the observed range of $CLM5_{grid}$ and $CLM5_{PFT}$ GPP skewnesses across PFTs is substantially lower than at the ICOS stations. The observed GPP skewness ascending order from ENF (lowest) to CRO (highest) is not represented well in the model data.

GPP from ENF, DBF, and GRA sites from $CLM5_{grid}$, $CLM5_{PFT}$, and GLASS agree with ICOS observations on the platykurtic nature of their distributions in ENF, DBF, and GRA. However, CRO is leptokurtic according to the

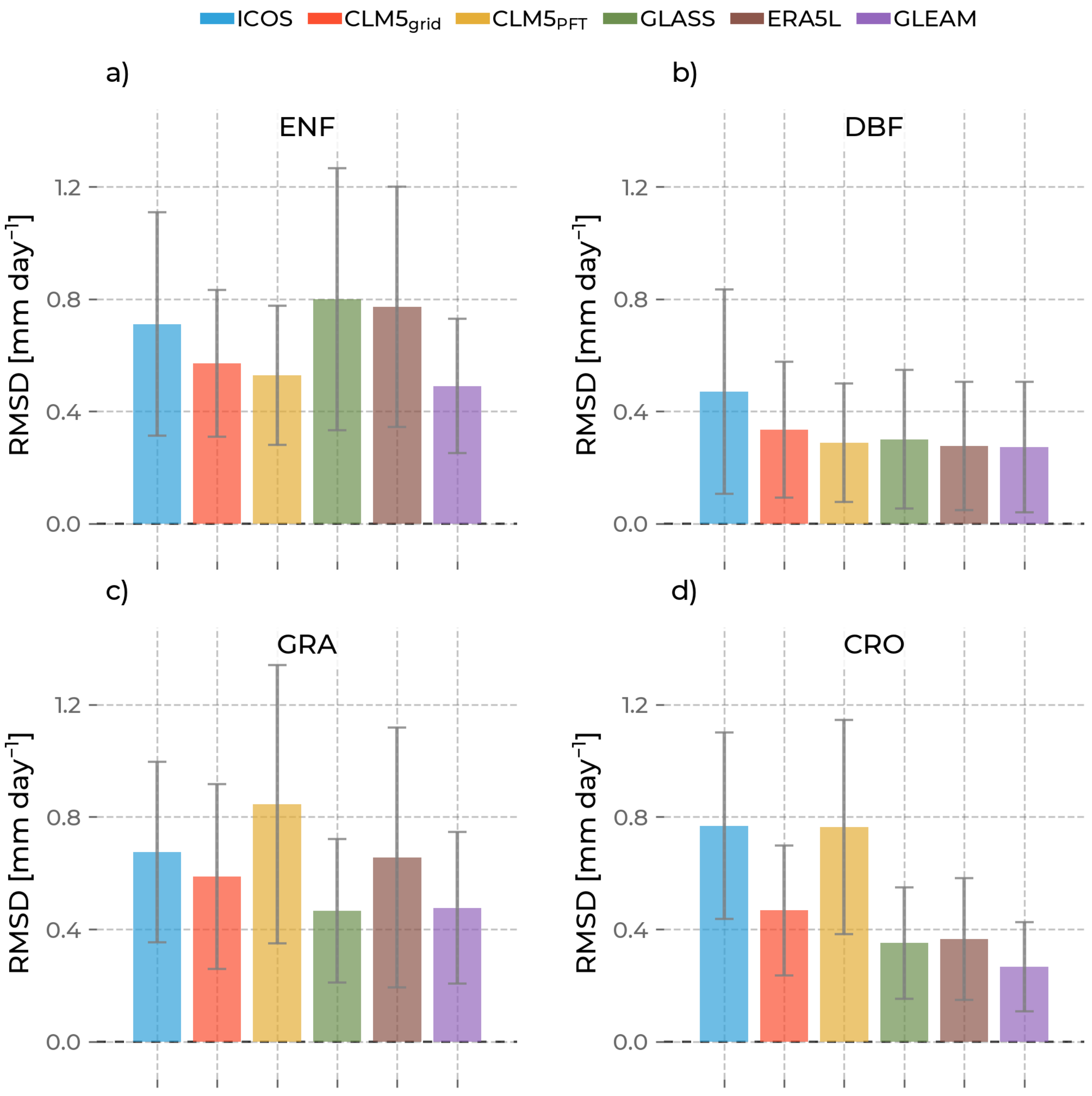
https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

ICOS observations but platykurtic for $CLM5_{PFT}$, $CLM5_{grid}$, and GLASS. Like the other moments of the GPP distributions, the spread of ICOS observation's kurtoses across PFTs is more significant than in $CLM5_{PFT}$, $CLM5_{grid}$, and GLASS. However, this is because of the remarkably different kurtosis of the CRO GPP to the other PFTs that show a very similar kurtosis of their distributions in ICOS around -0.6.

3.5. The inter-site similarity of PFT groups

In this section, we delineate the mean RMSD of each PFT per ET and GPP data sources. A low RMSD indicates that the stations corresponding to one PFT are similar, while a high RMSD hints at a great diversity within the PFT. By comparing the mean RMSD per PFT for ET and GPP across data sources, we can evaluate how much diversity is captured in the data of a particular PFT in the observations and models. The standard deviation of the RMSD for each PFT gives information on the spread of the inter-site RMSDs within the PFT group around that mean.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

