https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

Earth system models might overestimate the local plant productivity response to temperature–moisture extremes

Moritz Adam[1,2], Elisa Ziegler[1], Björn Gonzalez[1], Nils Weitzel[3,4,5], and Kira Rehfeld[1,2,6]

[1]Department of Geosciences, University of Tübingen, Tübingen, Germany
[2]Cluster of Excellence (EXC 3121): TERRA — Terrestrial Geo-Biosphere Interactions in a Changing World, University of Tübingen, Tübingen, Germany
[3]Research Center Trustworthy Data Science and Security of the University Alliance Ruhr, Dortmund, Germany
[4]Department of Statistics, TU Dortmund University, Dortmund, Germany
[5]School of Geographical Sciences, University of Bristol, Bristol, UK
[6]Department of Physics, University of Tübingen, Tübingen, Germany

Correspondence: Moritz Adam (moritz.adam@uni-tuebingen.de) and Elisa Ziegler (elisa.ziegler@uni-tuebingen.de)

Abstract. Compound temperature–moisture extremes, such as droughts or hot–wet extremes, have a pronounced and sometimes long-lasting impact on vegetation productivity. Accurate simulation of the involved processes by emission-driven Earth system models (ESMs) is crucial for inferring future terrestrial carbon uptake. However, ESMs often exhibit biases in the frequency and intensity of climate and weather extremes. Their ability to reproduce observed impacts of extreme atmospheric conditions on gross primary productivity (GPP) is therefore unclear. Comprehensive assessments of the statistical link between compound events and vegetation productivity beyond individual regions or event types are rare. Here, we scrutinize the relationship between temperature–moisture extremes and exceptionally low or high vegetation productivity in two state–of–the–art ESMs, CESM2 and MPI-ESM1.2, and gauge their performance relative to observation-constrained data. We find that temperature-moisture extremes modulate vegetation productivity in observations and models. The global-scale strength and timing of the statistical relationship agree well between observation-based data and model output. However, this agreement deteriorates towards smaller spatial scales, especially in the low latitudes. Here, an overestimated coupling strength by both models, likely related to biased rates of soil moisture change, suggests potentially unrealistic evaporative feedbacks, exaggerated drainage, or inadequate effective water-holding capacity in ESMs. Nevertheless, all data sources identify coherent significant relationships for all combinations of temperature–moisture and GPP extremes. This result highlights both beneficial and detrimental influences of temperature–moisture compound events on vegetation productivity and the importance of comprehensive assessments beyond single event types for capturing the net effect of climatic extremes on the biosphere. Further research should examine whether overestimated plant productivity responses to extreme conditions are a recurring phenomenon across all Earth system models. It could also investigate non-stationarity and nonlinearity of the relationships between climatic and vegetation extremes under climate change.

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

1 Introduction

Climate change goes far beyond shifting the Earth's mean state. Rising CO_2 levels increase the likelihood of temperature and hydroclimate extremes, threatening the planet's ecosystems and human livelihoods (Mahecha et al., 2024). Temperature and moisture anomalies can, individually, affect the biosphere. However, they tend to be particularly impactful when they co-occur as compound extreme events (Zscheischler et al., 2014b), which is increasingly likely (Ridder et al., 2022; Li et al., 2024). Examples of past compound events that severely damaged vegetation productivity include European heat–droughts diminishing carbon uptake by forests in 2018 and 2022 (Van Der Woude et al., 2023; Gharun et al., 2024; Zscheischler and Fischer, 2020), spring- and summertime cold–dry spells reducing gross primary production (GPP) in the mid- to high latitudes over the past decades (Li et al., 2022), and diminishing tropical GPP in response to droughts caused by El Niño in 2015/2016 (Bastos et al., 2018). At the same time, heat–drought events followed by high GPP anomalies in Siberia in the 2010s (Flach et al., 2021) and stronger tree growth in response to larger rainfall amounts from heavy precipitation days in semi-arid Australia (O'Donnell et al., 2021) are examples of climatic extremes that promoted vegetation growth. Disentangling these diverse, damaging, and beneficial impacts of temperature–moisture extremes on vegetation, as well as being able to model and project them under future climate change scenarios, is crucial (Zhou et al., 2019; Wu et al., 2021; Li et al., 2024), given the pivotal role of GPP for the land carbon sink (Sharma et al., 2023) and other ecosystem services provided by vegetation (Mahecha et al., 2024).

Earth system models (ESMs), i.e., comprehensive models of ocean, atmosphere, land surface, and the carbon cycle, struggle to capture the real-world frequency and intensity of temperature and moisture extremes both individually and in combination (Lorenz et al., 2016; Haren et al., 2015; Norris et al., 2021; Simpson et al., 2025; Reichstein et al., 2013). It is unclear how well ESMs capture the relationship between climatic (compound) events and vegetation productivity extremes. Both increased likelihoods of compound events and changes in their mechanistic link to vegetation extremes under climate change influence the land carbon sink and ecosystem functioning (Li et al., 2024; Tschumi et al., 2023). Abilities of ESMs to simulate the relation of moisture availability and vegetation extremes have been studied for specific extreme event combinations such as heat–droughts or heavy rainfall events (Harris et al., 2024; Flach et al., 2021; Li et al., 2022), but there are no comprehensive assessments across event types and regions for the co-occurrence of temperature–moisture and vegetation. Most often, bias assessments have focused on the drivers of vegetation extremes or their impacts, and were limited to specific areas (e.g., Villalobos-Herrera et al., 2021; Vogel et al., 2021). Assessing both on a global scale is crucial for model improvement and for developing statistical bias corrections, if necessary, to feed into climate change impact assessments. ESMs may continue to underestimate the likelihoods of temperature–moisture events and productivity extremes individually. However, if they accurately represented the statistical relationships between compound and GPP extremes, relative changes in these relationships could still provide valuable indications of shifting likelihoods of vegetation extremes under (hydro-)climate change.

ESMs draw on land surface models (LSMs) to simulate impacts of weather extremes on vegetation. LSMs simplify or disregard fine-grained, localized representation of plant species diversity, demographics, and ecosystem heterogeneities. Parametriza-

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

tions aggregate effects of vegetation–climate interactions on the carbon cycle at sub-grid-scales. This includes processes that affect a plant's ability to assimilate carbon, plant physiology, plant-soil interactions, nutrient availability, surface and subsurface water budgets, and climatic constraints (Blyth et al., 2021). The parameterizations employed in models from the sixth climate model intercomparison project (CMIP6; Eyring et al., 2016) and the imminent CMIP7 cycle typically generalize vegetation categories into functional types and average across heterogeneities, such as soil properties and root-zone nutrient dynamics (Blyth et al., 2021).

Land and vegetation models within ESMs operate at a high level of abstraction compared to real-world complexity. Calibration of the remaining free parameters allows ESMs to represent average land carbon stocks and dynamics well (Arora et al., 2020), at least in a transient mean state under present-day or recent-past climates. This suggests some predictive capabilities for regional-to-global mean changes in terrestrial carbon stocks under near-future climate change. However, models tend to exaggerate CO_2 fertilization trends (Kou-Giesbrecht et al., 2025). It is also challenging to model the highly variable land carbon sink since (hydro-)climate variability affects moisture availability and, therefore, carbon uptake through vegetation (Gentine et al., 2019). ESMs can represent the well-understood mechanisms by which moisture availability drives GPP, but results differ across models, and they often miss the precise timing, intensity, and location of compound event impacts (Harris et al., 2024). They also underestimate legacy effects on vegetation productivity compared to observations and paleovegetation reconstructions (Kolus et al., 2019; Dallmeyer et al., 2022). Whether these simplifications and model biases in occurrence and intensity of climate and biosphere extremes affect possible statistical links between the two types of extremes is unclear.

Here, we quantify model performance across combinations of compound and GPP extremes, for different regions, and spatial scales. As a step towards quantifying the net effects of compound events on GPP extremes, we study whether compound events increase the likelihood of subsequent GPP extremes, as well as potentially mitigating effects from the occurrence of compound events. For example, a prolonged cold-wet extreme in spring could increase soil moisture storage enough to reduce the likelihood of low GPP extremes in summer.

We explore the statistical relation between temperature–moisture compound events and GPP extremes in two CMIP6 models and assess them relative to gridded, observation-based products (Sect. 2). Employing a generalizable measure and statistical null model for this relationship (Sect. 3), we find reasonable agreement between models and observations (Sect. 4). However, this agreement depends on spatial aggregation and is subject to model-specific biases. In particular, both ESMs tend to overestimate the local strength of the coupling between climatic and vegetation extremes. We suggest that this overestimation might arise from an inaccurate soil moisture representation in the ESMs (Sect. 5.1). We find mechanistically comparable results across both data sources despite the biases (Sect. 4.3, Sect. 5.2), providing opportunities to study climate change impacts on vegetation productivity beyond the mean state in comprehensive model intercomparisons (Sect. 5.3).

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

2 Data

We scrutinize the capabilities of two frequently used, state-of-the-art ESMs from the CMIP6 cycle to capture statistical couplings between temperature–moisture compound events and extremes in GPP within simulations of the recent past (1979–2014 CE). We explore historical simulations with global coverage from the Max Planck Institute ESM version 1.2 (MPI-ESM) (Mauritsen et al., 2019; Wieners et al., 2019) and the Community ESM version 2 (CESM2) (Danabasoglu et al., 2020; Danabasoglu, 2019). Both fall into typical ranges of CMIP6 models and commonly perform well in different types of model evaluations (Tokarska et al., 2020; Kim et al., 2020; Arora et al., 2020; Wang et al., 2021). The models combine interacting submodels of the atmosphere, ocean, land surface (including vegetation), and marine and terrestrial biogeochemistry at atmospheric resolutions of 1.875° by 1.875° (MPI-ESM) and 1.25° by 0.9° (CESM2), respectively. MPI-ESM contains JSBACH version 3.2 as a land and dynamic global vegetation model (Reick et al., 2013). The community land model (CLM) version 5 (Lawrence et al., 2019) is the land and vegetation component of CESM2. Both land models are state-of-the-art for the CMIP6 cycle (Blyth et al., 2021) and represent the (sub-)surface energy and moisture budgets, vegetation composition through plant functional types (PFTs), nitrogen-limited photosynthesis with autotrophic respiration, and above- and below-ground heterotrophic respiration. They both use LUH2 (Hurtt et al., 2020) as a historical land-use constraint, but differ in their ability to resolve land use management practices as granular PFTs, for example, regarding irrigation.