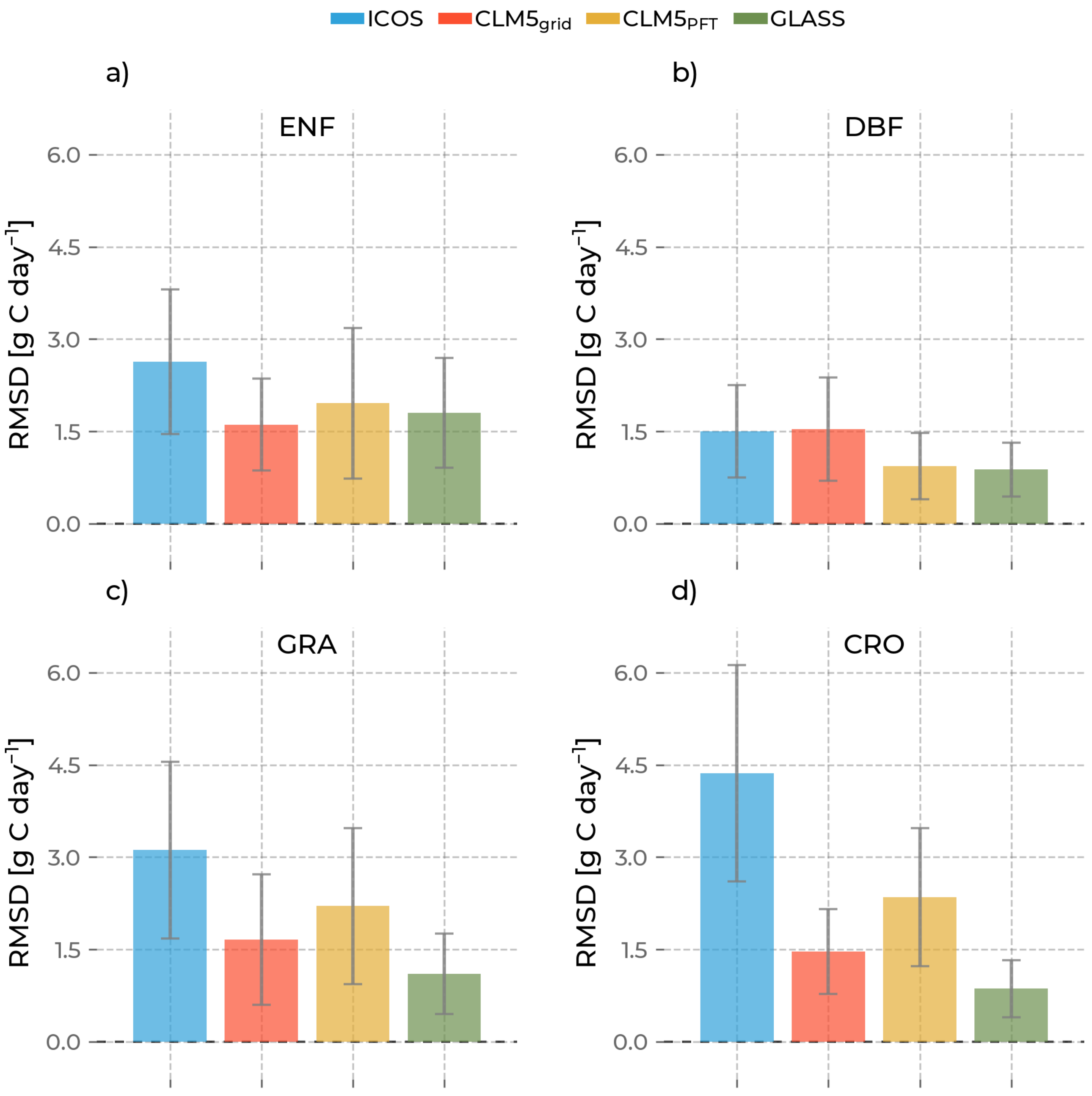


615 **Figure 8: The bars indicate the mean of the root mean square difference (RSMD) of evapotranspiration calculated for sites with the same plant functional type. The error bars are their standard deviation. Low values indicate high similarity between the sites, and high values show high dissimilarity. The color of the bars differentiates the data source.**

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

Figure 8 shows that $CLM5_{grid}$ has a lower difference in the ET time series between the corresponding sites for all PFT than ICOS. CLM_{PFT} has a lower mean RMSD than $CLM5_{grid}$ among ENF and DBF sites. Both $CLM5_{PFT}$ and $CLM5_{grid}$ underestimate the observed diversity of ET at ENF and DBF sites. Interestingly, the variation of ERA5L and GLASS ET time series for ENF is higher than observed, and they also show the most significant variation of RMSD. Meanwhile, DBF's mean RMSD of all models is lower than that of ICOS. $CLM5_{PFT}$ shows a higher diversity of ET between GRA sites and CRO sites than $CLM5_{grid}$. The $CLM5_{PFT}$ surpasses the observed mean RMSD for the GRA PFT. All other models underestimate it slightly ($CLM5_{grid}$, ERA5L) or more pronouncedly (GLASS, GLEAM). Particularly at CRO sites, the ET RMSD of $CLM5_{PFT}$ is substantially higher than the other models and at a similar level as ICOS observations. In contrast, all other models show significantly lower mean RSMDs there. Generally, a higher ET RMSD mean in a PFT group comes with a higher spread (higher standard deviation) for all data sources. The RSMD in ET between stations is lower for $CLM5_{grid}$ and GLEAM than for ICOS for all PFTs.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024



635 **Figure 9: The bars indicate the mean of the root mean square difference (RSMD) of gross primary production calculated for sites with the same plant functional type. The error bars are their standard deviation. Low values indicate high similarity between the sites, and high values show high dissimilarity. The color of the bars differentiates the data source.**

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

Figure 9 shows that for GPP, the models generally have a lower mean RSMD than ICOS across stations for all PFT, except for CLM5$_{grid}$ at DBF. CLM5$_{PFT}$ has a more diversely simulated ET across ENF, GRA, and CRO sites than CLM5$_{grid}$. Interestingly, the observed magnitude of the RMSD is lowest for DBF and highest for CRO and has a more extensive range across PFTs than the models. For example, the RMSDs of ICOS data differ by approximately 1.3 g C day^{-1} between GRA and CRO, while CLM$_{grid}$, CLM5$_{PFT}$, and GLASS indicate similar RSMDs for those PFTs. Especially CLM5$_{grid}$ shows a constant within-PFT variability of around 1.5 g C day^{-1} independent of the PFT. Higher mean GPP RMSD values also come with a higher standard deviation.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

4. Discussion

Our results show that $CLM5_{grid}$ and $CLM5_{PFT}$ approximate the ET observations from ICOS better than GLASS remote sensing and ERA5L reanalysis but worse than GLEAM reanalysis. Moreover, especially for $CLM5_{PFT}$ the systematic error in simulating ET is lower than all other evaluated data sets. For GPP, we found that $CLM5_{grid}$ and $CLM5_{PFT}$ performed worse than GLASS data, indicated by a larger PBIAS and larger RMSE. Surprisingly, $CLM5_{PFT}$ generally had a higher RMSE than CLM_{grid} but, at the same time, a lower PBIAS. Averaged ET and GPP phenologies were relatively well simulated but exhibited underestimations across all PFT, especially in DBF, compared to ICOS measurements. $CLM5_{PFT}$ better captured the PFT-specific mean and standard deviation of the ET and GPP annual dynamics than $CLM5_{grid}$, the reanalyses, and remote sensing data. The GPP and ET distributions analysis showed underestimations of their observed variability for all models, $CLM5_{grid}$, $CLM5_{PFT}$, GLASS, ERA5L, and GLEAM. Lastly, we found that for most PFTs, the modeled and remotely sensed data was too similar between stations of the same PFT group compared to the ICOS observations.

4.1. Uncertainty

4.1.1. Observations

Notably, the EC measurements carry uncertainties that might affect the results of this study, especially related to the systematic errors in the simulations. For instance, EC measurements neglect the energy from large eddies. To check for possible inconsistencies, we evaluated the energy balance corrected ET (ET_{corr}) from the ICOS sites (Pastorello et al., 2020). This methodology assumes a constant Bowen ratio to close the energy imbalance. Simulated ET underestimates ET_{corr} to a greater degree than the non-corrected ICOS ET (Figure S1, Figure S2), therefore suggesting a higher systematic error than in the analysis of non-corrected ET. Besides that, we discovered the same patterns with the corrected ET, concluding that the energy balance error did not introduce significant bias to our results and the interpretations. Furthermore, GPP is not directly measured but partitioned from NEE. The NEE partitioning method has an underlying uncertainty stemming from potentially unfulfilled assumptions that propagate to the GPP and ER variables in the ICOS data. So, we also ensured that our results remained consistent by evaluating the non-partitioned NEE and the ER variables (Figure S3, Figure S4, Figure S5, Figure S6). We discovered a substantial underestimation and missing variability in NEE and ER across PFTs in CLM5, confirming the systematic underestimation in our analysis of GPP. While we believe that our analyses have followed meticulous approaches to ensure robust results by applying the ICOS quality flags and comparing these additional variables, many studies still emphasized the biases arising from a shifting footprint with varying wind direction and wind speed and the energy balance correction method assuming a constant Bowen ratio (Jung et al., 2020; Eshonkulov et al., 2019; Chu et al., 2021). Therefore, we encourage the development and use of novel and more accurate energy balance closure methods (Zhang et al., 2024). Furthermore, dropping bad-quality gap-fill data from the ET and GPP time series might introduce a bias that underrepresents periods of low friction velocity and atmospheric inversion conditions. Lastly, based on the geographical distribution of the ICOS station network, the results might misrepresent Southern and Eastern Europe and semi-arid and arid hydro-climates (Figure 1). Those factors might have influenced the diversity of ET and GPP values and the ranges of their distributions.

4.1.2. Forcing

Importantly, discrepancies between the COSMO Reanalysis used to force the European CLM5 and the station observations might introduce deviation into our analyses that could hamper interpretations of our results regarding the model functionality. While the high-resolution forcing data already includes information from observations through data assimilation, particular locations and conditions might be less well represented than others, and a resulting bias in the meteorological variables would propagate to the simulation of ET and GPP. However, data assimilation approaches minimize the systematic error of the atmospheric model to the observations. Furthermore, the probability and potential influence of including a bias from the forcing of a single location is lowered by considering multiple sites in the performance and statistics of the PFTs. Nevertheless, we assessed the meteorological variables from the COSMO Reanalysis 6 (temperature, shortwave incoming radiation, precipitation, relative humidity) with the ICOS station data to scrutinize potential errors arising from the forcing. We used the same approach as for the GPP and ET evaluation (Figure S7 – Figure S14). We discovered that the forcing variables' average yearly dynamics and distributions represent the ICOS observations well. More minor yet notable misrepresentations include underestimations of shortwave downward radiation and precipitation in summer and relative humidity over GRA and CRO sites throughout the year compared to the measurements. This could explain some of our analyses' ET and GPP underestimations by CLM5. Notably, the mean and variance across the PFTs and their ranking are represented reasonably well for all forcing variables, as opposed to our results with GPP and ET. Furthermore, the skewness and kurtosis of the forcing temperature and shortwave downward radiation compare well to the ones from ICOS, indicating well-matching distributions between the COSMO Reanalysis 6 and the observations. However, in particular, the higher-degree moments of

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

the distribution are not well simulated for precipitation and relative humidity. These characteristics of the distributions affect the CLM5 simulations of GPP and ET and might have influenced our results. Further considerations, including ensemble simulations with perturbed forcings, are required to fully capture the uncertainty introduced into CLM5, but this is beyond the scope of this study.

4.1.3. Static information and initial conditions

The static surface information, including the soil texture, elevation, aspect, land unit, and PFT distributions, affect the simulation of ET and GPP in CLM5. The soil texture composition will define how water is stored and conducted in the soil, contributing to the evaporation from the soil, an essential ET component. Further, the soil texture will influence root water uptake if vegetation is present in the soil column, indirectly impacting plants' transpiration, another critical ET component. Further, ET is regulated by the available energy, which is determined by how the canopy, the elevation, and the aspect of that location influence the incoming radiation. Especially the diversity between these input variables across the locations of the ICOS stations might have played an essential role in the simulation of the PFT-specific ET and GPP distributions.

Lastly, particularly for $CLM5_{grid}$, GLASS, GLEAM, and ERA5L, the distribution of PFTs across the domain and in the grid cells corresponding to the ICOS stations define the equations and parameters that will be used for the calculation of ET and GPP. Consequently, if the grid cells corresponding to ICOS stations are dominated by PFTs that do not comply with the stations' footprints, the simulations of specific PFTs in the model are negatively affected. Importantly, this does not apply to the $CLM5_{PFT}$ because we could select the data that belongs to the adequate PFT. Therefore, interpretations of our results relating directly to vegetation functions implemented in CLM5 are here primarily focusing on the $CLM5_{PFT}$ data.

The initial conditions of the carbon cycle, most notably the size of the soil and vegetation carbon pools, are another source of uncertainty. Essentially, our spin-up and production simulations were restricted to the years where the high-resolution forcing was available (1995 - 2018). The spin-up simulations, therefore, recycle atmospheric forcings for a substantial period, which we also used in the production simulations. Hence, the production simulations adopted the equilibrium state (incoming carbon equals outgoing carbon) required to conclude the spin-up. However, in natural conditions, there was no carbon equilibrium in the simulated years. Instead, the carbon cycle experiences dynamic changes, such as long-term trends resulting from changing environmental conditions. Many European ecosystems exhibited a net carbon uptake, thus acting as a carbon sink

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

(Pilli et al., 2017; Winkler et al., 2023), measured in ICOS accordingly. The negative long-term mean NEE indicates carbon sources, which is evident across all PFTs in the EC observations (Figure S4 a). On the other hand, the simulations show a NEE close to zero for all PFTs, directly showing the effect of the equilibrium state of the land surface in the model. The results of DBF, which is the most significant carbon sink in the ICOS data and simultaneously shows the largest GPP underestimations by CLM5, underline a potentially important role of the carbon equilibrium in our results. Future work will conduct a more comprehensive spin-up under conditions closer to a real-world carbon equilibrium (the 1950s or earlier) and a transition run before the production simulations to capture the dynamic trends of the land surface processes. Possibly, the bias in the EC measurements towards conditions with low friction velocity and atmospheric inversion might also cause overestimations of GPP and the resulting carbon sink in ICOS.

4.2. PFT-specific evaluation

While $CLM5_{PFT}$ showed a smaller systematic error than CLM_{grid} for most PFT compared to the observations (lower absolute PBIAS), the ability to approximate the observation time series is worse (higher RMSE). A shifting sign in the bias of the CLM5PFT data explains these counterintuitive results. The presence of both positive and negative bias (in time and across stations) cancels out and yields an overall low PBIAS. In summary, we find in the evaluation that the ET time series of CLM_{PFT} are not closer to observations than $CLM5_{grid}$ for any PFT, but $CLM5_{PFT}$ generally approximates the ET sum over time better than $CLM5_{grid}$ for ENF, DBF, and CRO. However, it is also clear that, on average, the phenology of $CLM5_{PFT}$ is closer to the observed than $CLM5_{grid}$, for instance, for both ET and GPP at DBF and GRA sites. Importantly, critical PFT-specific characteristics, like DBF's steep spring GPP increase, are only captured by $CLM5_{PFT}$ and the inter-site variability of ET and GPP throughout a standard year. This discrepancy between the evaluation metrics and the vegetation phenology suggests that $CLM5_{PFT}$ could better capture the PFT-specific variability that ICOS observes. However, this variability is modeled in a way that did not contribute to a low RMSE, for instance, shifted in time or space, so the averaged PFT-specific comparisons (the phenology and the distribution moments) compare better with ICOS than CLM5grid. Further evidence for this explanation is that $CLM5_{PFT}$ generally captures more variability (higher ET and GPP standard deviation across sites throughout the year for ENF, GRA, and CRO, and higher variance for each PFT). This ability to capture more variability than the other models, which is closer to the observed variability, can potentially improve the represented variability in $CLM5_{PFT}$ if the suitable variation can be modeled at the right time and location. This spatiotemporal discrepancy of simulated and observed GPP and ET

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

variability could potentially be solved with optimized PFT parameters (Baker et al., 2022; Birch et al., 2021; Cheng et al., 2021; Dagon et al., 2020; Deng et al., 2021; Fisher et al., 2019b).

Several past studies also indicated the underestimation of ET and GPP in CLM5 compared to observations (Boas et al., 2023; Strebel et al., 2023; Cheng et al., 2021; Birch et al., 2021), which we confirm in this study. Parameter improvements could also alleviate these general underestimations of GPP and ET across PFTs, especially during summer (Dagon et al., 2020). However, optimal parameters might vary from site to site (Lin et al., 2015) even if they have the same PFT. Thus, CLM5, and more generally, LSMs that implement plant traits as parameters on the PFT level, cannot capture this intrinsic PFT variability resulting from these traits. Albeit optimized parameters might still reduce the bias on the continental level, a more comprehensive approach to the spatiotemporal variability of plant traits might improve regional simulations drastically (Anderegg et al., 2022; Van Bodegom et al., 2014; Kattge et al., 2011).

4.3. Inter-site similarity of PFT groups

For all models (CLM5$_{grid}$, CLM5$_{PFT}$, ERA5L, GLASS, GLEAM), the distributions of ET and GPP across PFTs are very similar, which is not the case for the observations. This is especially true for variances but also notable for the means, skewnesses, and kurtoses. We expected CLM5$_{PFT}$ to show more significant variability than CLM5$_{grid}$ and the other grid-scale models because the aggregated, mixed PFT data of the grid cell would homogenize the variables and cancel out some of the variability. While CLM5$_{PFT}$shows a more extensive range of variation of ET and GPP across PFTs than CLM5$_{grid}$, ERA5L, GLASS, and GLEAM, it still vastly underestimates the observed range of variance by ICOS, especially for GPP (Figure 5, Figure 7).

The mean RMSD across sites of the same PFT indicates that ET across sites can be as different in CLM$_{PFT}$ for GRA and CRO as in the observations. However, the ET differences across sites with the same PFT were underestimated at ENF and DBF. GPP differences across sites with the same PFTs were underestimated for all PFTs. This suggests the missed variance could mainly stem from missed PFT internal inter-site differences or unresolved differences in site-specific abiotic conditions (e.g., soil depth and texture). Possibly, this could not be improved through optimization of PFT-specific parameters, as these sites would still share the same set of parameters. An enhanced concept of functional types in vegetation, focusing on the spatiotemporal variability of observed plant traits, could better facilitate improvements that raise the simulated ET and GPP variance in space and time.