By selecting MPI-ESM and CESM2 for our case study, we bracket different characteristics of simulated extreme events. Heavy precipitation extremes in CESM2 are more intense and closer to observations than in many other models, whereas they are weaker in MPI-ESM than in most ESMs (Norris et al., 2021). This is reflected in surface and soil water budgets of JSBACH3.2, which show a low bias relative to observations and several other land models, including CLM5.0 (Seiler et al., 2022). These differences between the ESMs are attractive for assessing the link between temperature-moisture and GPP extremes, because they can provide mechanistic insights into potential biases and commonalities. In contrast to the differences in hydroclimate variability, the global carbon budget is close to the CMIP6 multi-model mean in both models, with a relatively larger importance of land carbon sinks and more pronounced land carbon sink feedbacks in CESM2 compared to MPI-ESM (Arora et al., 2020). Against observational benchmarks of annual GPP, net biosphere productivity, carbon stocks, and latent heat flux, CLM5.0 and JSBACH3.2 also show similar performance (Seiler et al., 2022). While JSBACH slightly outperforms CLM5.0 for GPP and carbon stock benchmarks, CLM5.0 in turn outperforms JSBACH on evaporative feedbacks reflected in the surface latent heat flux (Seiler et al., 2022).

For assessing the statistical coupling between climate and vegetation productivity extremes, we leverage gridded FLUXCOM GPP data as an observation-based benchmark (Tramontana et al., 2016; Jung et al., 2020). FLUXCOM, specifically the FLUXCOM+RS+METEO product, combines machine learning techniques, satellite observations, and meteorological data from reanalyses to propagate localized FluxNet eddy covariance measurements into a spatially resolved product (0.5° by 0.5°) (Jung et al., 2020). FLUXCOM is available with different reanalysis datasets used to upscale point measurements. We employ

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

the ERA5 version of FLUXCOM (Hersbach et al., 2020; Muñoz Sabater, 2019), because, unlike the other reanalyses, ERA5 provides volumetric soil water content in addition to surface temperature and precipitation. Temperature, precipitation, and soil moisture in ERA5 have been constrained by in-situ and remote sensing observations (Hersbach et al., 2020). We choose soil moisture as the primary indicator for moisture, as it directly reflects the water available to the plants (McColl et al., 2022).

FLUXCOM's mean GPP estimate was extensively tested against satellite observations, atmospheric inversions, and model datasets (Jung et al., 2020). Seasonal GPP fluxes in FLUXCOM agree well with observational constraints on water stress, despite deficiencies of soil moisture remote sensing data and FLUXCOM's tendency to underestimate interannual variability (Jung et al., 2020). Because FLUXCOM does not explicitly account for CO_2 fertilization effects on GPP, long-term CO_2 trends are not consistently realistic, especially in the late 1970s to early 1990s (Jung et al., 2020). However, since we are focusing on the statistical link between climatic compound events and GPP extremes in general and not on its state-dependency, this shortcoming is not limiting our assessment.

Our analysis covers the years 1979 to 2014, constrained by the ERA5 data used to produce the FLUXCOM dataset (starting in 1979) and the termination of MPI-ESM historical simulations (at the end of 2014). We draw on daily climate data (surface air temperature, soil moisture, precipitation), but GPP is available only at monthly resolution. During preprocessing, all data sources and variables are trimmed to this time interval and remapped to the spatial resolution of MPI-ESM1.2-LR.

3 Methods

We aim to quantify statistical relations between temperature–moisture extremes and extremes in vegetation productivity in FLUXCOM, CESM2, and MPI-ESM data. The analysis workflow comprises four steps: 1. Pre-processing (Sect. 2), 2. computation of temperature–moisture compound and vegetation productivity extremes (Sect. 3.1), 3. computation of co-occurrence rates for the extremes (COR, Sect. 3.2), 4. significance testing (Sect. 3.3). Significant CORs reflect the strength of the statistical coupling between temperature–moisture and vegetation productivity extremes and indicate characteristic response times to plant-physiologically demanding and beneficial conditions (Sect. 3.4).

3.1 Definition of climatic compound and GPP extremes

After removing the seasonal trend, we define extreme events based on monthly thresholds and local extreme conditions. We deliberately keep the anthropogenic CO_2 trend in the data, because we are mainly interested in the mechanistic link between temperature–moisture and GPP extremes and not on potential state dependencies of this link. Preliminary tests indicated that artifacts from detrending outweigh the benefits of reducing non-stationarity.

We determine extreme conditions based on the univariate monthly event distributions of daily-resolved temperature and soil moisture, and monthly-resolved GPP (Fig. 1a). For each variable, we define the 90th and 10th percentiles as thresholds for

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

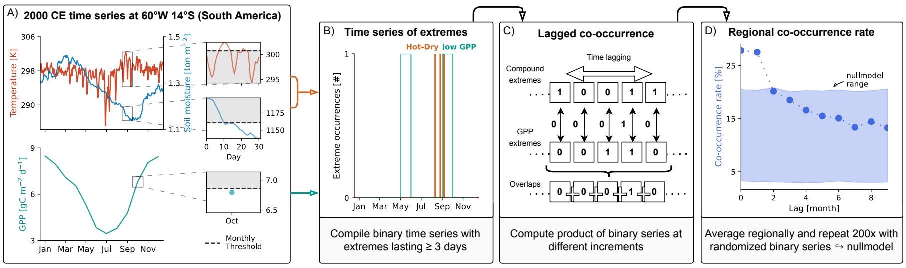

Figure 1. Definition and link of compound temperature–moisture events with GPP extremes for the example of hot–dry compound events followed by low GPP extremes at 60 °W 14 °S, which is part of the South-American monsoon reference region (SAM; Iturbide et al., 2020). Extreme conditions in grid cell surface air temperature, soil moisture, and GPP are defined separately for each variable with respect to local monthly quantile thresholds and the period from 1979 to 2014 (a). Threshold values and whether exceedance or undercutting is evaluated depend on the type of compound extreme. For a temperature–soil moisture compound event, extreme conditions must occur in both variables and persist for at least three consecutive days. The outcome of this event detection is two binary time series of GPP and compound extreme occurrences (b). If a compound and a GPP extreme occur with lag τ ($\tau \in [0,12]$ months), they are said to overlap (c). Using this (lagged) co-occurrence, dividing by the number of compound extremes, and averaging across regions yields the regional (lagged) co-occurrence rate (COR), here, for the SAM region (d). Because of the difference in resolution, a single GPP extreme can overlap with multiple daily temperature–precipitation compound events. After normalization by climate–extreme occurrences, the COR ranges from 0 to 1. The COR computation is repeated with 200 randomized extreme occurrence time series to compute the central 95% of the null model's probability distribution (see Sect. 3.3).

hot/cold, wet/dry, and high GPP/low GPP extremes, respectively. Those percentile thresholds allow the identification of locally extreme conditions relevant for local vegetation. Unique percentile thresholds for each location and dataset ensure we correctly interpret different interannual variability structures in the datasets (Jung et al., 2020). Vegetation productivity is modulated by combined and integrated temperature–moisture conditions. Therefore, we combine the univariate temperature and soil moisture extreme conditions and identify compound extreme conditions if the respective thresholds are crossed simultaneously (Fig. 1a). This definition of compound extremes is independent of the data sources and simple to calculate, even for distributions with multi-modality, zero inflation, and heavy tails, making it a robust measure of compound extremes (Hao et al., 2022). Short-lived compound extreme conditions, lasting for some hours to a couple of days, have minimal physiological impact on vegetation productivity. Therefore, we require temperature–moisture compound events to last for a minimum of three consecutive days.

The analysis primarily relies on soil moisture to identify moisture extremes. However, we repeat the entire analysis using precipitation to test robustness and mechanistic implications. Precipitation is indirectly linked to plant-available moisture, but directly indicates extreme atmospheric conditions. It also reduces the effective integration timescale of extreme hydrological conditions. For precipitation, we restrict the event population to non-zero values before determining the threshold for low-precipitation extremes, yet retain all time points in the analysis. This restriction is necessary to address regions with few precipitation events.

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

3.2 Co-occurrence rate (COR): measure of statistical coupling between temperature–moisture and GPP extremes

We compute the co-occurrence of both types of extremes with varying lag from 0 to 12 months to quantify the statistical relation between compound temperature–moisture events and GPP extremes. The occurrence of compound and GPP extremes yields two binary event time series, which contain 1 if conditions are extreme at a given time step, and 0 otherwise (Fig. 1b, c). Then, the co-occurrence of compound climate and GPP extremes at the same time is given by the product of both time series (Fig. 1c). To compare the daily temperature and soil moisture data with the monthly GPP fluxes, we sum all compound extreme occurrences within a month and compare the time series at monthly resolution. Therefore, several compound events can overlap with one vegetation productivity extreme, such that the values in the co-occurrence time series may exceed 1 at an individual time step.

Since vegetation productivity can react to changing climatological conditions with a delay, we extend the computation to cover non-zero time lags to include GPP extremes happening in the 12 months after extreme temperature–moisture conditions (Fig. 1c). For the resulting co-occurrences, a value of 1 indicates that a GPP extreme follows a compound climate extreme with lag τ. We sum the co-occurrences at lag τ for the whole time series and normalize the sum by the total count of compound extremes, yielding the *co-occurrence rate* $\mathrm{COR}(\tau)$ (Fig. 1d). The normalization ensures a COR between 0 and 1. The higher the COR, the higher the share of GPP extremes that follow after a compound climate extreme at a specific lag. For $\mathrm{COR}(\tau) = 1$ a GPP extreme would follow after every compound extreme with a lag τ, while for $\mathrm{COR}(\tau) = 0$ there would be no extreme GPP response.

3.3 Significance testing

We implement a two-sided randomization test to extract significant links between climatic precursors and GPP extremes. We construct a null model for the null hypothesis that the observed COR occurs by chance by shuffling monthly time blocks, effectively randomizing each event time series. With a block size of one month, we assume no persistent memory in the null model. Increasing the block size to several months has a negligible impact on the results (not shown). 200 null model realizations proved sufficient for the resulting distribution and significance thresholds to converge approximately (Fig. S1, Supplementary Information). We compute the CORs for each randomized time series in the ensemble, for each grid cell, and for each positive time lag between a compound event and a GPP extreme.

If the COR for the original data falls outside the central 90% of the null model probability distribution, we reject the null hypothesis (two-sided test) and consider the statistical link of GPP extremes and compound extremes to be significant. If the COR lies in the upper 5%, the occurrence of a compound event appears to significantly increase the likelihood of a subsequent GPP extreme. In turn, if the COR is below the null model's 5th percentile, the occurrence of a compound event significantly reduces the likelihood of a subsequent GPP extreme.

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

3.4 Interpreting the co-occurrence rate through its peak height and timing

We use the maximum offset of the COR to the null model ("*peak height*") to quantify the coupling strength between compound and GPP extremes in the observation-based and model data. The corresponding time lag ("*peak timing*") then describes the characteristic time for GPP fluxes to reach extreme values in response to temperature–moisture extremes. This response time does not necessarily imply that a GPP extreme will always or exclusively occur within that period, but that it is most likely for a GPP extreme to occur at that time. Grid cell-level pattern correlations of peak height and timing provide a measure of spatial agreement between modeled and observed responses.