4.4. Data requirements

As outlined above, beyond parameter optimizations, a comprehensive implementation of functional ecosystem diversity could significantly improve the LSM simulation outputs regarding multiple aspects of their distributions. This could introduce a state-of-the-art understanding of vegetation function into LSMs, which is essential to evaluate different theories of plan trait evolution and their effect on current and future energy, water, and carbon cycles.

In that light, we encourage sites to co-locate research infrastructures (Futter et al., 2023), like ICOS and the Integrated European Long-Term Ecosystem, critical zone, and socio-ecological Research Infrastructure (eLTER-RI). Thereby, sites cover additional observation spheres like biodiversity (e.g., functional diversity of plants) and socio-ecology (through forest and crop management and driving land use change) and establish a strong base for studies to increase the understanding of the whole system (Mirtl et al., 2018; Mirtl et al., 2021; Baatz et al., 2018). Further, this would promote large-scale observations needed to introduce more trait variability into LSMs. Lastly, combining LSMs and these holistic observations by data assimilation, going beyond decoupled modeling efforts (Bloom et al., 2020) and resulting in an ecosystem reanalysis (Baatz et al., 2021), would provide essential and specific data on the carbon cycle, which are currently unavailable.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

5. Conclusions

We evaluated the simulated evapotranspiration (ET) and gross primary production (GPP) from a 3 km resolved Community Land Model v5 (CLM5) set up over the European CORDEX domain. We differentiate the model outputs between the grid scale ($CLM5_{grid}$) and the plant functional type scale ($CLM5_{PFT}$) and compare them with ICOS station data as ground truth data. Furthermore, we compare with ET and GPP from remote sensing derived data from the Global Land Surface Satellite (GLASS) and reanalysis products such as the European Centre for Medium-Range Weather Forecast Reanalysis 5 - Land (ERA5L) and the Global Land Evaporation Amsterdam Model (GLEAM). $CLM5_{grid}$ and $CLM5_{PFT}$ exhibit promising skills in approximating the observations and often perform better than ERA5L, GLASS, and GLEAM. $CLM5_{PFT}$ showed a lower systematic bias (lower percent bias) but approximated the ICOS observations generally worse (larger root mean square error) than $CLM5_{grid}$ (Table 2, Table 3). ET and GPP are systematically underestimated for both model scales across all PFTs throughout the year. Especially during summer at DBF sites, GPP is substantially lower for $CLM5_{PFT}$ and $CLM5_{grid}$ than for ICOS observations (Figure 2, Figure 3).

Essentially, $CLM5_{PFT}$ and, to a greater degree, $CLM5_{grid}$, ERA5L, GLEAM, and GLASS show a lower spatiotemporal variability of ET and GPP than the measurements exhibited by a lower range of all the modeled ET and GPP distribution moments across PFTs than in ICOS. This smaller range and a lower root mean square difference between sites of one PFT group suggests that $CLM5_{grid}$, and more surprisingly, $CLM5_{PFT}$, simulate GPP and ET more similarly across PFTs than the ICOS measurements.

Further studies should investigate whether optimizing parameters in $CLM5_{PFT}$ with observation data increases the diversity of ET and GPP values or whether this is a structurally induced bias. This work provides essential insights for studies that aim to find optimized parameters and meaningful context for analyses of more specific ET and GPP dynamics using the evaluated data.

Availabilities

Code availability

A frozen version of the CLM5 version used here is stored here: https://doi.org/10.5281/zenodo.11091890. The case setup for the European 3 km simulation as well as a post-processing script is available under https://doi.org/10.5281/zenodo.11091845. Analysis, processing and plotting scripts and are available at https://doi.org/10.5281/zenodo.11091898, which requires the helper scripts in this additional repository: https://doi.org/10.5281/zenodo.11091813.

Data availability

We used publicly available data, namely the Warm-Winter-2020 data set from the Integrated Carbon Observation System (ICOS, https://www.icos-cp.eu/data-products/2G60-ZHAK), the ERA5-Land reanalysis (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land), Global Land Surface Satellite (GLASS) data derived from remote sensing (http://www.glass.umd.edu/index.html) and reanalysis data from the Global Evaporation Amsterdam Model (GLEAM, https://www.gleam.eu/). Intermediary tabular data in parquet format corresponding to the location of the ICOS stations are stored in https://doi.org/10.5281/zenodo.11091898 for each data source used here, including $CLM5_{grid}$ and $CLM5_{PFT}$. The raw CLM5 outputs over the whole European domain, which were not used in this study, can be made available upon request (approx. 8 terabyte).

Author contribution

C.P.T., B.S.N., and H.J.H.F. conceived and designed the study. C.P.T. processed the data and performed the analyses. B.S.N., H.J.H.F., R.B., and H.V. suggested the analyses and helped interpret the results. C.P.T. wrote the manuscript and edited the suggestions from all co-authors.

Competing interests

The authors declare that no competing interests are present.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

Acknowledgements

This project has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement No 871128 (eLTER PLUS).

The authors gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JURECA at Jülich Supercomputing Centre (JSC).

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

Bibliography

Ali, A. A., Xu, C., Rogers, A., Fisher, R. A., Wullschleger, S. D., Massoud, E. C., Vrugt, J. A., Muss, J. D., McDowell, N. G., Fisher, J. B., Reich, P. B., and Wilson, C. J.: A global scale mechanistic model of photosynthetic capacity (LUNA V1.0), Geosci. Model Dev., 9, 587–606, https://doi.org/10.5194/gmd-9-587-2016, 2016.

Anderegg, L. D. L., Griffith, D. M., Cavender-Bares, J., Riley, W. J., Berry, J. A., Dawson, T. E., and Still, C. J.: Representing plant diversity in land models: An evolutionary approach to make “Functional Types” more functional, Global Change Biology, 28, 2541–2554, https://doi.org/10.1111/gcb.16040, 2022.

Baatz, R., Sullivan, P. L., Li, L., Weintraub, S. R., Loescher, H. W., Mirtl, M., Groffman, P. M., Wall, D. H., Young, M., White, T., Wen, H., Zacharias, S., Kühn, I., Tang, J., Gaillardet, J., Braud, I., Flores, A. N., Kumar, P., Lin, H., Ghezzehei, T., Jones, J., Gholz, H. L., Vereecken, H., and Van Looy, K.: Steering operational synergies in terrestrial observation networks: opportunity for advancing Earth system dynamics modellingB6, Earth System Dynamics, 9, 593–609, https://doi.org/10.5194/esd-9-593-2018, 2018.

Baatz, R., Hendricks Franssen, H. J., Euskirchen, E., Sihi, D., Dietze, M., Ciavatta, S., Fennel, K., Beck, H., De Lannoy, G., Pauwels, V. R. N., Raiho, A., Montzka, C., Williams, M., Mishra, U., Poppe, C., Zacharias, S., Lausch, A., Samaniego, L., Van Looy, K., Bogena, H., Adamescu, M., Mirtl, M., Fox, A., Goergen, K., Naz, B. S., Zeng, Y., and Vereecken, H.: Reanalysis in Earth System Science: Toward Terrestrial Ecosystem Reanalysis, Rev Geophys, 59, https://doi.org/10.1029/2020RG000715, 2021.

Baker, E., Harper, A. B., Williamson, D., and Challenor, P.: Emulation of high-resolution land surface models using sparse Gaussian processes with application to JULES, Geosci. Model Dev., 15, 1913–1929, https://doi.org/10.5194/gmd-15-1913-2022, 2022.

Birch, L., Schwalm, C. R., Natali, S., Lombardozzi, D., Keppel-Aleks, G., Watts, J., Lin, X., Zona, D., Oechel, W., Sachs, T., Black, T. A., and Rogers, B. M.: Addressing biases in Arctic–boreal carbon cycling in the Community Land Model Version 5, Geosci. Model Dev., 14, 3361–3382, https://doi.org/10.5194/gmd-14-3361-2021, 2021.

Bloom, A. A., Bowman, K. W., Liu, J., Konings, A. G., Worden, J. R., Parazoo, N. C., Meyer, V., Reager, J. T., Worden, H. M., Jiang, Z., Quetin, G. R., Smallman, T. L., Exbrayat, J.-F., Yin, Y., Saatchi, S. S.,

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

Williams, M., and Schimel, D. S.: Lagged effects regulate the inter-annual variability of the tropical carbon balance, Biogeosciences, 17, 6393–6422, https://doi.org/10.5194/bg-17-6393-2020, 2020.

Boas, T., Bogena, H. R., Ryu, D., Vereecken, H., Western, A., and Hendricks Franssen, H.-J.: Seasonal soil moisture and crop yield prediction with fifth-generation seasonal forecasting system (SEAS5) long-range meteorological forecasts in a land surface modelling approach, Hydrol. Earth Syst. Sci., 27, 3143–3167, https://doi.org/10.5194/hess-27-3143-2023, 2023.

Bollmeyer, C., Keller, J. D., Ohlwein, C., Wahl, S., Crewell, S., Friederichs, P., Hense, A., Keune, J., Kneifel, S., Pscheidt, I., Redl, S., and Steinke, S.: Towards a high-resolution regional reanalysis for the European CORDEX domain: High-Resolution Regional Reanalysis for the European CORDEX Domain, Q.J.R. Meteorol. Soc., 141, 1–15, https://doi.org/10.1002/qj.2486, 2015.

Bonan, G. B., Levis, S., Kergoat, L., and Oleson, K. W.: Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models: PLANT FUNCTIONAL TYPES AND CLIMATE MODELS, Global Biogeochem. Cycles, 16, 5-1-5-23, https://doi.org/10.1029/2000GB001360, 2002.

Bonan, G. B., Lawrence, P. J., Oleson, K. W., Levis, S., Jung, M., Reichstein, M., Lawrence, D. M., and Swenson, S. C.: Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data, J. Geophys. Res., 116, G02014, https://doi.org/10.1029/2010JG001593, 2011.

Bonan, G. B., Williams, M., Fisher, R. A., and Oleson, K. W.: Modeling stomatal conductance in the earth system: linking leaf water-use efficiency and water transport along the soil–plant–atmosphere continuum, Geosci. Model Dev., 7, 2193–2222, https://doi.org/10.5194/gmd-7-2193-2014, 2014.

Caldararu, S., Purves, D. W., and Smith, M. J.: The effect of using the plant functional type paradigm on a data-constrained global phenology model, Biogeochemistry: Modelling, Terrestrial, 2015.

Cheng, Y., Huang, M., Zhu, B., Bisht, G., Zhou, T., Liu, Y., Song, F., and He, X.: Validation of the Community Land Model Version 5 Over the Contiguous United States (CONUS) Using In Situ and Remote Sensing Data Sets, JGR Atmospheres, 126, e2020JD033539, https://doi.org/10.1029/2020JD033539, 2021.

Chu, H., Luo, X., Ouyang, Z., Chan, W. S., Dengel, S., Biraud, S. C., Torn, M. S., Metzger, S., Kumar, J., Arain, M. A., Arkebauer, T. J., Baldocchi, D., Bernacchi, C., Billesbach, D., Black, T. A., Blanken, P. D., Bohrer, G., Bracho, R., Brown, S., Brunsell, N. A., Chen, J., Chen, X., Clark, K., Desai, A. R., Duman, T., Durden, D., Fares, S., Forbrich, I., Gamon, J. A., Gough, C. M., Griffis, T., Helbig, M., Hollinger, D., Humphreys, E., Ikawa, H., Iwata, H., Ju, Y., Knowles, J. F., Knox, S. H., Kobayashi, H., Kolb, T., Law, B., Lee, X., Litvak, M., Liu, H., Munger, J. W., Noormets, A., Novick, K., Oberbauer, S. F., Oechel, W., Oikawa, P., Papuga, S. A., Pendall, E., Prajapati, P., Prueger, J., Quinton, W. L., Richardson, A. D., Russell, E. S., Scott, R. L., Starr, G., Staebler, R., Stoy, P. C., Stuart-Haëntjens, E., Sonnentag, O., Sullivan, R. C., Suyker, A., Ueyama, M., Vargas, R., Wood, J. D., and Zona, D.: Representativeness of Eddy-Covariance flux footprints for areas surrounding AmeriFlux sites, Agricultural and Forest Meteorology, 301-302, 108350, https://doi.org/10.1016/j.agrformet.2021.108350, 2021.

Copernicus Climate Change Service: ERA5-Land hourly data from 2001 to present, https://doi.org/10.24381/CDS.E2161BAC, 2019.

Cranko Page, J., Abramowitz, G., De Kauwe, Martin. G., and Pitman, A. J.: Are Plant Functional Types Fit for Purpose?, Geophysical Research Letters, 51, e2023GL104962, https://doi.org/10.1029/2023GL104962, 2024.

Dagon, K., Sanderson, B. M., Fisher, R. A., and Lawrence, D. M.: A machine learning approach to emulation and biophysical parameter estimation with the Community Land Model, version 5, Adv. Stat. Clim. Meteorol. Oceanogr., 6, 223–244, https://doi.org/10.5194/ascmo-6-223-2020, 2020.

De Pue, J., Wieneke, S., Bastos, A., Barrios, J. M., Liu, L., Ciais, P., Arboleda, A., Hamdi, R., Maleki, M., Maignan, F., Gellens-Meulenberghs, F., Janssens, I., and Balzarolo, M.: Temporal variability of observed and simulated gross primary productivity, modulated by vegetation state and hydrometeorological drivers, Biogeochemistry: Modelling, Terrestrial, 2023.

Deng, M., Meng, X., Lu, Y., Li, Z., Zhao, L., Hu, Z., Chen, H., Shang, L., Wang, S., and Li, Q.: Impact and Sensitivity Analysis of Soil Water and Heat Transfer Parameterizations in Community Land Surface Model on the Tibetan Plateau, J Adv Model Earth Syst, 13, https://doi.org/10.1029/2021MS002670, 2021.

Earth Resources Observation And Science (EROS) Center: Global Topographic 30 Arc-Second Hydrologic Digital Elevation Model 1 km, https://doi.org/10.5066/F77P8WN0, 2017.

Eshonkulov, R., Poyda, A., Ingwersen, J., Wizemann, H.-D., Weber, T. K. D., Kremer, P., Högy, P., Pulatov, A., and Streck, T.: Evaluating multi-year, multi-site data on the energy balance closure of eddy-covariance flux measurements at cropland sites in southwestern Germany, Biogeosciences, 16, 521–540, https://doi.org/10.5194/bg-16-521-2019, 2019.

Farquhar, G. D., von Caemmerer, S., and Berry, J. A.: A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species, Planta, 149, 78–90, https://doi.org/10.1007/BF00386231, 1980.

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

Fisher, R. A., Wieder, W. R., Sanderson, B. M., Koven, C. D., Oleson, K. W., Xu, C., Fisher, J. B., Shi, M., Walker, A. P., and Lawrence, D. M.: Parametric Controls on Vegetation Responses to Biogeochemical Forcing in the CLM5, J Adv Model Earth Syst, 11, 2879–2895, https://doi.org/10.1029/2019MS001609, 2019a.