In summary, co-occurrence rates quantify the strength and timing of the statistical relationship between GPP extremes in the succession of temperature–moisture extremes. A significant COR implies that this relation is stronger than expected by chance. Depending on the context, significant statistical links either indicate an increase or a decrease in the likelihood of a GPP extreme following a climatic extreme event. To comprehensively analyze the relationships between compound events and GPP extremes, we separately analyze the coupling strength and timing for all combinations of hot/cold, dry/wet, and low/high GPP extremes.

4 Results

We investigate the relationships between temperature-precipitation compound extremes and GPP flux extremes in CESM2, MPI-ESM1.2, and FLUXCOM observations. Based on the co-occurrence rate of compound events and GPP extremes, we quantify regional variations in the mechanistic response and the agreement between simulation and observation data. We first discuss the characteristics of the statistical relationship between cold–wet compound events and high GPP extremes as an example (Sect. 4.1). Subsequently, we aggregate the results at regional and global scale to systematically evaluate model biases across all event combinations (Sect. 4.2). Finally, we explore mechanistic regimes that could explain the similarities and differences between the data sources (Sect. 4.3). Note that we use shorthand expressions: "hot–dry–low" would, for example, indicate co-occurrences of hot–dry climatic compound events followed by low GPP extremes.

4.1 Significant local links between cold–wet compound events and high GPP extremes

To provide an intuitive understanding of the COR as a measure of statistical coupling, we first explore how often cold–wet events are followed by very high vegetation productivity (cold–wet–high). The peak height of the co-occurrence rate (Fig. 2, left column) reflects the maximum coupling strength between GPP extremes and compound events. The corresponding peak timing (Fig. 2, right column) provides a characteristic response time of GPP to compound extreme conditions in temperature and soil moisture (Sect. 3). For cold–wet–high, all datasets exhibit distinct regional and latitudinal patterns in peak height and timing (Fig. 2). Mid- and low-latitude semi-arid to arid regions exhibit the strongest and spatially most extensive link between cold–wet events and high GPP extremes, reflecting the positive impact of additional moisture and lower evaporative pressure

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

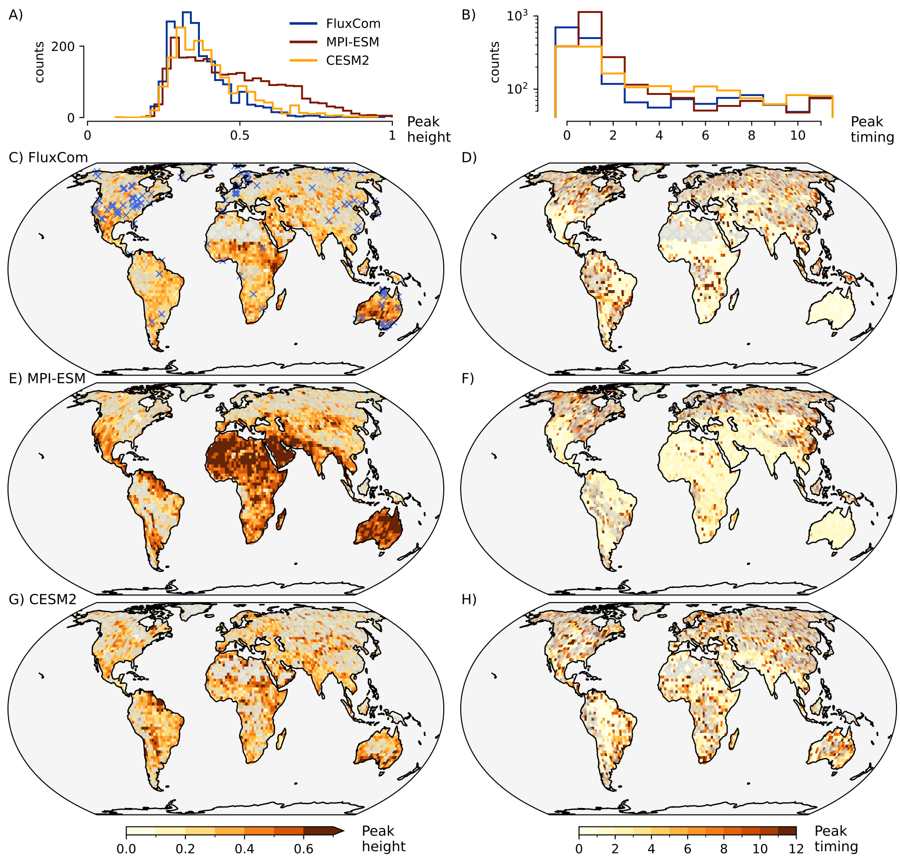

Figure 2. Highest share of co-occurrences (peak height) between cold–wet compound events and high GPP extremes (left column) and the corresponding timing (right column) in FLUXCOM and simulations. Histograms in (a), (b) show all significant grid cells with respect to the null model. Note that counts of peak timing are given on a logarithmic scale in (b). Gray shading in (c)–(h) masks these insignificant data points. Crosses in (c) mark the FluxNet observation sites contributing to the FLUXCOM dataset. The 52 Central European FluxNet sites are depicted as a single thick cross.

in moisture-limited environments. In contrast, there are almost no significant statistical links at high latitudes. The significant coupling is particularly pronounced in MPI-ESM (Fig. 2a, e). Although the spatial patterns are often more pronounced in the simulations, they are largely similar to those in FLUXCOM. This can be seen in pattern correlation coefficients between the datasets: 0.48 (FLUXCOM with MPI-ESM), 0.39 (FLUXCOM and CESM), and 0.44 (CESM and MPI-ESM). These correlations give an indication of the average local agreement of spatial patterns between our datasets, regardless of systematic biases (Fig. 3). For peak height, pattern correlations are medium to high for all combinations of compound events and GPP extremes, ranging from 0.39 to 0.76, with hot–wet and cold–dry ranking higher than hot–dry and cold–wet events (Fig. 3a). Both ESMs largely agree with FluxCom, more so than with each other.

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

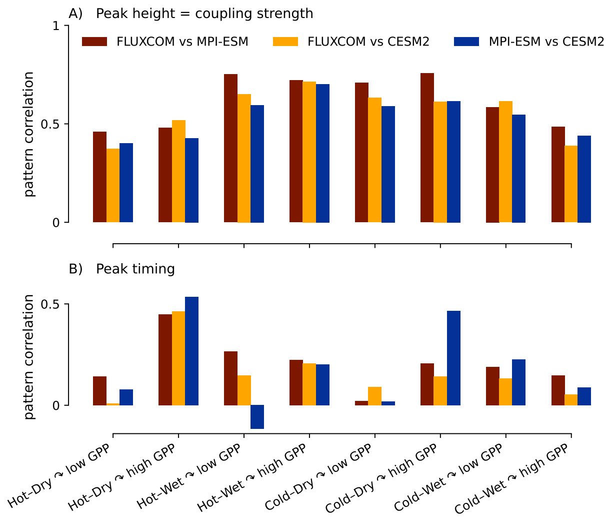

Figure 3. Global spatial pattern correlation coefficients of COR peak heights (a) and peak timings (b) for all combinations of extremes and datasets. Note that pattern correlations were only computed for grid cells that are significant in both datasets. The share of overlapping patterns is generally high, but varies by extreme type and dataset (Fig. S2).

Where the statistical link between cold–wet events and high GPP extremes is strong, the peak timing is typically small, ranging from zero to two months. Notable exceptions are scattered across South America and Africa in FLUXCOM and CESM2, and Australia in CESM2, where peak timings lie between 4 and 12 months. There are some local model-specific discrepancies. For example, MPI-ESM simulations show lead times of 6 to 12 months in northern North America, whereas CESM2 tends to produce occasionally higher lead times across all regions (Fig. 2b, h). Additionally, the spatial patterns for peak timing tend to be more noisy than for peak height. This is also reflected in low pattern correlations of the peak timing (Fig. 3b). Across event types, average local agreement is mostly low for peak timing (Fig. 3b), unlike for peak height. As an exception, across all dataset combinations we consistently find that hot–dry–high extremes behave similarly, with pattern correlations between 0.45 and 0.54.

4.2 Regional variations in coupling strength and timing between types of extremes

To systematically assess global and regional biases between the datasets, we aggregate the spatially resolved results for peak heights and timings across all combinations of extremes. We use the IPCC AR6 land reference regions, which decompose the global land area into 46 regions with coherent climatic conditions (Iturbide et al., 2020).

As an example, Northeastern South America (NES) shows a rapidly increasing co-occurrence rate of hot–dry or cold–wet compound extremes (Fig. 4a, b). In this region, which is characterized by a tropical non-seasonal climate, low moisture-high-temperature conditions appear to have an up to eight-month legacy impact on vegetation productivity in MPI-ESM and CESM.

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

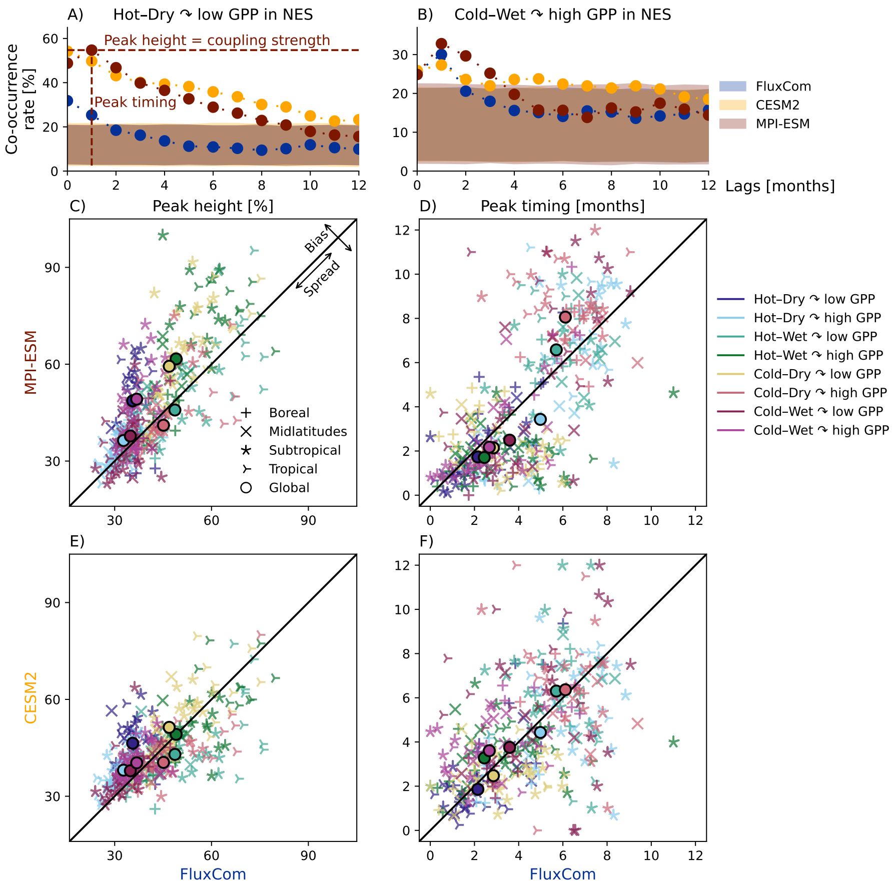

Figure 4. Agreement in regionally averaged co-occurrence rate and timing between FLUXCOM and the MPI-ESM and CESM model simulations. Average co-occurrence rates in North Eastern South America (NES) for hot–dry compounds followed by low GPP extremes (a) and cold–wet events followed by high GPP extremes (b). Horizontal bands represent the 95% range of the null model realizations. Peak height (c,e) and timing (d, f), as illustrated in (a), compared between FLUXCOM data and the MPI-ESM (c, d) and CESM simulations (e, f) for all combinations of compound climate events and GPP extremes (color-coded). Only significant co-occurrences are shown. Marker styles categorize IPCC AR6 land reference regions (Figure S5). Diagonal lines indicate perfect agreement between datasets. Global mean agreement for any extreme type combinations (marked by different colors) is marked by circles.

In FLUXCOM, the strength of this effect is lower, and it declines more rapidly (Fig. 4a). By contrast, cold conditions associated with excess moisture appear to support vegetation productivity more immediately, and consistently between observations and simulations (Fig. 4b). While the peak height and timing vary between regions, datasets and variables, the share of co-occurrences gradually declines in regions where a notable initial peak above the significance threshold was found. In NES for example, both models show a peak height of around 55% for hot–dry–low, whereas the FLUXCOM data peaks around 31% (Figure 4a). In contrast, all data sources agree on a lower peak height at around 30% for cold–wet–high (Figure 4b).

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

4.2.1 Coupling strength

Overall, regional and especially globally averaged couplings between climate and vegetation extremes appear to be similar in observation-based and model data (Fig. 4c, e). All data sources consistently suggest the highest statistical association of hot–wet–high and cold–dry–low extremes. However, these combinations, along with hot–wet–low extremes, also exhibit the greatest regional spread in coupling strength. In FLUXCOM, globally averaged hot–wet–low and cold–dry–low extremes exhibit coupling strengths similar to those of their high GPP counterparts. In contrast, in MPI-ESM, all combinations of dry events followed by low GPP extremes and wet events followed by high GPP extremes are biased high compared to FLUXCOM. Therefore, unlike for FLUXCOM, all of these combinations rank higher than hot–wet–high and cold–dry–high. Global mean statistical coupling strengths in CESM2 do not differ as much between event types as in MPI-ESM. However, this model also exhibits a positive mean bias for hot–dry–low extremes.

Biases observed on the global scale are more pronounced regionally, where both models tend to overestimate statistical coupling strength between temperature–moisture and GPP extremes compared to FLUXCOM. This regional bias is particularly evident in MPI-ESM for hot–dry–low, cold–wet–high, cold–dry–low, and hot–wet–high extremes. In CESM2 simulations, the direction of biases is similar to MPI-ESM, but their amplitude tends to be smaller.

Cold–dry and hot–wet temperature–moisture extremes appear to have a larger regional spread than hot–dry and cold–wet extremes. The pronounced latitudinal structure likely contributes to this as hot–wet extremes exhibit the highest coupling strength in the tropics and subtropics (e.g., hot–wet–high mean (standard deviation): 56 (19)% in MPI-ESM, 47 (12)% in CESM2, and 48 (13)% in FLUXCOM) compared to the higher latitudes (41 (8)% in MPI-ESM, 41 (6)% in CESM2, and 40 (5)% in FLUXCOM). To a slightly lesser extent, this pattern holds for cold–dry extremes (e.g., for cold–dry–low in FLUXCOM, 44 (12)% in the subtropics and tropics and 38 (4)% in high latitudes).

4.2.2 Timing

The capacity of vegetation and soils to buffer extreme conditions, and the type of growth-limiting regime (energy, moisture, nutrients) influence the duration over which vegetation can sustain extremes and the time it takes for the GPP flux to respond to them. This is reflected in the COR's peak timing. Global model-observation biases are small, between -1.6 months (model leading FLUXCOM data) and +1.9 months (model lagging FLUXCOM) for MPI-ESM simulations and between -0.6 and +0.9 for CESM2 (Fig. 4d, f). Timing is short for immediate responses when buffering capacity is low, and supplies are limited. MPI-ESM tends towards faster response times than FLUXCOM for all but the two slowest cases (hot–wet–low and cold–dry–high events), highlighting how an immediate response can occur both for extreme events that damage and benefit vegetation productivity. Hot–dry compound events followed by low GPP extremes exhibit the shortest response time across the datasets at 1.7 to 2.2 months. Hot–wet–high, cold–dry–low, and cold–wet with low or high GPP events have similarly short response times, ranging from 1.7 to 3.8 months. The event combinations with the slowest response in all datasets are hot–wet–low (MPI-ESM:

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

6.6, CESM2: 6.3, FLUXCOM: 5.7) and cold–dry–high extremes (MPI-ESM: 8.1, CESM2: 6.4, FLUXCOM: 6.1). Such long response times can occur when extremes expose supply limitations after prolonged buffering or due to seasonality, or when they stimulate supplies for a future growing season.

Regionally, there are notable differences in model–observation biases and a substantial spread across the regions. Some combinations of events exhibit latitudinal patterns in these biases. For example, for cold–dry–high extremes, biases are mostly positive in the tropics and subtropics, indicating that the model response there is slower than that shown by FLUXCOM. In the mid-latitudes and boreal regions, on the other hand, the models tend to respond faster for this event combination. When cold–dry events are followed by low GPP extremes, the opposite latitudinal pattern tends to appear.

4.3 Regionally opposing patterns of vegetation productivity as a consequence of changing extreme moisture-temperature relationships

To comprehensively differentiate the spatial response of vegetation productivity to temperature–moisture extremes, we classify the effects of these extremes into three classes according to their GPP response: (i) damaging to vegetation productivity due to an increase (reduction) of the likelihood of low (high) GPP extremes, (ii) beneficial to vegetation productivity due to an increase (reduction) of the likelihood of high (low) GPP extremes, or (iii) damaging and beneficial at different lags after the compound extreme. The latter is included only for completeness, as it occurs rarely. Fig. 5 shows the spatial patterns of the classified effects of compound extremes on vegetation productivity for the FLUXCOM dataset across all combinations of temperature, soil moisture, and GPP extremes. The patterns are generally similar for the model data (Fig. S14, S15).

Under extraordinarily wet conditions, temperature and moisture limitations materialize differently in the productivity response. Exceptionally high temperatures stimulate productivity in the boreal and some temperate regions (Fig. 5a), whereas low temperature extremes damage productivity in the same regions (Fig. 5f). However, under exceptionally wet conditions, other than in arid regions, the moisture supply and latent heat transport remain sufficient to outweigh the damaging potential of enhanced evaporation during hot extremes (Fig. 5a, h). Likewise, contrasting their influence under extremely dry conditions, low temperature extremes contribute to beneficial growth environments by attenuating evaporation in moisture-limited, low-latitude regions (Fig. 5c, e). This shift in evaporative balance is also why vegetation is less likely to experience low GPP extremes under cold–wet conditions in some (sub-)tropical regions (Fig. 5f).

In cold–dry extreme conditions, reduced evaporation in transition zones that outweighs the moisture shortage could explain beneficial effects on high GPP extremes in some subtropical and midlatitude regions (Fig. 5d). Dry conditions have adverse effects elsewhere, both by reducing the likelihood of exceptionally high productivity and increasing the likelihood of extraordinarily low productivity for short lags (Fig. 5b, e, g). Here, a lack of moisture supply dominates the response in mid- to low-latitude regions under hot–dry conditions (Fig. 5e, g). In contrast, a shortened growing season or unfavorable growth con-

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

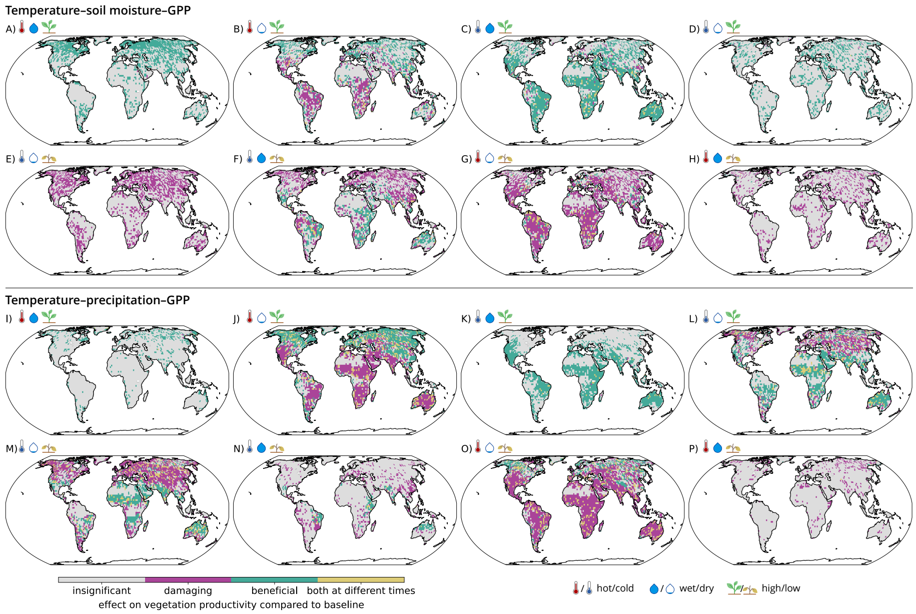

Figure 5. Categorization of enhanced and muted co-occurrence rates between temperature–soil moisture and GPP extremes (a-h) and between temperature–precipitation and GPP extremes (i-p) in the FLUXCOM dataset. Panels display the categorization for the different types of compound events from hot–wet–high (a, i) through hot–wet–low (h, p) as indicated by icons and aggregated for all time lags. Co-occurrence rates are diagnosed as "damaging"/"beneficial" when they exceed/undercut the null model's 95% range for low GPP extremes and undercut/exceed the range for high GPP extremes. This categorization presumes that an increased/decreased statistical coupling between temperature–moisture extremes and low/high GPP extremes corresponds to less favorable conditions for plant growth and carbon uptake. Vice versa, a decreased/increased coupling with low/high GPP extremes reflects more favorable conditions for vegetation productivity. In rare cases, co-occurrence rates both exceed and undercut the null model range at different lags. This behavior is categorized as "both at different times."

ditions caused by low temperatures explain the high-latitude response (Fig. 5e).

We find consistent and sometimes mirrored patterns between regions and types of extremes (Fig. 5). This mirroring can be found for all combinations of extremes, when contrasting hot and cold, dry and wet, and high and low GPP extremes. For example, hot–dry extremes significantly increase the likelihood of high GPP events in the Northern Hemisphere's high and mid latitudes (Fig. 5b), indicating beneficial conditions for plant growth with increased warmth. At the same time, hot–dry events reduce the likelihood of high GPP extremes in the subtropics and tropics, where a lack of moisture supply damages vegetation productivity. Vice versa, cold–wet events create less favorable growth conditions in mid- and high latitudes, indicated by an increased likelihood of low GPP extremes (Fig. 5f). In low latitudes, cold–wet extremes reduce the likelihood of extremely low plant productivity, particularly in rainforest regions, thereby having a positive effect on plant growth. Combining the two (nearly) complementary patterns illustrates how vegetation productivity benefits from one type of temperature–moisture event,

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

but gets damaged by an opposing type of event.