Fisher, R. A., Wieder, W. R., Sanderson, B. M., Koven, C. D., Oleson, K. W., Xu, C., Fisher, J. B., Shi, M., Walker, A. P., and Lawrence, D. M.: Parametric Controls on Vegetation Responses to Biogeochemical Forcing in the CLM5, J. Adv. Model. Earth Syst., 11, 2879–2895, https://doi.org/10.1029/2019MS001609, 2019b.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P., Le Quéré, C., Luijkx, I. T., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., Barbero, L., Bates, N. R., Becker, M., Bellouin, N., Decharme, B., Bopp, L., Brasika, I. B. M., Cadule, P., Chamberlain, M. A., Chandra, N., Chau, T.-T.-T., Chevallier, F., Chini, L. P., Cronin, M., Dou, X., Enyo, K., Evans, W., Falk, S., Feely, R. A., Feng, L., Ford, D. J., Gasser, T., Ghattas, J., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Heinke, J., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Jacobson, A. R., Jain, A., Jarníková, T., Jersild, A., Jiang, F., Jin, Z., Joos, F., Kato, E., Keeling, R. F., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Körtzinger, A., Lan, X., Lefèvre, N., Li, H., Liu, J., Liu, Z., Ma, L., Marland, G., Mayot, N., McGuire, P. C., McKinley, G. A., Meyer, G., Morgan, E. J., Munro, D. R., Nakaoka, S.-I., Niwa, Y., O'Brien, K. M., Olsen, A., Omar, A. M., Ono, T., Paulsen, M., Pierrot, D., Pocock, K., Poulter, B., Powis, C. M., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T. M., Schwinger, J., Séférian, R., et al.: Global Carbon Budget 2023, Earth Syst. Sci. Data, 15, 5301–5369, https://doi.org/10.5194/essd-15-5301-2023, 2023.

Futter, M. N., Dirnböck, T., Forsius, M., Bäck, J. K., Cools, N., Diaz-Pines, E., Dick, J., Gaube, V., Gillespie, L. M., Högbom, L., Laudon, H., Mirtl, M., Nikolaidis, N., Poppe Terán, C., Skiba, U., Vereecken, H., Villock, H., Weldon, J., Wohner, C., and Alam, S. A.: Leveraging research infrastructure co-location to evaluate constraints on terrestrial carbon cycling in northern European forests, Ambio, https://doi.org/10.1007/s13280-023-01930-4, 2023.

Giorgi, F., Jones, C., and Asrar, G. R.: Addressing Climate Information Needs at the Regional Level: the CORDEX Framework, World Meteorological Organization Bulletin, 58, 175–183, 2009.

Graf, A., Klosterhalfen, A., Arriga, N., Bernhofer, C., Bogena, H., Bornet, F., Brüggemann, N., Brümmer, C., Buchmann, N., Chi, J., Chipeaux, C., Cremonese, E., Cuntz, M., Dušek, J., El-Madany, T. S., Fares, S., Fischer, M., Foltýnová, L., Gharun, M., Ghiasi, S., Gielen, B., Gottschalk, P., Grünwald, T., Heinemann, G., Heinesch, B., Heliasz, M., Holst, J., Hörtnagl, L., Ibrom, A., Ingwersen, J., Jurasinski, G., Klatt, J., Knohl, A., Koebsch, F., Konopka, J., Korkiakoski, M., Kowalska, N., Kremer, P., Kruijt, B., Lafont, S., Léonard, J., De Ligne, A., Longdoz, B., Loustau, D., Magliulo, V., Mammarella, I., Manca, G., Mauder, M., Migliavacca, M., Mölder, M., Neirynck, J., Ney, P., Nilsson, M., Paul-Limoges, E., Peichl, M., Pitacco, A., Poyda, A., Rebmann, C., Roland, M., Sachs, T., Schmidt, M., Schrader, F., Siebicke, L., Šigut, L., Tuittila, E.-S., Varlagin, A., Vendrame, N., Vincke, C., Völksch, I., Weber, S., Wille, C., Wizemann, H.-D., Zeeman, M., and Vereecken, H.: Altered energy partitioning across terrestrial ecosystems in the European drought year 2018, Phil. Trans. R. Soc. B, 375, 20190524, https://doi.org/10.1098/rstb.2019.0524, 2020.

Graf, A., Wohlfahrt, G., Aranda-Barranco, S., Arriga, N., Brümmer, C., Ceschia, E., Ciais, P., Desai, A. R., Di Lonardo, S., Gharun, M., Grünwald, T., Hörtnagl, L., Kasak, K., Klosterhalfen, A., Knohl, A., Kowalska, N., Leuchner, M., Lindroth, A., Mauder, M., Migliavacca, M., Morel, A. C., Pfennig, A., Poorter, H., Terán, C. P., Reitz, O., Rebmann, C., Sanchez-Azofeifa, A., Schmidt, M., Šigut, L., Tomelleri, E., Yu, K., Varlagin, A., and Vereecken, H.: Joint optimization of land carbon uptake and albedo can help achieve moderate instantaneous and long-term cooling effects, Commun Earth Environ, 4, 298, https://doi.org/10.1038/s43247-023-00958-4, 2023.

Green, J. K., Zhang, Y., Luo, X., and Keenan, T. F.: Systematic Underestimation of Canopy Conductance Sensitivity to Drought by Earth System Models, AGU Advances, 5, e2023AV001026, https://doi.org/10.1029/2023AV001026, 2024.

Harding, B., Tremblay, C., and Cousineau, D.: Standard errors: A review and evaluation of standard error estimators using Monte Carlo simulations, TQMP, 10, 107–123, https://doi.org/10.20982/tqmp.10.2.p107, 2014.

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Klein Goldewijk, K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., Lawrence, P., Ma, L., Mertz, O., Pongratz, J., Popp, A., Poulter, B., Riahi, K., Shevliakova, E., Stehfest, E., Thornton, P., Tubiello, F. N., van Vuuren, D. P., and Zhang, X.: Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6, Geosci. Model Dev., 13, 5425–5464, https://doi.org/10.5194/gmd-13-5425-2020, 2020.

ICOS RI: Ecosystem final quality (L2) product in ETC-Archive format - release 2021-1, 911 MB, https://doi.org/10.18160/FZMY-PG92, 2021.

IGBP: Global Soil Data Task (IGBP-DIS, ISO-image of CD), https://doi.org/10.1594/PANGAEA.869912, 2000.

Jafari, M., Tavili, A., Panahi, F., Zandi Esfahan, E., and Ghorbani, M.: Introduction, in: Reclamation of Arid Lands, Springer International Publishing, Cham, 1–19, 2018.

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

Kattge, J., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M., Reich, P. B., Wright, I. J., Cornelissen, J. H. C., Violle, C., Harrison, S. P., Van Bodegom, P. M., Reichstein, M., Enquist, B. J., Soudzilovskaia, N. A., Ackerly, D. D., Anand, M., Atkin, O., Bahn, M., Baker, T. R., Baldocchi, D., Bekker, R., Blanco, C. C., Blonder, B., Bond, W. J., Bradstock, R., Bunker, D. E., Casanoves, F., Cavender-Bares, J., Chambers, J. Q., Chapin Iii, F. S., Chave, J., Coomes, D., Cornwell, W. K., Craine, J. M., Dobrin, B. H., Duarte, L., Durka, W., Elser, J., Esser, G., Estiarte, M., Fagan, W. F., Fang, J., Fernández-Méndez, F., Fidelis, A., Finegan, B., Flores, O., Ford, H., Frank, D., Freschet, G. T., Fyllas, N. M., Gallagher, R. V., Green, W. A., Gutierrez, A. G., Hickler, T., Higgins, S. I., Hodgson, J. G., Jalili, A., Jansen, S., Joly, C. A., Kerkhoff, A. J., Kirkup, D., Kitajima, K., Kleyer, M., Klotz, S., Knops, J. M. H., Kramer, K., Kühn, I., Kurokawa, H., Laughlin, D., Lee, T. D., Leishman, M., Lens, F., Lenz, T., Lewis, S. L., Lloyd, J., Llusià, J., Louault, F., Ma, S., Mahecha, M. D., Manning, P., Massad, T., Medlyn, B. E., Messier, J., Moles, A. T., Müller, S. C., Nadrowski, K., Naeem, S., Niinemets, Ü., Nöllert, S., Nüske, A., Ogaya, R., Oleksyn, J., Onipchenko, V. G., Onoda, Y., Ordoñez, J., Overbeck, G., et al.: TRY - a global database of plant traits: TRY - A GLOBAL DATABASE OF PLANT TRAITS, Global Change Biology, 17, 2905–2935, https://doi.org/10.1111/j.1365-2486.2011.02451.x, 2011.