Juxtaposing cold–wet–high extremes (Fig. 5c) and hot–dry–low extremes (Fig. 5g) shows that cold–wet events increase the chance for high GPP extremes in the same regions, effectively associated with benefiting plant growth. In contrast, hot–dry events increase the likelihood of low GPP extremes, reducing vegetation productivity, particularly in mid- and low latitudes. Similar patterns exist for cold–dry–low extremes (Fig. 5e) with damaging effects east in the Andes and throughout Eurasia and North America, in contrast to the beneficial effects of hot–wet–high extremes in similar regions (Fig. 5a).

Following from the regional agreement in peak height and timing, the response patterns of vegetation productivity are similar in the three datasets, albeit with regionally different biases and extent in the models compared to FLUXCOM (Fig. S14, S15). In CESM2, productivity benefits are more pronounced for hot–dry–high extremes in the high latitudes compared to FLUXCOM (Fig. S14). CESM2 also overemphasizes damages in the mid- and low latitudes when low GPP extremes follow hot–dry and cold–dry events, reducing the latitudinal gradient for cold–dry–low extremes. MPI-ESM responds analogously for hot–dry–high extremes, but overestimates spatial extent and coherence compared to FLUXCOM (Fig. S15). It matches FLUXCOM patterns in the high latitudes somewhat better than CESM2, particularly in Eastern Siberia. Similar observations hold for cold–wet–low. Here, MPI-ESM matches FLUXCOM's response patterns reasonably well, but again overestimates spatial extent and coherence. CESM2, by contrast, tends to underestimate the spatial extent of response patterns for this event combination (Fig. S14). Notably, compared with FLUXCOM and CESM2, MPI-ESM also predicts a much stronger response in and around the Sahara region for low GPP extremes following dry events and for high GPP extremes following wet events.

Overall, at an aggregated level, a lack of moisture is predominantly detrimental to vegetation productivity. In this case, modulating effects of temperature extremes outweigh moisture shortages only at a few locations in the high latitudes for hot–dry–high extremes and in the transition zones for cold–dry–high extremes. On the other hand, high temperature extremes mostly benefit plant growth when moisture is abundant, particularly in the high and mid-latitudes. In turn, (extreme) moisture abundance will generally support vegetation productivity, unless temperature limits access to water in boreal zones or confines the evaporative water budget in some subtropical and temperate climatological transition zones.

4.4 Evaporative effects and the moisture integrating time modulate the vegetation response to climatic extremes

Almost all plants take up necessary moisture from the soil, most of which comes from rainwater. Precipitation, as an alternative indicator of moisture availability, reduces the effective integration time of the climatic compound events compared to soil moisture. It comes with lower latent temperature control on wet/dry extreme conditions through evaporative effects, as these affect soil moisture but not (directly) precipitation. As a result, responses to hot–wet–low, cold–dry–high, cold–dry–low, and hot–dry–low extreme conditions qualitatively differ for precipitation across all datasets (Fig. 5).

Patterns of advantageous and adverse conditions for vegetation productivity are often opposed for contrasting temperature–soil moisture extreme conditions. For example, plant growth benefits from hot–wet–high extremes, but suffers from cold–dry–low extremes (Fig.5a, e). The mirrored patterns disappear with temperature–precipitation extremes because soil moisture and

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

precipitation emphasize different mechanistic aspects. With precipitation, typically only one of the complementary extreme conditions exhibits a coherent spatial pattern, as is the case for cold–dry–low (Fig.5m), whereas hot–wet extremes do not increase the likelihood of high GPP extremes (Fig.5i).

Response maps for temperature–precipitation extremes also no longer show the globally uniform response of high and low GPP extremes to cold–dry extremes (Fig. 5l, m). Instead, they exhibit strong dipoles. While vegetation productivity in the tropics benefits from cold–dry temperature–precipitation extremes, these events have damaging effects in the mid- and high latitudes. Similarly, the almost uniform damaging effect of hot–dry conditions shifts toward a dipole pattern (Fig. 5j, o). In this case, hot–dry temperature–precipitation events support growth conditions in the high latitudes. On the other hand, there is barely any significant link between compound hot–wet temperature–precipitation events and high/low GPP extremes (Fig. 5i, p), unlike for soil moisture with a substantially increased likelihood of high GPP extremes in the mid- and high latitudes (Fig. 5a). More subtle differences between the two moisture indicators occur for hot–dry and cold–wet events. While statistical coupling with precipitation as an indicator is lower for hot–dry–low and higher for cold–wet–low extremes compared to soil moisture (Fig. S13), response patterns are more extensive and coherent in space, particularly for hot–dry extremes that benefit productivity in the high latitudes (Fig. 5j, o).

Despite the mirrored patterns occurring only for temperature–soil moisture extremes, significant beneficial and detrimental patterns for vegetation productivity often concur well between temperature–soil moisture and temperature–precipitation extremes (Fig.5). It is also notable that precipitation-based CORs are substantially higher under wet than dry conditions in all datasets, and biases between observations and simulations are reduced (Fig. S13c, d). However, global-scale model response times tend to be biased high with precipitation, i.e., the peak heights are reached later in the models than in FLUXCOM (Fig. S13e, f). This bias is most pronounced for event combinations with short characteristic response times. Biases for peak times are also larger when using precipitation and the regional spread increases for all datasets. By contrast, for peak heights model biases are reduced when using precipitation instead of soil moisture.

In contrast to temperature-soil moisture extremes, patterns including temperature–precipitation extremes are inherently consistent, i.e., the same type of temperature–precipitation extreme creates damaging (beneficial) effects through changes in both, low and high GPP extremes, or affects only one of the extreme types. In contrast, there are several occasions where the same type of temperature–soil moisture extreme is both damaging by lowering (increasing) the likelihood of high (low) GPP extremes and beneficial by reducing (increasing) the likelihood of low (high) GPP at the same location.

5 Discussion

5.1 Similarity between observations and model data

Both models appear capable of capturing essential dependencies that influence vegetation productivity a few months into the future. Spatial aggregation impacts the degree of consistency of the statistical link between extremes in temperature–moisture and

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

vegetation productivity between MPI-ESM, CESM2, and the observation-based data from FLUXCOM. Models and FLUXCOM agree most at the global and regional scale, despite some model-specific biases (Sect. 4.2). The agreement decreases at the grid-cell level. Aggregating localized statistical links leads to more reliable estimates from the ESMs. This highlights the complexity of developing and calibrating state-of-the-art ESMs that represent local climate–biosphere interactions beyond the mean state well (Reichstein et al., 2013; Villalobos-Herrera et al., 2021).

Contrasting the qualitative agreement, there is a substantial spread in regional performance, with model-specific biases. Raw, not bias-corrected ESM data can be considered reliable only in some regions. For example, there is a clear latitudinal gradient in local model performance. Both MPI-ESM and CESM2 tend to perform better in the mid-latitudes and boreal regions compared to the (sub-)tropics (Fig. 3, S9). Much of this regional spread stems from both ESMs systematically overestimating the strength and spatial coherence of the statistical coupling between climate and vegetation extremes on the local level. This partly carries over to regional aggregations (Sect. 4.2). Spatially variable climate variability could influence the regional spread, but is unlikely to cause substantial biases over a 36-year analysis period.

ESMs mostly underestimate the intensity of daily precipitation extremes and underestimate the dry day frequency in comparison to observations (Seneviratne et al., 2021; Norris et al., 2021). It may appear counterintuitive that climate extremes pre-condition vegetation towards productivity extremes more strongly in the ESMs than in the observational data. However, the two phenomena are not directly linked to overestimated vegetation–climate coupling, which may well be due to deficiencies in the land models, particularly the soil. Biases in both models occur under extremely dry conditions. This points to oversensitive evaporative feedbacks, exaggerated drainage, or too low soil water holding capacity, leading to accelerated evaporation of soil moisture. Higher soil moisture decay rates in both ESMs compared to ERA5 support this interpretation of exaggerated dry-out (Figure 6). MPI-ESM shows additional biases under extremely wet conditions, which could arise from an overly simplified representation of precipitation input to soil water pools, missing moisture routing above and below the surface, or misaligned soil water-holding capacity, which is compatible with other model benchmarking results (Seiler et al., 2022). The overestimated rates of soil moisture change in MPI-ESM are not just connected to biases in the coupling between climatic and vegetation extremes, but also translate into mean GPP biases (Fig. S16). This could indicate an influence of the biased extreme representation on the mean state biases, given the strong influence of productivity extreme events on terrestrial carbon stocks (Zhou et al., 2019).

Intriguingly, the ESMs do not overestimate coupling strengths when considering temperature–precipitation rather than temperature–soil moisture compound extremes. This could indicate that biases in the soil moisture–vegetation relationship may be compensated by biases in precipitation–soil moisture relationships. Alternatively, temperature–precipitation extremes could drive GPP extremes through different mechanistic links than temperature–soil moisture extremes. This interpretation would be supported by the differing orders of global mean peak heights and timings between extreme types when using pre-

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

cipitation rather than soil moisture (Figure 3).

5.2 Potential biases of observations and models

The observation-constrained FLUXCOM data aggregates uncertainties from the upscaling process and underlying FluxNet site data (Jung et al., 2020). FLUXCOM also underestimates interannual GPP variability and exhibits partly unrealistic GPP growth rates (Jung et al., 2020). Both of these shortcomings could lead to FLUXCOM biases being mistaken for model biases in our comparison, if they alter the GPP event distribution. However, the comparison between precipitation-based and soil moisture-based estimates indicates a mechanistic explanation for the largest part of the model–observation differences. Deficiencies in the HTESSEL soil model used to produce ERA5 soil moisture estimates (Balsamo et al., 2015) could also enhance perceived ESM biases. However, HTESSEL performs well with respect to observations (Balsamo et al., 2015), and assimilation of soil moisture observations from in-situ and satellite observations in ERA5 should further reduce soil moisture biases (Hersbach et al., 2020). Given that the observed biases have model-specific characteristics, the mechanistic shortcomings are likely located on the side of the ESMs.

Both CESM2 and MPI-ESM show substantial regional spread in the statistical relationships between hot-wet and high GPP extremes. MPI-ESM overestimates the strength and underestimates the lag between climatic driver and vegetation response, whereas CESM2 underestimates the strength and overestimates the lag. At the same time, both models tend to exaggerate the spatial extent of significant coupling, particularly within (sub-)tropical regions (Fig. S6). Both suggest that GPP extremes are significantly affected by hot–wet compound events in around half of the (sub-)tropics (Fig. S3), which is more than twice the area based on FLUXCOM/ERA5 (Fig. S3). The MPI-ESM land component JSBACH was tuned for a high moisture sensitivity of vegetation in the Sahel to study the phenomenon of the African humid period in the late Pleistocene and early Holocene (Groner et al., 2018). Sensitive, perhaps oversensitive, vegetation productivity in the (sub-)tropics is most likely also the reason for the large regional spread observed in both ESMs for cold–dry compound events that lead to low GPP extremes and for high GPP extremes following cold–wet events in MPI-ESM (Fig. S3, S7, S8). However, FLUXCOM tends to overestimate tropical mean GPP and net ecosystem exchange, partly due to weaknesses in FluxNet site data (Jung et al., 2020). Although this bias does not necessarily affect the GPP event distribution in a way that propagates into the COR, it is crucial to better constrain observations and models in this pivotal region for the global carbon budget.

Our analysis builds on CORs to quantify how well ESMs relate climatic and productivity extremes. The COR imposes no constraints on causality and latent variables. This is advantageous for our impact-oriented analysis. However, it will not necessarily identify if a model performs well for the wrong reasons. To assess causality, rather than correlation, the COR would need to be applied incrementally to driver and response variables. This would result in large data requirements for intermediate state variables that are rarely available, especially on the observation side.

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

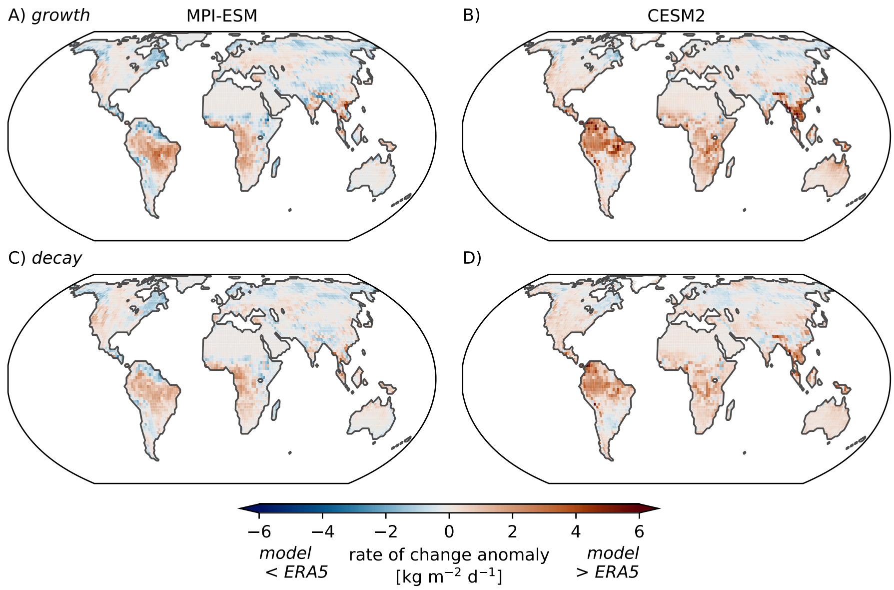

Figure 6. Divergent change rates in soil moisture between ERA5 and the ESMs. Panels depict smoothed annual maximum rates of change averaged over 1979–2014 as anomalies to ERA5 for increasing (a, b) and decreasing soil moisture (c, d).

5.3 Mechanisms

We find that both beneficial and detrimental preconditioning effects of temperature–moisture extremes for vegetation growth conditions manifest as coherent spatial patterns in the observation-based and model data (Fig. 5). The spatial response patterns provide mechanistic insights into the repercussions of temperature–moisture extremes on vegetation growth around the globe. On the one hand, it is evident that temperature–soil moisture extremes can benefit vegetation productivity through an increased likelihood of high GPP extremes in many regions. An exception are hot–dry events where extreme climatic conditions have damaging effects in the tropics and subtropics. On the other hand, climatic extremes almost always increase the likelihood of low GPP extremes, with harmful consequences for vegetation productivity. In this case, cold–wet extremes are the exception, as they reduce the likelihood of low GPP extremes in the (sub-)tropics, thereby benefiting plant growth conditions. These opposing responses suggest differing mechanistic links for high compared to low GPP extremes.

Our results suggest that process-based links have to be contextualized on an event-type and region-specific basis. In addition, seasonality is an important factor to consider, particularly in explaining varying peak timings across events and regions (Zscheischler et al., 2014a; Guan et al., 2015). Overall, the spatial patterns mostly reflect the dominant temperature limitation for growth in the high latitudes and moisture limitations in the mid- and low latitudes. For example, in the context of extremely dry but hot conditions, these limitations lead to increased likelihoods for beneficial high GPP extremes in boreal latitudes (Fig. 5b), either because they free up frozen water early in the growing season, or because of the growth promoting effect of

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

high temperature outweighing the relatively reduced moisture (Flach et al., 2021) (Fig. 3, 5). In contrast, hot–dry events are significantly linked to damaging effects, as they increase the likelihood of low and reduce the likelihood of high GPP extremes in moisture-limited regions (Fig. 5b,g). If insufficient moisture supply persists, it can severely degrade growing conditions (Green et al., 2019).

Compounding temperature–soil moisture extremes require soil moisture to integrate sustained atmospheric drought (high precipitation) and high (low) evaporation facilitated by high (low) temperature. At the same time, to promote extreme conditions, temperatures must remain relatively high (low) for an extended period, which is less likely than shorter temperature extremes. This is different for temperature–precipitation extremes, which only require temporal coincidence, but no build-up time. Soil moisture also has a prolonged memory, potentially causing legacy effects, although these are typically under-resolved in ESMs (Kolus et al., 2019; Anderegg et al., 2015). In contrast, atmospheric dynamics that drive precipitation extremes exhibit low temporal autocorrelation. These differences may partly explain why we observe more extensive and coherent spatial patterns with precipitation. Additionally, precipitation as an indicator of atmospheric drought and soil moisture content as a measure of soil drought partly represent different mechanisms of how plants experience drought (Harris et al., 2022; Wu et al., 2022), which could contribute to the different spatial patterns.

It would be worthwhile to expand our approach to additional moisture indicators and latent variables in the future, such as vapor pressure deficit or soil conditions. Our first-order approach can only implicitly account for the potential of compound extremes to cause persistent physical damage to vegetation from floods, fires, or other natural disturbances (Jain et al., 2022). In addition, changes in anthropogenic land use and land management practices, such as irrigation, can significantly alter physical surface properties and the moisture budget, affecting the likelihood of extremes (Findell et al., 2017). Future work could therefore also explore how both natural and anthropogenic disturbances shape the preconditioning of GPP extremes under climate change.

6 Conclusions

The interplay between temperature–moisture extremes and vegetation productivity will co-determine resilience and vulnerability of the land carbon sink and terrestrial ecosystems to future climate change. ESMs are essential tools for projecting this link between climatic and biosphere extremes and the resulting impacts on carbon stocks. While research efforts often focus either on extreme drivers, impacts, or on specific event types, we comprehensively benchmarked the statistical link between temperature–moisture and GPP extremes across all event combinations in two state-of-the-art ESMs against an observation-constrained dataset. We introduced a generalizable metric that can be translated to more complex networks of extreme event interactions. Expanding our analysis to comprehensive model-data and inter-model comparisons in the context of CMIP6 and CMIP7 would add confidence in the generalizability of our results.

While there is substantial regional variation in models and data, with spurious deviations in the lag between climatic drivers and GPP responses, there is reasonable agreement on the strength of the statistical relationship for regionally and especially globally aggregated results. We found that MPI-ESM and CESM2 locally overestimate the association between temperature–moisture and GPP extremes. We hypothesize that this discrepancy arises due to overestimated rates of change in soil moisture in the ESMs. This may perhaps be due to unrealistic evaporative feedbacks, exaggerated drainage, or inadequate effective water-holding capacity in the model soils. Assessments of legacy- and state-dependencies of this statistical coupling under climate change can build on general agreement found regionally and globally and the quantified local biases. However, as GPP integrates short-lived disturbances, biases in the statistical link can propagate into the annual mean GPP, as appears to be the case for MPI-ESM. The link between climatic and GPP extremes could potentially be exploited further as a proxy to test mean GPP biases beyond the immediate effects of atmospheric extremes.

Model and observation-based data indicated both adverse and advantageous effects of temperature–moisture extremes on vegetation productivity in different regions. There is justified and wide interest in extreme events that damage vegetation and ecosystems. However, as our results showed, beneficial impacts on vegetation can occur across all latitudes and cover a similar share of land as detrimental effects. This situation could change, either due to shifts in the frequency of different types of climate extremes or changes in the link between climate extremes and GPP extremes. Whether the beneficial impacts will persist under future climate change or evolve asymmetrically to an increasing likelihood of GPP decline should be further quantified and could be highly relevant for the strongly variable and interconnected net land carbon uptake.

Code and data availability. ERA5 data is available through the Copernicus Climate Change Service (C3S) Climate Data Store (CDS) (Copernicus Climate Change Service (C3S), 2019, 2023). Monthly data from the FLUXCOM initiative are available via the Max Planck Institute for Biogeochemistry data portal (Boenisch, 2020) upon free registration (Jung et al., 2020; Tramontana et al., 2016). CESM2 output (Version 20190312; Danabasoglu, 2019) and MPI-ESM1.2-LR output (Version 20210921; Wieners et al., 2019) are publicly available through the Earth System Grid Federation (ESGF), for example, on their server at DKRZ: https://esgf-metagrid.cloud.dkrz.de/search/cmip6-dkrz/. Analysis code, along with intermediate data, will be made available on Zenodo upon acceptance.

Author contributions. MA, EZ, NW, and KR conceptualized the project. MA, EZ, BG, and NW developed the methodology. BG developed the software with input from MA and EZ. BG and MA analyzed the data and performed the visualization. NW and KR acquired funding, and KR administered the project. MA and EZ wrote the original draft. All authors contributed to writing and editing the manuscript.

Competing interests. KR is an editorial board member of *Earth System Dynamics*. Besides, the authors have no other competing interests to declare.

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

Acknowledgements. MA acknowledges funding from the German Federal Environmental Foundation (DBU), the Landesgraduiertenförderung Baden–Württemberg, and the Cluster of Excellence (EXC 3121) TERRA – Terrestrial Geo-Biosphere Interactions in a Changing World funded by the German Research Foundation (DFG). EZ and BG were supported by the German Federal Ministry of Research, Technology, and Space (BMFTR), a Research for Sustainability initiative (FONA, www.fona.de) through the PalMod project (Grant number: 01LP2310A). KR is a member of the Cluster of Excellence TERRA and the Cluster of Excellence (EXC 2064): Machine Learning: New Perspectives for Science, also funded under Germany's Excellence Strategy. NW was supported by the Engineering and Physical Sciences Research Council (EPSRC, Grant number: EP/Z002591/1).

We acknowledge the FLUXCOM initiative for providing free access to their data, the CESM2 and MPI-ESM modeling groups for making their data available through the ESGF, and the Copernicus Climate Change Service (C3S) Climate Data Store (CDS) for providing the ERA5 data. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains. We thank Julia Brugger for helpful discussions on precipitation and soil moisture.