Kühn, N., Tovar, C., Carretero, J., Vandvik, V., Enquist, B. J., and Willis, K. J.: Globally important plant functional traits for coping with climate change, Frontiers of Biogeography, 13, https://doi.org/10.21425/F5FBG53774, 2021.

Lawrence, D. M., Fisher, R., Koven, C., Oleson, K., Swenson, S., Mariana Vertenstein, Ben Andre, Gordon Bonan, Bardan Ghimire, Leo van Kampenhout, Daniel Kennedy, Erik Kluzek, Ryan Knox, Peter Lawrence, Fang Li, Hongyi Li, Danica Lombardozzi, Yaqiong Lu, Justin Perket, William Riley, William Sacks, Mingjie Shi, Will Wieder, Chonggang Xu, Ashehad Ali, Andrew Badger, Gautam Bisht, Patrick Broxton, Michael Brunke, Jonathan Buzan, Martyn Clark, Tony Craig, Kyla Dahlin, Beth Drewniak, Louisa Emmons, Josh Fisher, Mark Flanner, Pierre Gentine, Jan Lenaerts, Sam Levis, L. Ruby Leung, William Lipscomb, Jon Pelletier, Daniel M. Ricciuto, Ben Sanderson, Jacquelyn Shuman, Andrew Slater, Zachary Subin, Jinyun Tang, Ahmed Tawfik, Quinn Thomas, Simone Tilmes, Francis Vitt, and Xubin Zeng: CLM5 Documentation (Technical Note), https://www2.cesm.ucar.edu/models/cesm2/land/CLM50_Tech_Note.pdf, February 2018.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., Broeke, M., Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Val Martin, M., and Zeng, X.: The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty, J. Adv. Model. Earth Syst., 11, 4245–4287, https://doi.org/10.1029/2018MS001583, 2019.

Lawrence, P. J. and Chase, T. N.: Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0), J. Geophys. Res., 112, G01023, https://doi.org/10.1029/2006JG000168, 2007.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

Lehner, F., Coats, S., Stocker, T. F., Pendergrass, A. G., Sanderson, B. M., Raible, C. C., and Smerdon, J. E.: Projected drought risk in 1.5°C and 2°C warmer climates, Geophysical Research Letters, 44, 7419–7428, https://doi.org/https://doi.org/10.1002/2017GL074117, 2017.

Liang, S., Cheng, J., Jia, K., Jiang, B., Liu, Q., Xiao, Z., Yao, Y., Yuan, W., Zhang, X., Zhao, X., and Zhou, J.: The Global Land Surface Satellite (GLASS) Product Suite, Bulletin of the American Meteorological Society, 102, E323–E337, https://doi.org/10.1175/BAMS-D-18-0341.1, 2021.

Lin, Y.-S., Medlyn, B. E., Duursma, R. A., Prentice, I. C., Wang, H., Baig, S., Eamus, D., De Dios, V., Mitchell, P., Ellsworth, D. S., De Beeck, M. O., Wallin, G., Uddling, J., Tarvainen, L., Linderson, M.-L., Cernusak, L. A., Nippert, J. B., Ocheltree, T. W., Tissue, D. T., Martin-StPaul, N. K., Rogers, A., Warren, J. M., De Angelis, P., Hikosaka, K., Han, Q., Onoda, Y., Gimeno, T. E., Barton, C. V. M., Bennie, J., Bonal, D., Bosc, A., Löw, M., Macinins-Ng, C., Rey, A., Rowland, L., Setterfield, S. A., Tausz-Posch, S., Zaragoza-Castells, J., Broadmeadow, M. S. J., Drake, J. E., Freeman, M., Ghannoum, O., Hutley, Lindsay B., Kelly, J. W., Kikuzawa, K., Kolari, P., Koyama, K., Limousin, J.-M., Meir, P., Lola Da Costa, A. C., Mikkelsen, T. N., Salinas, N., Sun, W., and Wingate, L.: Optimal stomatal behaviour around the world, Nature Clim Change, 5, 459–464, https://doi.org/10.1038/nclimate2550, 2015.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geosci. Model Dev., 10, 1903–1925, https://doi.org/10.5194/gmd-10-1903-2017, 2017.

Medlyn, B. E., Duursma, R. A., Eamus, D., Ellsworth, D. S., Prentice, I. C., Barton, C. V. M., Crous, K. Y., De Angelis, P., Freeman, M., and Wingate, L.: Reconciling the optimal and empirical approaches to modelling stomatal conductance: RECONCILING OPTIMAL AND EMPIRICAL STOMATAL MODELS, Global Change Biology, 17, 2134–2144, https://doi.org/10.1111/j.1365-2486.2010.02375.x, 2011.

Migliavacca, M., Musavi, T., Mahecha, M. D., Nelson, J. A., Knauer, J., Baldocchi, D. D., Perez-Priego, O., Christiansen, R., Peters, J., Anderson, K., Bahn, M., Black, T. A., Blanken, P. D., Bonal, D., Buchmann, N., Caldararu, S., Carrara, A., Carvalhais, N., Cescatti, A., Chen, J., Cleverly, J., Cremonese, E., Desai, A. R., El-Madany, T. S., Farella, M. M., Fernández-Martínez, M., Filippa, G., Forkel, M., Galvagno, M., Gomarasca, U., Gough, C. M., Göckede, M., Ibrom, A., Ikawa, H., Janssens, I. A., Jung, M., Kattge, J., Keenan, T. F., Knohl, A., Kobayashi, H., Kraemer, G., Law, B. E., Liddell, M. J., Ma, X., Mammarella, I., Martini, D., Macfarlane, C., Matteucci, G., Montagnani, L., Pabon-Moreno, D. E., Panigada, C., Papale, D., Pendall, E., Penuelas, J., Phillips, R. P., Reich, P. B., Rossini, M., Rotenberg, E., Scott, R. L., Stahl, C., Weber, U., Wohlfahrt, G., Wolf, S., Wright, I. J., Yakir, D., Zaehle, S., and Reichstein, M.: The three major axes of terrestrial ecosystem function, Nature, https://doi.org/10.1038/s41586-021-03939-9, 2021.

Mirtl, M., T. Borer, E., Djukic, I., Forsius, M., Haubold, H., Hugo, W., Jourdan, J., Lindenmayer, D., McDowell, W. H., Muraoka, H., Orenstein, D. E., Pauw, J. C., Peterseil, J., Shibata, H., Wohner, C., Yu, X., and Haase, P.: Genesis, goals and achievements of Long-Term Ecological Research at the global scale: A critical review of ILTER and future directions, Science of The Total Environment, 626, 1439–1462, https://doi.org/10.1016/j.scitotenv.2017.12.001, 2018.

Mirtl, M., Kuhn, I., Montheith, D., Bäck, J., Orenstein, D., Provenzale, A., Zacharias, S., Haase, P., and Shachak, M.: Whole System Approach for in-situ research on Life Supporting Systems in the Anthropocene (WAILS), pico, 2021.

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., 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. de, Ligne, A. D., De Oliveira, R. C., Delpierre, N., Desai, A. R., Di Bella, C. M., Tommasi, P. di, 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., et al.: 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, 2020.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

Pilli, R., Grassi, G., Kurz, W. A., Fiorese, G., and Cescatti, A.: The European forest sector: past and future carbon budget and fluxes under different management scenarios, Biogeosciences, 14, 2387–2405, https://doi.org/10.5194/bg-14-2387-2017, 2017.

Poppe Terán, C., Naz, B. S., Graf, A., Qu, Y., Hendricks Franssen, H.-J., Baatz, R., Ciais, P., and Vereecken, H.: Rising water-use efficiency in European grasslands is driven by increased primary production, Commun Earth Environ, 4, 95, https://doi.org/10.1038/s43247-023-00757-x, 2023.