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

References

Anderegg, W. R. L., Schwalm, C., Biondi, F., Camarero, J. J., Koch, G., Litvak, M., Ogle, K., Shaw, J. D., Shevliakova, E., Williams, A. P., Wolf, A., Ziaco, E., and Pacala, S.: Pervasive drought legacies in forest ecosystems and their implications for carbon cycle models, Science, 349, 528–532, https://doi.org/10.1126/science.aab1833, 2015.

Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., Schwinger, J., Bopp, L., Boucher, O., Cadule, P., Chamberlain, M. A., Christian, J. R., Delire, C., Fisher, R. A., Hajima, T., Ilyina, T., Joetzjer, E., Kawamiya, M., Koven, C. D., Krasting, J. P., Law, R. M., Lawrence, D. M., Lenton, A., Lindsay, K., Pongratz, J., Raddatz, T., Séférian, R., Tachiiri, K., Tjiputra, J. F., Wiltshire, A., Wu, T., and Ziehn, T.: Carbon–concentration and carbon–climate feedbacks in CMIP6 models and their comparison to CMIP5 models, Biogeosciences, 17, 4173–4222, https://doi.org/10.5194/bg-17-4173-2020, 2020.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz-Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., and Vitart, F.: ERA-Interim/Land: a global land surface reanalysis data set, Hydrology and Earth System Sciences, 19, 389–407, https://doi.org/10.5194/hess-19-389-2015, 2015.

Bastos, A., Friedlingstein, P., Sitch, S., Chen, C., Mialon, A., Wigneron, J.-P., Arora, V. K., Briggs, P. R., Canadell, J. G., Ciais, P., Chevallier, F., Cheng, L., Delire, C., Haverd, V., Jain, A. K., Joos, F., Kato, E., Lienert, S., Lombardozzi, D., Melton, J. R., Myneni, R., Nabel, J. E. M. S., Pongratz, J., Poulter, B., Rödenbeck, C., Séférian, R., Tian, H., Van Eck, C., Viovy, N., Vuichard, N., Walker, A. P., Wiltshire, A., Yang, J., Zaehle, S., Zeng, N., and Zhu, D.: Impact of the 2015/2016 El Niño on the terrestrial carbon cycle constrained by bottom-up and top-down approaches, Philosophical Transactions of the Royal Society B: Biological Sciences, 373, 20170 304, https://doi.org/10.1098/rstb.2017.0304, 2018.

Blyth, E. M., Arora, V. K., Clark, D. B., Dadson, S. J., De Kauwe, M. G., Lawrence, D. M., Melton, J. R., Pongratz, J., Turton, R. H., Yoshimura, K., and Yuan, H.: Advances in Land Surface Modelling, Current Climate Change Reports, 7, 45–71, https://doi.org/10.1007/s40641-021-00171-5, 2021.

Boenisch, G.: Max Planck Institute for Biogeochemistry, Data Portal, https://www.bgc-jena.mpg.de/geodb/projects/Home.php, 2020.

Copernicus Climate Change Service (C3S): ERA5-Land hourly data from 1950 to present, https://doi.org/10.24381/cds.e2161bac, 2019.

Copernicus Climate Change Service (C3S): ERA5 monthly averaged data on single levels from 1940 to present, https://doi.org/10.24381/cds.f17050d7, 2023.

Dallmeyer, A., Kleinen, T., Claussen, M., Weitzel, N., Cao, X., and Herzschuh, U.: The deglacial forest conundrum, Nature Communications, 13, 6035, https://doi.org/10.1038/s41467-022-33646-6, 2022.

Danabasoglu, G.: IPCC DDC: NCAR CESM2 model output prepared for CMIP6 CMIP historical, https://doi.org/10.26050/WDCC/AR6.C6CMNRCES2hi, 2019.

Danabasoglu, G., Lamarque, J., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., Kampenhout, L. v., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019ms001916, 2020.

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

Findell, K. L., Berg, A., Gentine, P., Krasting, J. P., Lintner, B. R., Malyshev, S., Santanello, J. A., and Shevliakova, E.: The impact of anthropogenic land use and land cover change on regional climate extremes, Nature Communications, 8, 989, https://doi.org/10.1038/s41467-017-01038-w, 2017.

Flach, M., Brenning, A., Gans, F., Reichstein, M., Sippel, S., and Mahecha, M. D.: Vegetation modulates the impact of climate extremes on gross primary production, Biogeosciences, 18, 39–53, https://doi.org/10.5194/bg-18-39-2021, 2021.

Gentine, P., Green, J. K., Guérin, M., Humphrey, V., Seneviratne, S. I., Zhang, Y., and Zhou, S.: Coupling between the terrestrial carbon and water cycles—a review, Environmental Research Letters, 14, 083 003, https://doi.org/10.1088/1748-9326/ab22d6, 2019.

Gharun, M., Shekhar, A., Xiao, J., Li, X., and Buchmann, N.: Effect of the 2022 summer drought across forest types in Europe, Biogeosciences, 21, 5481–5494, https://doi.org/10.5194/bg-21-5481-2024, 2024.

Green, J. K., Seneviratne, S. I., Berg, A. M., Findell, K. L., Hagemann, S., Lawrence, D. M., and Gentine, P.: Large influence of soil moisture on long-term terrestrial carbon uptake, Nature, 565, 476–479, https://doi.org/10.1038/s41586-018-0848-x, 2019.

Groner, V. P., Raddatz, T., Reick, C. H., and Claussen, M.: Plant functional diversity affects climate–vegetation interaction, Biogeosciences, 15, 1947–1968, https://doi.org/10.5194/bg-15-1947-2018, 2018.

Guan, K., Pan, M., Li, H., Wolf, A., Wu, J., Medvigy, D., Caylor, K. K., Sheffield, J., Wood, E. F., Malhi, Y., Liang, M., Kimball, J. S., Saleska, S., Berry, J., Joiner, J., and Lyapustin, A. I.: Photosynthetic seasonality of global tropical forests constrained by hydroclimate, Nature Geoscience, 8, 284–289, https://doi.org/10.1038/ngeo2382, 2015.

Hao, Z., Hao, F., Xia, Y., Feng, S., Sun, C., Zhang, X., Fu, Y., Hao, Y., Zhang, Y., and Meng, Y.: Compound droughts and hot extremes: Characteristics, drivers, changes, and impacts, Earth-Science Reviews, 235, 104 241, https://doi.org/https://doi.org/10.1016/j.earscirev.2022.104241, 2022.

Haren, R. V., Haarsma, R. J., Vries, H. D., Oldenborgh, G. J. V., and Hazeleger, W.: Resolution dependence of circulation forced future central European summer drying, Environmental Research Letters, 10, 055 002, https://doi.org/10.1088/1748-9326/10/5/055002, 2015.

Harris, B. L., Taylor, C. M., Weedon, G. P., Talib, J., Dorigo, W., and Van Der Schalie, R.: Satellite-Observed Vegetation Responses to Intraseasonal Precipitation Variability, Geophysical Research Letters, 49, e2022GL099 635, https://doi.org/10.1029/2022GL099635, 2022.

Harris, B. L., Quaife, T., Taylor, C. M., and Harris, P. P.: Contrasting responses of vegetation productivity to intraseasonal rainfall in Earth system models, Earth System Dynamics, 15, 1019–1035, https://doi.org/10.5194/esd-15-1019-2024, 2024.

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

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,

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

D. P., and Zhang, X.: Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6, Geoscientific Model Development, 13, 5425–5464, https://doi.org/10.5194/gmd-13-5425-2020, 2020.

Iturbide, M., Gutiérrez, J. M., Alves, L. M., Bedia, J., Cerezo-Mota, R., Cimadevilla, E., Cofiño, A. S., Di Luca, A., Faria, S. H., Gorodetskaya, I. V., Hauser, M., Herrera, S., Hennessy, K., Hewitt, H. T., Jones, R. G., Krakovska, S., Manzanas, R., Martínez-Castro, D., Narisma, G. T., Nurhati, I. S., Pinto, I., Seneviratne, S. I., Van Den Hurk, B., and Vera, C. S.: An update of IPCC climate reference regions for subcontinental analysis of climate model data: definition and aggregated datasets, Earth System Science Data, 12, 2959–2970, https://doi.org/10.5194/essd-12-2959-2020, 2020.

Jain, P., Castellanos-Acuna, D., Coogan, S. C. P., Abatzoglou, J. T., and Flannigan, M. D.: Observed increases in extreme fire weather driven by atmospheric humidity and temperature, Nature Climate Change, 12, 63–70, https://doi.org/10.1038/s41558-021-01224-1, 2022.

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.

Kim, Y.-H., Min, S.-K., Zhang, X., Sillmann, J., and Sandstad, M.: Evaluation of the CMIP6 multi-model ensemble for climate extreme indices, Weather and Climate Extremes, 29, 100 269, https://doi.org/10.1016/j.wace.2020.100269, 2020.

Kolus, H. R., Huntzinger, D. N., Schwalm, C. R., Fisher, J. B., McKay, N., Fang, Y., Michalak, A. M., Schaefer, K., Wei, Y., Poulter, B., Mao, J., Parazoo, N. C., and Shi, X.: Land carbon models underestimate the severity and duration of drought's impact on plant productivity, Scientific Reports, 9, 2758, https://doi.org/10.1038/s41598-019-39373-1, 2019.

Kou-Giesbrecht, S., Reis Ely, C. R., Perakis, S. S., Cleveland, C. C., Menge, D. N. L., Reed, S. C., Taylor, B. N., Batterman, S. A., Crews, T. E., Dynarski, K. A., Gei, M., Gundale, M. J., Herridge, D. F., Jovan, S. E., Peoples, M. B., Piipponen, J., Rodríguez-Caballero, E., Salmon, V. G., Soper, F. M., Staccone, A. P., Weber, B., Williams, C. A., and Wurzburger, N.: Overestimated natural biological nitrogen fixation translates to an exaggerated CO_2 fertilization effect in Earth system models, Proceedings of the National Academy of Sciences, 122, e2514628 122, https://doi.org/10.1073/pnas.2514628122, 2025.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., Van Kampenhout, L., Kennedy, D., 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., Van Den 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, Journal of Advances in Modeling Earth Systems, 11, 4245–4287, https://doi.org/10.1029/2018MS001583, 2019.

Li, J., Bevacqua, E., Chen, C., Wang, Z., Chen, X., Myneni, R. B., Wu, X., Xu, C.-Y., Zhang, Z., and Zscheischler, J.: Regional asymmetry in the response of global vegetation growth to springtime compound climate events, Communications Earth & Environment, 3, 123, https://doi.org/10.1038/s43247-022-00455-0, 2022.

Li, J., Zhang, Y., Bevacqua, E., Zscheischler, J., Keenan, T. F., Lian, X., Zhou, S., Zhang, H., He, M., and Piao, S.: Future increase in compound soil drought-heat extremes exacerbated by vegetation greening, Nature Communications, 15, 10 875, https://doi.org/10.1038/s41467-024-55175-0, 2024.

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

Lorenz, R., Argüeso, D., Donat, M. G., Pitman, A. J., Van Den Hurk, B., Berg, A., Lawrence, D. M., Chéruy, F., Ducharne, A., Hagemann, S., Meier, A., Milly, P. C. D., and Seneviratne, S. I.: Influence of land-atmosphere feedbacks on temperature and precipitation extremes in the GLACE-CMIP5 ensemble, Journal of Geophysical Research: Atmospheres, 121, 607–623, https://doi.org/10.1002/2015JD024053, 2016.

Mahecha, M. D., Bastos, A., Bohn, F. J., Eisenhauer, N., Feilhauer, H., Hickler, T., Kalesse-Los, H., Migliavacca, M., Otto, F. E. L., Peng, J., Sippel, S., Tegen, I., Weigelt, A., Wendisch, M., Wirth, C., Al-Halbouni, D., Deneke, H., Doktor, D., Dunker, S., Duveiller, G., Ehrlich, A., Foth, A., García-García, A., Guerra, C. A., Guimarães-Steinicke, C., Hartmann, H., Henning, S., Herrmann, H., Hu, P., Ji, C., Kattenborn, T., Kolleck, N., Kretschmer, M., Kühn, I., Luttkus, M. L., Maahn, M., Mönks, M., Mora, K., Pöhlker, M., Reichstein, M., Rüger, N., Sánchez-Parra, B., Schäfer, M., Stratmann, F., Tesche, M., Wehner, B., Wieneke, S., Winkler, A. J., Wolf, S., Zaehle, S., Zscheischler, J., and Quaas, J.: Biodiversity and Climate Extremes: Known Interactions and Research Gaps, Earth's Future, 12, e2023EF003 963, https://doi.org/10.1029/2023EF003963, 2024.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenéz-de-la Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., M¨obis, B., M\textbackslash "uller, W. A., Nabel, J. E., Nam, C. C., Notz, D., Nyawira, S. S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J. u., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., Storch, J. S. v., Tian, F., Voigt, A., Vrese, P., Wieners, K. H., Wilkenskjeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2, Journal of Advances in Modeling Earth Systems, 11, 998–1038, https://doi.org/10.1029/2018ms001400, 2019.

McColl, K. A., Roderick, M. L., Berg, A., and Scheff, J.: The terrestrial water cycle in a warming world, Nature Climate Change, 12, 604–606, https://doi.org/10.1038/s41558-022-01412-7, 2022.

Muñoz Sabater, J.: ERA5-Land hourly data from 1950 to present, https://doi.org/10.24381/cds.e2161bac, 2019.

Norris, J., Hall, A., Neelin, J. D., Thackeray, C. W., and Chen, D.: Evaluation of the Tail of the Probability Distribution of Daily and Subdaily Precipitation in CMIP6 Models, Journal of Climate, 34, 2701–2721, https://doi.org/10.1175/JCLI-D-20-0182.1, 2021.

O'Donnell, A. J., Renton, M., Allen, K. J., and Grierson, P. F.: The role of extreme rain events in driving tree growth across a continental-scale climatic range in Australia, Ecography, 44, 1086–1097, https://doi.org/10.1111/ecog.05671, 2021.

Reichstein, M., Bahn, M., Ciais, P., Frank, D., Mahecha, M. D., Seneviratne, S. I., Zscheischler, J., Beer, C., Buchmann, N., Frank, D. C., Papale, D., Rammig, A., Smith, P., Thonicke, K., Van Der Velde, M., Vicca, S., Walz, A., and Wattenbach, M.: Climate extremes and the carbon cycle, Nature, 500, 287–295, https://doi.org/10.1038/nature12350, 2013.

Reick, C. H., Raddatz, T., Brovkin, V., and Gayler, V.: Representation of natural and anthropogenic land cover change in MPI-ESM, Journal of Advances in Modeling Earth Systems, 5, 459–482, https://doi.org/10.1002/jame.20022, 2013.

Ridder, N. N., Ukkola, A. M., Pitman, A. J., and Perkins-Kirkpatrick, S. E.: Increased occurrence of high impact compound events under climate change, npj Climate and Atmospheric Science, 5, 3, https://doi.org/10.1038/s41612-021-00224-4, 2022.

Seiler, C., Melton, J. R., Arora, V. K., Sitch, S., Friedlingstein, P., Anthoni, P., Goll, D., Jain, A. K., Joetzjer, E., Lienert, S., Lombardozzi, D., Luyssaert, S., Nabel, J. E. M. S., Tian, H., Vuichard, N., Walker, A. P., Yuan, W., and Zaehle, S.: Are Terrestrial Biosphere Models Fit for Simulating the Global Land Carbon Sink?, Journal of Advances in Modeling Earth Systems, 14, e2021MS002 946, https://doi.org/10.1029/2021MS002946, 2022.

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh, M., Vicente-Serrano, S. M., Wehner, M., and Zhou, B.: Weather and Climate Extreme Events in a Changing Climate, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., pp. 1513–1766, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/9781009157896.013, 2021.

Sharma, B., Kumar, J., Ganguly, A. R., and Hoffman, F. M.: Carbon cycle extremes accelerate weakening of the land carbon sink in the late 21st century, Biogeosciences, 20, 1829–1841, https://doi.org/10.5194/bg-20-1829-2023, 2023.

Simpson, I. R., Shaw, T. A., Ceppi, P., Clement, A. C., Fischer, E., Grise, K. M., Pendergrass, A. G., Screen, J. A., Wills, R. C. J., Woollings, T., Blackport, R., Kang, J. M., and Po-Chedley, S.: Confronting Earth System Model trends with observations, Science Advances, 11, eadt8035, https://doi.org/10.1126/sciadv.adt8035, 2025.

Tokarska, K. B., Stolpe, M. B., Sippel, S., Fischer, E. M., Smith, C. J., Lehner, F., and Knutti, R.: Past warming trend constrains future warming in CMIP6 models, SCIENCE ADVANCES, 2020.

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

Tschumi, E., Lienert, S., Bastos, A., Ciais, P., Gregor, K., Joos, F., Knauer, J., Papastefanou, P., Rammig, A., Van Der Wiel, K., Williams, K., Xu, Y., Zaehle, S., and Zscheischler, J.: Large Variability in Simulated Response of Vegetation Composition and Carbon Dynamics to Variations in Drought-Heat Occurrence, Journal of Geophysical Research: Biogeosciences, 128, e2022JG007 332, https://doi.org/10.1029/2022JG007332, 2023.

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, Nature Communications, 14, 6218, https://doi.org/10.1038/s41467-023-41851-0, 2023.

Villalobos-Herrera, R., Bevacqua, E., Ribeiro, A. F. S., Auld, G., Crocetti, L., Mircheva, B., Ha, M., Zscheischler, J., and De Michele, C.: Towards a compound-event-oriented climate model evaluation: a decomposition of the underlying biases in multivariate fire and heat stress hazards, Natural Hazards and Earth System Sciences, 21, 1867–1885, https://doi.org/10.5194/nhess-21-1867-2021, 2021.

Vogel, J., Paton, E., and Aich, V.: Seasonal ecosystem vulnerability to climatic anomalies in the Mediterranean, Biogeosciences, 18, 5903–5927, https://doi.org/10.5194/bg-18-5903-2021, 2021.

Wang, Z., Zhan, C., Ning, L., and Guo, H.: Evaluation of global terrestrial evapotranspiration in CMIP6 models, Theoretical and Applied Climatology, 143, 521–531, https://doi.org/10.1007/s00704-020-03437-4, 2021.

Wieners, K.-H., Giorgetta, M., Jungclaus, J., Reick, C., Esch, M., Bittner, M., Legutke, S., Schupfner, M., Wachsmann, F., Gayler, V., Haak, H., Vrese, P. d., Raddatz, T., Mauritsen, T., Storch, J.-S. v., Behrens, J., Brovkin, V., Claussen, M., Crueger, T., Fast, I., Fiedler, S., Hagemann, S., Hohenegger, C., Jahns, T., Kloster, S., Kinne, S., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Müller, W., Nabel, J., Notz, D., Gehlen, K. P.-v., Pincus, R., Pohlmann, H., Pongratz, J., Rast, S., Schmidt, H., Schnur, R., Schulzweida, U., Six, K., Stevens, B., Voigt, A., and Roeckner, E.: MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 CMIP historical, https://doi.org/10.22033/ESGF/CMIP6.6595, 2019.

https://doi.org/10.5194/egusphere-2026-626
Preprint. Discussion started: 30 March 2026

Wu, M., Manzoni, S., Vico, G., Bastos, A., De Vries, F. T., and Messori, G.: Drought Legacy in Sub-Seasonal Vegetation State and Sensitivity to Climate Over the Northern Hemisphere, Geophysical Research Letters, 49, e2022GL098 700, https://doi.org/10.1029/2022GL098700, 2022.

Wu, Y., Miao, C., Sun, Y., AghaKouchak, A., Shen, C., and Fan, X.: Global Observations and CMIP6 Simulations of Compound Extremes of Monthly Temperature and Precipitation, GeoHealth, 5, e2021GH000 390, https://doi.org/10.1029/2021GH000390, 2021.

Zhou, S., Zhang, Y., Park Williams, A., and Gentine, P.: Projected increases in intensity, frequency, and terrestrial carbon costs of compound drought and aridity events, Science Advances, 5, eaau5740, https://doi.org/10.1126/sciadv.aau5740, 2019.

Zscheischler, J. and Fischer, E. M.: The record-breaking compound hot and dry 2018 growing season in Germany, Weather and Climate Extremes, 29, 100 270, https://doi.org/10.1016/j.wace.2020.100270, 2020.

Zscheischler, J., Mahecha, M. D., Von Buttlar, J., Harmeling, S., Jung, M., Rammig, A., Randerson, J. T., Schölkopf, B., Seneviratne, S. I., Tomelleri, E., Zaehle, S., and Reichstein, M.: A few extreme events dominate global interannual variability in gross primary production, Environmental Research Letters, 9, 035 001, https://doi.org/10.1088/1748-9326/9/3/035001, 2014a.

Zscheischler, J., Michalak, A. M., Schwalm, C., Mahecha, M. D., Huntzinger, D. N., Reichstein, M., Berthier, G., Ciais, P., Cook, R. B., El-Masri, B., Huang, M., Ito, A., Jain, A., King, A., Lei, H., Lu, C., Mao, J., Peng, S., Poulter, B., Ricciuto, D., Shi, X., Tao, B., Tian, H., Viovy, N., Wang, W., Wei, Y., Yang, J., and Zeng, N.: Impact of large-scale climate extremes on biospheric carbon fluxes: An intercomparison based on MsTMIP data, Global Biogeochemical Cycles, 28, 585–600, https://doi.org/10.1002/2014GB004826, publisher: John Wiley & Sons, Ltd, 2014b.