Prein, A. F., Gobiet, A., Truhetz, H., Keuler, K., Goergen, K., Teichmann, C., Fox Maule, C., van Meijgaard, E., Déqué, M., Nikulin, G., Vautard, R., Colette, A., Kjellström, E., and Jacob, D.: Precipitation in the EURO-CORDEX 0.11° and 0.44° simulations: high resolution, high benefits?, Clim Dyn, 46, 383–412, https://doi.org/10.1007/s00382-015-2589-y, 2016.

Rahmati, M., Graf, A., Poppe Terán, C., Amelung, W., Dorigo, W., Franssen, H.-J. H., Montzka, C., Or, D., Sprenger, M., Vanderborght, J., Verhoest, N. E. C., and Vereecken, H.: Continuous increase in evaporative demand shortened the growing season of European ecosystems in the last decade, Commun Earth Environ, 4, 236, https://doi.org/10.1038/s43247-023-00890-7, 2023.

Reichstein, M., Stoy, P. C., Desai, A. R., Lasslop, G., and Richardson, A. D.: Partitioning of Net Fluxes, in: Eddy Covariance, edited by: Aubinet, M., Vesala, T., and Papale, D., Springer Netherlands, Dordrecht, 263–289, 2012.

Rousi, E., Kornhuber, K., Beobide-Arsuaga, G., Luo, F., and Coumou, D.: Accelerated western European heatwave trends linked to more-persistent double jets over Eurasia, Nat Commun, 13, 3851, https://doi.org/10.1038/s41467-022-31432-y, 2022.

Scott, D. W.: Multivariate Density Estimation: Theory, Practice, and Visualization, 1st ed., Wiley, 1992.

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., Doney, S. C., Graven, H., Heinze, C., Huntingford, C., Levis, S., Levy, P. E., Lomas, M., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Arneth, A., Bonan, G., Bopp, L., Canadell, J. G., Chevallier, F., Ciais, P., Ellis, R., Gloor, M., Peylin, P., Piao, S. L., Le Quéré, C., Smith, B., Zhu, Z., and Myneni, R.: Recent trends and drivers of regional sources and sinks of carbon dioxide, Biogeosciences, 12, 653–679, https://doi.org/10.5194/bg-12-653-2015, 2015.

Solomon, A. M. and Shugart, H. H. (Eds.): Vegetation dynamics & global change, Chapman & Hall ; IIASA], New York : [Laxenburg, Austria, 338 pp., 1993.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

Song, J., Miller, G. R., Cahill, A. T., Aparecido, L. M. T., and Moore, G. W.: Modeling land surface processes over a mountainous rainforest in Costa Rica using CLM4.5 and CLM5, Geosci. Model Dev., 13, 5147–5173, https://doi.org/10.5194/gmd-13-5147-2020, 2020.

Strebel, L., Bogena, H., Vereecken, H., Andreasen, M., Aranda, S., and Hendricks Franssen, H.-J.: Evapotranspiration prediction for European forest sites does not improve with assimilation of in-situ soil water content data, Vadose Zone Hydrology/Modelling approaches, 2023.

Ukkola, A. M., De Kauwe, M. G., Pitman, A. J., Best, M. J., Abramowitz, G., Haverd, V., Decker, M., and Haughton, N.: Land surface models systematically overestimate the intensity, duration and magnitude of seasonal-scale evaporative droughts, Environ. Res. Lett., 11, 104012, https://doi.org/10.1088/1748-9326/11/10/104012, 2016.

Umair, M., Kim, D., and Choi, M.: Impact of climate, rising atmospheric carbon dioxide, and other environmental factors on water-use efficiency at multiple land cover types, Sci Rep, 10, 11644, https://doi.org/10.1038/s41598-020-68472-7, 2020.

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

Van Bodegom, P. M., Douma, J. C., and Verheijen, L. M.: A fully traits-based approach to modeling global vegetation distribution, Proc. Natl. Acad. Sci. U.S.A., 111, 13733–13738, https://doi.org/10.1073/pnas.1304551110, 2014.

Van Der Woude, A. M., Peters, W., Joetzjer, E., Lafont, S., Koren, G., Ciais, P., Ramonet, M., Xu, Y., Bastos, A., Botía, S., Sitch, S., De Kok, R., Kneuer, T., Kubistin, D., Jacotot, A., Loubet, B., Herig-Coimbra, P.-H., Loustau, D., and Luijkx, I. T.: Temperature extremes of 2022 reduced carbon uptake by forests in Europe, Nat Commun, 14, 6218, https://doi.org/10.1038/s41467-023-41851-0, 2023.

Vautard, R., Cattiaux, J., Happé, T., Singh, J., Bonnet, R., Cassou, C., Coumou, D., D’Andrea, F., Faranda, D., Fischer, E., Ribes, A., Sippel, S., and Yiou, P.: Heat extremes in Western Europe increasing faster than simulated due to atmospheric circulation trends, Nat Commun, 14, 6803, https://doi.org/10.1038/s41467-023-42143-3, 2023.

https://doi.org/10.5194/egusphere-2024-978
Preprint. Discussion started: 7 May 2024

Wahl, S., Bollmeyer, C., Crewell, S., Figura, C., Friederichs, P., Hense, A., Keller, J. D., and Ohlwein, C.: A novel convective-scale regional reanalysis COSMO-REA2: Improving the representation of precipitation, metz, 26, 345–361, https://doi.org/10.1127/metz/2017/0824, 2017.

Warm Winter 2020 Team and ICOS Ecosystem Thematic Centre: 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, 1 February 2022.

Winkler, K., Yang, H., Ganzenmüller, R., Fuchs, R., Ceccherini, G., Duveiller, G., Grassi, G., Pongratz, J., Bastos, A., Shvidenko, A., Araza, A., Herold, M., Wigneron, J.-P., and Ciais, P.: Changes in land use and management led to a decline in Eastern Europe's terrestrial carbon sink, Commun Earth Environ, 4, 237, https://doi.org/10.1038/s43247-023-00893-4, 2023.

Wozniak, M. C., Bonan, G. B., Keppel-Aleks, G., and Steiner, A. L.: Influence of Vertical Heterogeneities in the Canopy Microenvironment on Interannual Variability of Carbon Uptake in Temperate Deciduous Forests, JGR Biogeosciences, 125, e2020JG005658, https://doi.org/10.1029/2020JG005658, 2020.

Wu, H., Fu, C., Wu, H., and Zhang, L.: Plant Hydraulic Stress Strategy Improves Model Predictions of the Response of Gross Primary Productivity to Drought Across China, JGR Atmospheres, 125, e2020JD033476, https://doi.org/10.1029/2020JD033476, 2020.

Xie, X., Li, A., Tan, J., Lei, G., Jin, H., and Zhang, Z.: Uncertainty analysis of multiple global GPP datasets in characterizing the lagged effect of drought on photosynthesis, Ecological Indicators, 113, 106224, https://doi.org/10.1016/j.ecolind.2020.106224, 2020.

Zhang, L., Ning, F., Bai, X., Zeng, X., and He, C.: Performance evaluation of CLM5.0 in simulating liquid soil water in high mountainous area, Northwest China, J. Mt. Sci., 20, 1865–1883, https://doi.org/10.1007/s11629-022-7803-x, 2023.

Zhang, W., Nelson, J. A., Miralles, D. G., Mauder, M., Migliavacca, M., Poyatos, R., Reichstein, M., and Jung, M.: A New Post-Hoc Method to Reduce the Energy Imbalance in Eddy Covariance Measurements, Geophysical Research Letters, 51, e2023GL107084, https://doi.org/10.1029/2023GL107084, 2024.

Zheng, Y., Shen, R., Wang, Y., Li, X., Liu, S., Liang, S., Chen, J. M., Ju, W., Zhang, L., and Yuan, W.: Improved estimate of global gross primary production for reproducing its long-term variation, 1982–2017, Earth Syst. Sci. Data, 12, 2725–2746, https://doi.org/10.5194/essd-12-2725-2020, 2020.